

FACULTAD DE ESTUDIOS ESTADÍSTICOS

MÁSTER EN CIENCIA DE DATOS E INTELIGENCIA DE NEGOCIOS

Curso 2024/2025

Trabajo de Fin de Máster

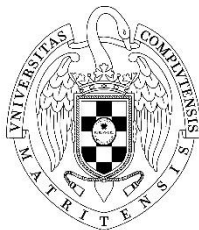
TÍTULO: *Priorización de proyectos de inversión pública
bajo restricciones presupuestarias:*

*Un enfoque híbrido basado en AHP y Aprendizaje
automático no supervisado*

Alumno: Diego Enrique Lassis González

Tutor: Javier Portela García – Miguel

Septiembre 09 de 2025



UNIVERSIDAD COMPLUTENSE
MADRID



Declaración Responsable sobre Autoría y Uso Ético de Herramientas de Inteligencia Artificial (IA)

Yo, Lassis González-Enrique, Diego

Con DNI/NIE/PASAPORTE: RD7498190

Declaro de manera responsable que el/la presente:

- Trabajo de Fin de Grado (TFG)
- Trabajo de Fin de Máster (TFM)
- Tesis Doctoral

Titulado/a

Priorización de proyectos de inversión pública bajo restricciones presupuestarias:

Un enfoque híbrido basado en AHP, regresión supervisada y LTR

es el resultado de mi trabajo intelectual personal y creativo, y ha sido elaborado de acuerdo con los principios éticos y las normas de integridad vigentes en la comunidad académica y, más específicamente, en la Universidad Complutense de Madrid.

Soy pues, autor del material aquí incluido y, cuando no ha sido así y he tomado el material de otra fuente, lo he citado o bien he declarado su procedencia de forma clara -incluidas, en su caso, herramientas de inteligencia artificial-. Las ideas y aportaciones principales incluidas en este trabajo, y que acreditan la adquisición de competencias, son mías y no proceden de otras fuentes o han sido reescritas usando material de otras fuentes.

Asimismo, aseguro que los datos y recursos utilizados son legítimos, verificables y han sido obtenidos de fuentes confiables y autorizadas. Además, he tomado medidas para garantizar la confidencialidad y privacidad de los datos utilizados, evitando cualquier tipo de sesgo o discriminación injusta en el tratamiento de la información.

En Madrid a 09 días de septiembre de 2025



FIRMA

Resumen

Este Trabajo de Fin de Máster propone y valida un modelo híbrido para la priorización de proyectos del Plan Nacional Plurianual de Inversión Pública (PNPIP) de la República Dominicana. El enfoque integra: (i) un puntaje normativo calculado con AHP según el marco SNIP/PNPIP; (ii) aprendizaje no supervisado (K-Means, Ward, GMM y HDBSCAN) para identificar tipologías empíricas de proyectos; y (iii) una optimización lineal entera mixta (MILP) bajo techo presupuestario que articula reglas de diversidad y cuotas. A partir de la evidencia, se adoptan siete clústeres Ward como capa diagnóstica, consolidados en tres estratos (AHP_bajo/medio/alto) para la toma de decisiones. Se comparan cinco estrategias: S0 (AHP puro), S1 (cap institucional), S2 (cuotas por estrato), S3 (cuotas por clúster) y S4 (premios de cobertura institucional/territorial en la función objetivo). La evaluación combina eficiencia (\sum score AHP), concentración (HHI institucional y provincial) y coberturas. Los resultados muestran que, en techos intermedios, S2–S4 reducen concentración y amplían coberturas con pérdidas de eficiencia marginales ($\leq 2-3\%$, ver Anexo 4) respecto a S0; S1 es útil para contener acaparamientos, pero requiere balancearse para no trasladar concentración al plano territorial. El modelo no sustituye al AHP: lo complementa y hace gobernable, transparentando el precio de la diversidad y ofreciendo carteras óptimas condicionadas por reglas explícitas.

Palabras clave: AHP; clustering; Ward; MILP; optimización con restricción presupuestaria; inversión pública; diversidad institucional y territorial; República Dominicana; Aprendizaje automático no supervisado.

Abstract

This thesis develops and validates a hybrid model for prioritizing public investment projects in the Dominican Republic's PNPIP. The approach integrates: (i) a normative score computed with AHP (aligned with the national SNIP/PNPIP framework); (ii) unsupervised learning (K-Means, Ward, GMM, and HDBSCAN) to uncover empirical project typologies; and (iii) a Mixed-Integer Linear Programming (MILP) model under a budget cap that embeds diversity rules and quotas. Evidence supports seven Ward clusters as a diagnostic layer, consolidated into three strata (low/medium/high AHP) for decision-making. Five strategies are compared: S0 (pure AHP), S1 (institutional cap), S2 (strata quotas), S3 (cluster quotas), and S4 (institutional/territorial coverage incentives in the objective). Performance is assessed through efficiency (\sum AHP score), concentration (institutional and provincial HHI), and coverage. Results indicate that, at intermediate budgets, S2–S4 decrease concentration and expand coverage with marginal efficiency losses ($\leq 2-3\%$, see Annex 4) relative to S0; S1 curbs institutional dominance but should be combined with other rules to avoid territorial concentration. The model does not replace AHP: it complements and operationalizes it, making the price of diversity explicit and delivering optimal portfolios conditioned by transparent rules.

Keywords: AHP; clustering; Ward; MILP; budget-constrained optimization; public investment; institutional and territorial diversity; Dominican Republic; unsupervised machine learning.

Contenido

1. Introducción al problema y objetivos del estudio	1
1.1. Contextualización del problema	1
1.2. Marco institucional y normativo.....	1
1.3. Planteamiento del problema	2
1.4. Justificación del estudio.....	3
1.4.1. Perspectiva académica y metodológica.....	3
1.4.2. Perspectiva institucional.....	4
1.5. Objetivos.....	4
1.5.1. Objetivo general	4
1.5.2. Objetivos específicos	4
1.6. Estructura del documento	5
2. Marco teórico	6
2.1. Planificación pública e inversión en el contexto de priorización.....	6
2.2. Métodos de decisión multicriterio (MCDM / AHP)	7
2.2.1. Fundamentos teóricos del AHP	7
2.2.2. Propiedades, supuestos y limitaciones.....	8
2.2.3. Evidencia aplicada en inversión pública y portafolios	8
2.3. Diseño del score AHP para el PNPIP	8
2.3.1. Definición de criterios C1–C7 y alineación normativa.....	8
2.3.2. Jerarquía, comparaciones y cálculo de pesos.....	9
2.3.3. Indicadores, escalas y normalización	9
2.3.4. Integración con el flujo SNIP–PNPIP	9
2.4. Aprendizaje automático en la priorización de proyectos	9
2.4.1. Enfoques supervisados: alcances y limitaciones	10
2.4.2. Aprendizaje no supervisado: fundamentos y técnicas de clustering	10
2.4.3. Validación y calidad de los agrupamientos.....	12
2.4.4. Interpretabilidad y transparencia en la inversión pública	13
2.5. Integración y evaluación del enfoque híbrido	13
2.5.1. Integración normativo–empírico	13
2.5.2. Evaluación normativa	13
2.5.3. Evaluación técnica	14
3. Metodología general	14
3.1. Etapa 1: Preparación de datos.....	14

3.1.1.	Fuentes de datos utilizadas	14
3.1.2.	Procesos de limpieza e imputación	15
3.1.3.	Ingeniería de características	16
3.1.4.	Variables seleccionadas	17
3.1.5.	Análisis descriptivo	17
3.2.	Etapa 2: Construcción del score normativo (AHP).....	18
3.2.1.	Definición de criterios.....	18
3.2.2.	Matriz de comparación de pares.....	19
3.2.3.	Cálculo de pesos normalizados	21
3.2.4.	Validación del Índice de consistencia	21
3.2.5.	Asignación del score a cada proyecto.....	21
3.3.	Etapa 3: Modelado no supervisado.....	22
3.3.1.	Preprocesamiento de datos para modelado	22
3.3.2.	Clustering con K-Means	23
3.3.3.	Clustering Jerárquico (Ward)	24
3.3.4.	Mezclas Gaussianas (GMM).....	25
3.3.5.	DBSCAN/HDBSCAN (detección de outliers).....	26
3.3.6.	Selección del modelo de clusters	27
3.4.	Etapa 4: Integración AHP – ML	29
3.4.1.	Comparación de clusters con score AHP	29
3.4.2.	Identificación de tipologías de proyectos.....	30
3.4.3.	Interpretabilidad de clusters (radar plots, perfiles socioeconómicos) ...	30
3.4.4.	Diversidad institucional y territorial de clusters.....	30
3.4.5.	Preparación de datos para escenarios de priorización.....	31
3.5.	Etapa 5: Diseño experimental de optimización con restricción presupuestaria	31
3.5.1.	Lógica del experimento y enfoque de análisis.....	31
3.5.2.	Datos operativos y notación	31
3.5.3.	De la motivación institucional a la formulación del problema.....	32
3.5.4.	Parametrización experimental (dos escenarios)	34
3.5.5.	Procedimiento computacional (resolución y reporte)	34
3.5.6.	Marco de evaluación (métricas y criterios de lectura)	35
3.5.7.	Expectativas ex ante y pertinencia para la DGIP	35
4.	Resultados.....	35

4.1.	Etapa 2: AHP	36
4.2.	Etapa 3: Modelado no supervisado	37
4.2.1.	Clustering con K-Means	38
4.2.2.	Clustering jerárquico (Ward)	39
4.2.3.	Mezclas Gaussianas (GMM).....	40
4.2.4.	Clustering DBSCAN/HDBSCAN	41
4.2.5.	Selección del modelo óptimo	43
4.3.	Etapa 4: Integración AHP - ML.....	45
4.3.1.	Comparación de estratos (7 y 3).....	45
4.3.2.	Interpretabilidad.....	46
4.4.	Etapa 5: Diseño experimental de optimización con restricción presupuestaria 48	
4.5.	Discusión.....	49
5.	Conclusiones y líneas futuras de investigación	50
5.1.	Conclusiones técnicas.....	50
5.2.	Conclusiones institucionales.....	51
5.3.	Limitaciones del estudio	51
5.4.	Recomendaciones de política pública y de investigación futura.....	52
	Bibliografía	53
	Anexos	57

Índice de figuras

Figura 1. Esquema jerárquico de criterios	19
Figura 2. Detección de outliers del Score AHP (%)	36
Figura 3. Distribución del score AHP (%)	36
Figura 4. Heatmap de solapamiento KDE por estratos	37
Figura 5. Curvas de evaluación k dominante (K-Means)	38
Figura 6. Curvas de evaluación k dominante (Ward)	39
Figura 7. Dendograma (Ward)	39
Figura 8. Curvas de BIC/AIC por k (GMM)	41
Figura 9. Distribución de clusters (HDBSCAN)	42
Figura 10. Pruebas de estrés k dominante (HDBSCAN)	42
Figura 11. Configuraciones de parámetros en HDBSCAN	42
Figura 12. Boxplot de distribución de estratos (7 y 3)	45
Figura 13. Identificación de tipologías de estratos con radar plot (7 estratos)	47
Figura 14. Identificación de tipologías de estratos con radar plot (3 estratos)	47
Figura 15. Score AHP acumulado vs presupuesto (ajustes fuertes)	48
Figura 16. HHI provincial vs presupuesto (ajustes fuertes)	48
Figura 17. HHI por estrato (Institución y provincia)	49

Índice de tablas

Tabla 1. Comparativa de modelos de machine learning no supervisados	11
Tabla 2. Métricas de validación en clustering	14
Tabla 3. Diccionario de variables	14
Tabla 4. Escala de Saaty de importancia relativa	19
Tabla 5. Comparación de criterios	20
Tabla 6. Resultados de pesos AHP	36
Tabla 7. Pruebas de separación de clusters/estratos	37
Tabla 8. Pruebas de estrés k dominante (K-Means)	38
Tabla 9. Pruebas de estrés k dominante (Ward)	40
Tabla 10. Pruebas de estrés k dominante (GMM)	41
Tabla 11. Comparación base modelo óptimo (Ranking básico)	43
Tabla 12. Comparación de modelo óptimo (Ranking policy-aware)	44
Tabla 13. Robustez de selección de modelo óptimo (Sensibilidad de pesos/umbrales)	45
Tabla 14. Resumen del score AHP para 7 clusters	46
Tabla 15. Resumen del score AHP para 3 estratos	46

Capítulo 1

1. Introducción al problema y objetivos del estudio

1.1. Contextualización del problema

La inversión pública desempeña un papel fundamental en el desarrollo económico y social de los países, pues constituye el mecanismo mediante el cual el Estado traduce los planes de desarrollo en bienes, servicios e infraestructuras que generan valor público. En la República Dominicana, este proceso está enmarcado en un conjunto normativo e institucional que busca garantizar que la asignación de recursos responda a los principios de eficiencia, equidad, transparencia y sostenibilidad (Ley 498-06, 2006).

No obstante, la realidad muestra tensiones estructurales. La cartera de proyectos de inversión que se formula en cada ciclo supera sistemáticamente la disponibilidad de recursos fiscales. Esta brecha entre demanda y presupuesto genera la necesidad de definir mecanismos claros de priorización que permitan seleccionar, entre múltiples propuestas, aquellas que más contribuyan a los objetivos de desarrollo nacional. En este sentido, el Decreto 493-07 establece que, frente a la insuficiencia de recursos, deben priorizarse los proyectos de arrastre sobre los nuevos, garantizando la continuidad de las inversiones ya iniciadas. Sin embargo, la normativa no precisa cómo operacionalizar esta regla en términos metodológicos, lo que abre espacio a prácticas heterogéneas y a veces discrecionales entre instituciones ejecutoras.

La Estrategia Nacional de Desarrollo 2030 (END 2030) refuerza esta exigencia, al establecer que las inversiones deben contribuir al cumplimiento de sus ejes estratégicos y metas de largo plazo. No obstante, la ausencia de criterios homogéneos para evaluar y priorizar proyectos dificulta asegurar una alineación sistemática de la cartera de inversión con la visión de desarrollo definida. Como señala Boix Cots (2023), la falta de metodologías científicas de apoyo a la decisión en la administración pública genera “*proyectos innecesarios, sobrecostos y retrasos, con impactos sociales y ambientales negativos*” (p. 2).

En este contexto, la problemática central que aborda este trabajo es la necesidad de diseñar un modelo de priorización de proyectos de inversión pública que combine el marco normativo nacional con metodologías multicriterio y herramientas de inteligencia artificial, para superar las limitaciones de los enfoques tradicionales y garantizar una asignación más eficiente y transparente de los recursos.

1.2. Marco institucional y normativo

El sistema de planificación e inversión pública dominicano está soportado por un entramado legal que constituye el fundamento de la presente investigación.

En primer lugar, la Ley 498-06 de Planificación e Inversión Pública crea el Sistema Nacional de Planificación e Inversión Pública (SNIP) y establece que toda inversión debe ser evaluada, priorizada y monitoreada en coherencia con los planes nacionales, sectoriales y territoriales. El artículo 4 establece que el sistema busca “*asegurar la*

calidad del gasto público mediante la identificación, formulación, evaluación, priorización y seguimiento de los proyectos de inversión” (Ley 498-06, 2006, Art. 4).

El Decreto 493-07, que aprueba su reglamento de aplicación, define la operatividad del sistema y las funciones de las Unidades Institucionales de Planificación y Desarrollo (UIP), encargadas de formular y evaluar proyectos en cada organismo del Estado. Este reglamento dispone expresamente que, en escenarios de restricción presupuestaria, la inclusión de nuevos proyectos queda supeditada a la continuidad de los proyectos de arrastre (Decreto 493-07, 2007, Art. 11).

La Ley 1-12 que aprueba la Estrategia Nacional de Desarrollo 2030 (2011) constituye el marco de referencia obligatorio para la política pública, al establecer una visión de largo plazo con cuatro ejes estratégicos: fortalecimiento institucional, cohesión social, competitividad económica y sostenibilidad ambiental. El artículo 6 establece que los planes y presupuestos deben estar alineados con sus objetivos, convirtiéndola en un instrumento vinculante para la priorización de inversiones.

Finalmente, el documento del Sistema Nacional de Planificación (2007) organiza la institucionalidad de la planificación en niveles nacional, regional y municipal, y articula los planes de desarrollo con los presupuestos, disponiendo que las inversiones se integren en el Plan Nacional Plurianual del Sector Público (PNPSP) y, específicamente, en el Plan Nacional Plurianual de Inversión Pública (PNPIP).

1.3. Planteamiento del problema

En la República Dominicana, la Dirección General de Inversión Pública enfrenta cada año el desafío de evaluar y priorizar un conjunto amplio de proyectos sometidos por instituciones ejecutoras y gobiernos locales para su inclusión en el PNPIP. La brecha entre el volumen de proyectos propuestos y la disponibilidad presupuestaria obliga a descartar un número significativo de iniciativas, aun cuando muchas de ellas respondan a demandas legítimas de la ciudadanía.

La normativa vigente establece lineamientos generales de priorización —como la continuidad de proyectos de arrastre frente a los nuevos—, pero carece de un marco metodológico estandarizado para evaluar de manera comparativa la contribución de cada proyecto a los objetivos de desarrollo. En la práctica, la DGIP ha implementado un esquema de priorización basado en el Analytic Hierarchy Process (AHP). No obstante, este modelo presenta una limitación fundamental: aunque se definieron criterios normativos para evaluar los proyectos, no se desarrolló un sistema de ponderación, de manera que todos los criterios se trataron con el mismo peso relativo.

Actualmente, el proceso de priorización se apoya en herramientas institucionales que, aunque normativamente definidas, presentan limitaciones metodológicas. No se cuenta con un modelo automatizado que permita ponderar múltiples dimensiones de forma simultánea, ni con mecanismos de optimización que permitan ordenar los proyectos de forma dinámica bajo restricciones reales, como el techo presupuestario. Este diseño incompleto genera varios problemas:

- **Menor objetividad en la priorización**, ya que no se distinguen adecuadamente los criterios estratégicos de aquellos de menor relevancia.

- **Homogeneización de proyectos distintos**, pues los criterios con impacto desigual sobre el desarrollo se consideran equivalentes.
- **Desalineación con el PNPIP y los lineamientos de política pública**, al no reflejar prioridades diferenciales como la equidad territorial, la sostenibilidad financiera o la alineación estratégica.
- **Riesgo de ineficiencia en la asignación presupuestaria**, al no discriminar con suficiente precisión entre proyectos de alto impacto y proyectos de impacto marginal.

En síntesis, aunque el uso del AHP ha representado un avance hacia la estandarización de la priorización, la ausencia de un sistema de ponderaciones lo convierte en un mecanismo normativo insuficiente. Asimismo, este enfoque puramente normativo no aprovecha el potencial de los datos históricos ni permite identificar patrones empíricos o agrupamientos naturales de proyectos que podrían complementar el análisis y mejorar la toma de decisiones.

1.4. Justificación del estudio

1.4.1. Perspectiva académica y metodológica

La priorización de proyectos de inversión es un problema de asignación bajo restricciones presupuestarias. Tradicionalmente, el análisis costo-beneficio ha sido el instrumento más difundido para evaluar proyectos, dado su énfasis en la eficiencia económica. Sin embargo, diversos autores han señalado que este enfoque es limitado, pues *“tiende a privilegiar únicamente dimensiones monetarias y a excluir criterios sociales, territoriales y ambientales”* (Lobo Flores, 2024, p. 3, citando a Ramírez, 2007, y Arancibia, 2016).

Frente a ello, la literatura sobre métodos de decisión multicriterio (MCDM) ha propuesto herramientas que permiten integrar criterios múltiples y heterogéneos en la toma de decisiones. Entre estas, el Proceso de Análisis Jerárquico (AHP), desarrollado por Saaty en la década de 1970, se ha consolidado como una de las metodologías más aplicadas en contextos de inversión pública, al posibilitar la descomposición de decisiones complejas en estructuras jerárquicas y la asignación de pesos relativos mediante comparaciones por pares.

La evidencia empírica respalda su utilidad Ng Corrales et al. (2024) demostraron en un estudio en la Comarca Ngäbe-Buglé (Panamá) que la aplicación de AHP en proyectos de inversión social *“facilita la clasificación, ordenamiento y priorización de alternativas complejas”*. A nivel internacional, el modelo SEESIM, desarrollado en Italia, integra criterios de sostenibilidad social, económica y ambiental bajo la metodología AHP para orientar la ejecución de planes nacionales de recuperación y resiliencia (Baffo et al., 2024). En paralelo, el Infrastructure Prioritization Framework (IPF) del Banco Mundial constituye un instrumento pragmático para países en desarrollo, que combina criterios sociales, ambientales y económicos en un marco multicriterio flexible y transparente (Darwin Marcelo et al., 2016). En España, Boix Cots (2023) desarrolló el Sistema de Evaluación y Priorización de Inversiones Públicas (SEPIP), que incorpora además gestión de incertidumbre y toma de decisiones en grupo.

Esta literatura sustenta que los modelos de decisión multicriterio no solo aportan objetividad y transparencia para el abordaje de este tipo de problemáticas, sino que también son capaces de aislar los sesgos heurísticos que determinan los contextos estructurales políticos.

1.4.2. Perspectiva institucional

El desarrollo de un modelo híbrido que combine técnicas normativas y analíticas se considera fundamental para superar las limitaciones actuales en el proceso de priorización de inversiones públicas en la República Dominicana. Desde la práctica institucional en la Dirección General de Inversión Pública se ha constatado que, si bien el AHP con un sistema de ponderación correctamente estimado aporta un marco transparente y objetivo de calificación normativa, presenta limitaciones para capturar la diversidad territorial e institucional de la cartera de proyectos.

En este sentido, la integración de técnicas de aprendizaje automático no supervisado, como el clustering, permite identificar tipologías de proyectos, revelar agrupamientos empíricos y analizar patrones de formulación. Este enfoque complementario ofrece varias ventajas:

- **Corregir la debilidad del AHP institucional actual**, al incorporar ponderaciones normativas consistentes con la estrategia de desarrollo.
- **Aumentar la transparencia del proceso**, al mostrar cómo se agrupan los proyectos según sus características.
- **Incorporar criterios de diversidad** institucional y territorial más allá de la calificación normativa.
- **Detectar sesgos en la formulación de proyectos**, al contrastar clusters empíricos con el ranking AHP.
- **Explorar escenarios alternativos** de priorización que integren el enfoque normativo con la evidencia empírica.

De esta manera, el modelo propuesto no sustituye el marco normativo vigente, sino que lo fortalece y lo complementa, generando mayor objetividad en la priorización y garantizando un uso más eficiente y equitativo del presupuesto disponible.

1.5. Objetivos

1.5.1. Objetivo general

Diseñar y validar un modelo híbrido de priorización de proyectos de inversión pública, que integre métodos de decisión multicriterio (AHP) y técnicas de aprendizaje automático no supervisado orientados a la optimización del ranking bajo restricciones presupuestarias reales, en el marco normativo e institucional del Sistema Nacional de Inversión Pública de la República Dominicana.

1.5.2. Objetivos específicos

- i. Construir score de priorización partiendo de un esquema de criterios y ponderaciones mediante el Proceso de Análisis Jerárquico (AHP), utilizando criterios normativos definidos por el Comité de Priorización del MEPyD.

- ii. Implementar técnicas de aprendizaje automático no supervisado (clustering) para identificar tipologías de proyectos, patrones de formulación y agrupamientos empíricos en la cartera del PNPIP.
- iii. Analizar la diversidad institucional y territorial de los clusters obtenidos y su relación con los resultados del AHP.
- iv. Desarrollar un diseño experimental de optimización con restricciones presupuestarias, mediante la integración de los resultados del AHP y el clustering en un modelo híbrido.
- v. Evaluar el desempeño técnico, interpretabilidad y aplicabilidad institucional del sistema de priorización resultante.

1.6. Estructura del documento

El presente trabajo se organiza en cinco capítulos. En el Capítulo 1 se introduce la investigación, presentando la contextualización del problema, el marco institucional y normativo en forma de síntesis estratégica, la justificación académica e institucional y los objetivos del estudio.

En el Capítulo 2 se desarrolla el marco teórico, el cual se abre con una sección de fundamentación normativa que profundiza en el Sistema Nacional de Planificación e Inversión Pública (SNIP), la Ley 498-06, su reglamento de aplicación (Decreto 493-07) y la Estrategia Nacional de Desarrollo 2030 (Ley 1-12), para luego abordar los fundamentos conceptuales de la planificación pública, los métodos multicriterio de decisión —con énfasis en el Proceso de Análisis Jerárquico (AHP)— y las técnicas de aprendizaje automático no supervisado aplicables al problema de priorización de proyectos.

En el Capítulo 3 se presenta la metodología de investigación, detallando las etapas de preparación y depuración de los datos, la construcción del score AHP con ponderaciones derivadas de la normativa, la aplicación de algoritmos de clustering, la integración de resultados en un modelo híbrido y el diseño experimental de validación.

El Capítulo 4 expone los resultados obtenidos, complementados con un análisis crítico y discusión en relación con la literatura revisada y el marco institucional.

Finalmente, en el Capítulo 5 se formulan las conclusiones, se destacan las contribuciones académicas e institucionales del estudio, se señalan las limitaciones encontradas y se proponen recomendaciones tanto para futuras investigaciones como para la aplicación práctica del modelo en la Dirección General de Inversión Pública y en el marco del Sistema Nacional de Inversión Pública.

Capítulo 2

2. Marco teórico

2.1. Planificación pública e inversión en el contexto de priorización

En entornos de limitación presupuestaria, la asignación eficiente del gasto de capital constituye una prioridad para los gobiernos que buscan maximizar el impacto del desarrollo económico y social. En este marco, la planificación pública adquiere una dimensión estratégica, orientada no solo a definir objetivos y metas, sino también a establecer mecanismos institucionales que permitan priorizar la inversión pública en función de criterios técnicos, territoriales y sectoriales.

La República Dominicana aborda esta necesidad mediante la creación del Sistema Nacional de Planificación e Inversión Pública (SNPIP), formalizado en la Ley No. 498-06, el cual articula *“el conjunto de principios, normas, órganos y procesos a través de los cuales se fijan las políticas, objetivos, metas y prioridades del desarrollo económico y social evaluando su cumplimiento”* (Ley 498-06, 2006, Art. 1). Este sistema se concibe como la columna vertebral del modelo de gestión pública orientado al desarrollo, estableciendo las condiciones necesarias para que los recursos públicos se asignen conforme a un plan previamente concertado y estructurado.

Dentro de este marco, el instrumento operativo central es *“...el Plan Nacional Plurianual de Inversión Pública, que forma parte del Plan Nacional Plurianual del Sector Público (PNPIP)”*, definido en la Sección III del Reglamento de Aplicación (Decreto 493-07, 2007, Art.s 36-38). El PNPIP organiza los proyectos de inversión a ser ejecutados por las entidades públicas, ordenándolos de acuerdo con criterios de prioridad definidos institucionalmente.

El diseño de este plan está regulado por el Reglamento de Aplicación No. 1, instituido mediante el Decreto No. 493-07, que introduce un conjunto de mecanismos técnico-normativos para garantizar la consistencia, viabilidad y sostenibilidad de la inversión pública. En particular, el Artículo 77 establece que únicamente podrán incorporarse al PNPIP aquellos proyectos que cuenten con un código SNIP y un dictamen de aceptación para cada fase de su ciclo de vida, exigencia que procura asegurar la madurez técnica y financiera de las propuestas. Para tal fin, se crea un Comité de Priorización, presidido por el Viceministro de Planificación e integrado por las direcciones generales de Inversión Pública, Desarrollo Económico y Social, y Ordenamiento Territorial, quienes formulan propuestas de jerarquización conforme a criterios técnicos y de política pública (Decreto 493-07, 2007, Art. 77).

Este modelo de priorización es reforzado por el Artículo 78, que establece que el orden de prelación debe respetar la estructura integral del proyecto, sus costos, componentes y cronograma, evitando su fragmentación artificial. A su vez, el artículo 79 vincula este proceso con la programación presupuestaria, al establecer que la actualización anual del PNPIP debe ser presentada antes del 31 de mayo de cada año, y que la asignación de

recursos debe otorgar prioridad a los proyectos en ejecución (arrastre) sobre los nuevos. Esta estructura está diseñada para maximizar la eficiencia en el uso de recursos, dando continuidad a compromisos contractuales y evitando la dispersión del gasto.

El Capítulo V del reglamento consolida el proceso al vincular la planificación plurianual con los instrumentos presupuestarios, garantizando coherencia entre lo programado y lo ejecutado. Así, la planificación y la inversión pública quedan integradas en un único flujo operativo, normado y jerarquizado, que constituye el modelo institucional vigente de priorización.

No obstante, a pesar de la solidez técnica del marco legal, la realidad presupuestaria impone desafíos adicionales, entre ellos, la necesidad de evaluar grandes volúmenes de proyectos con rapidez, objetividad y transparencia. En este contexto, el desarrollo de modelos de inteligencia artificial para apoyar el proceso de priorización se plantea como un complemento a los instrumentos existentes. El objetivo no es sustituir el marco normativo, sino aportar herramientas de analítica avanzada que permitan optimizar las decisiones bajo restricción presupuestaria, respetando los criterios y etapas ya estipulados en la normativa vigente.

2.2. Métodos de decisión multicriterio (MCDM / AHP)

Los métodos de decisión multicriterio (MCDM) se definen como un conjunto de herramientas que permiten evaluar alternativas considerando múltiples criterios heterogéneos, tanto cuantitativos como cualitativos. Su aplicación resulta especialmente útil en contextos públicos, donde las decisiones de inversión no pueden reducirse únicamente a indicadores económicos, sino que deben integrar criterios sociales, territoriales y estratégicos.

2.2.1. Fundamentos teóricos del AHP

Los métodos de decisión multicriterio (MCDM) se definen como un conjunto de herramientas que permiten evaluar alternativas considerando múltiples criterios heterogéneos, tanto cuantitativos como cualitativos. Su aplicación resulta especialmente útil en contextos públicos, donde las decisiones de inversión no pueden reducirse únicamente a indicadores económicos, sino que deben integrar criterios sociales, territoriales y estratégicos.

Formalmente, se construye una matriz de comparaciones por pares $A = [a_{ij}]$, donde cada elemento $a_{ji} = 1/a_{ij}$ y $a_{ii} = 1$. El vector de prioridades w se obtiene como el autovector principal de A , normalizado a $\sum w_i = 1$.

En la formulación estándar, el vector de pesos w satisface:

$$A w = \lambda_{max} w, \text{ con } \sum_{i=1}^n w_i = 1$$

La consistencia de los juicios se verifica mediante el Índice de Consistencia (CI) y la Razón de Consistencia (CR), definidos como:

$$CI = \frac{\lambda_{max} - n}{n - 1}, CR = \frac{CI}{RI}$$

Donde λ_{max} es el valor propio máximo de A y n el número de criterios. Este valor se contrasta con Índice Aleatorio RI , y está asociado al tamaño n de la matriz. Se considera aceptable un nivel de CR inferior a los umbrales propuestos por Saaty, garantizando la coherencia interna de las comparaciones.

2.2.2. Propiedades, supuestos y limitaciones

El AHP se apoya en supuestos de: (i) comparabilidad entre criterios y alternativas en la jerarquía; (ii) reciprocidad y transitividad aproximada de los juicios; y (iii) independencia entre criterios al momento de comparar. Su verificación de consistencia reduce el riesgo de contradicción en juicios; no obstante, cuando la consistencia es baja, se requiere revisar y ajustar comparaciones. Entre las limitaciones se señalan la sensibilidad a inconsistencias y la necesidad de una definición clara de criterios para evitar solapamientos. Estos aspectos se abordan con procedimientos de control de calidad (revisión de CR , pruebas de sensibilidad de pesos) y protocolos de documentación metodológica (Saaty, 1987).

2.2.3. Evidencia aplicada en inversión pública y portafolios

El AHP ha sido utilizado en múltiples contextos de política pública. En la selección de contratos de obras públicas, se ha demostrado su utilidad para integrar criterios técnicos, económicos y de riesgo en un marco de decisión trazable y replicable (Bertolini et al., 2006). También se ha empleado en la priorización de proyectos de inversión social en la Comarca Ngäbe-Buglé (Panamá), donde permitió clasificar y ordenar iniciativas en función de su contribución al desarrollo territorial (Corrales et al., 2024).

Asimismo, se ha aplicado en la construcción de modelos multicriterio más avanzados, como el SEESIM, orientado a evaluar sostenibilidad social, económica y ambiental en planes nacionales (Baffo et al., 2024), o el Infrastructure Prioritization Framework (IPF) desarrollado por el Banco Mundial (Darwin Marcelo et al., 2016). En gestión de portafolios, también se detalla un procedimiento práctico que incluye definición del objetivo, estructuración jerárquica, generación de matrices de comparación, cálculo de pesos y verificación de consistencia, con énfasis en la trazabilidad de cada juicio y la comunicación de resultados (Ricardo Viana Vargas, 2010).

2.3. Diseño del score AHP para el PNPIP

2.3.1. Definición de criterios C1–C7 y alineación normativa

En el proceso de priorización del Plan Nacional Plurianual de Inversión Pública (PNPIP) se emplea un conjunto de criterios institucionales estandarizados, identificados como C1–C7, definidos por el Comité de Priorización del MEPyD en el marco del Sistema Nacional de Inversión Pública (SNIP), que aborda el PNPIP 2025-2028 publicado por la DGIP. Dichos criterios responden a los principios establecidos en la Ley 498-06 y en su reglamento de aplicación (Decreto 493-07), que exige registro SNIP, dictámenes por etapa y priorización anual con prelación de proyectos de continuidad/arrastre sobre nuevos, conservando la integridad de costos, componentes y cronograma (arts. 77–79 del Reglamento; arts. 36–38 de la Ley).

En términos conceptuales, el diseño del score responde a tres dimensiones: (i) conformidad regulatoria (madurez del ciclo de vida y documentación técnica), (ii)

coherencia estratégica (alineación con la END y planes sectoriales/institucionales), y (iii) trazabilidad institucional (incorporación efectiva en la programación plurianual del PNPS) (art. 36 del Reglamento).

El detalle operativo de cada criterio (definiciones, indicadores, escalas de valoración) se presenta en el Capítulo 3 (Metodología), donde se documenta su aplicación en la construcción del score AHP.

2.3.2. Jerarquía, comparaciones y cálculo de pesos

El score AHP para el PNPIP se estructura en tres niveles jerárquicos: objetivo general (priorizar proyectos bajo restricción presupuestaria), criterios C1–C7 y alternativas (proyectos de inversión). Se construye la matriz de comparaciones A de tamaño $n = 7$ siguiendo la escala fundamental de Saaty y se deriva el vector de prioridades w por el método del autovector principal, normalizando a $\sum_i w_i = 1$.

La consistencia se verifica con CI y CR , y se recalibran entradas si CR excede el umbral aceptable (Saaty, 1987). Cada elemento a_{ij} de la matriz A se define en función del orden de prelación que establece el marco legal e institucional de la planificación e inversión pública (Ley 498-06, Decreto 493-07 y Estrategia Nacional de Desarrollo) para identificar los principios y reglas que deben guiar el proceso de priorización. Esto también se detalla en el capítulo 3.

2.3.3. Indicadores, escalas y normalización

Para cada criterio se definen indicadores observables y reglas de normalización acordes con su naturaleza (beneficio/costo) y con la trazabilidad institucional (registros SNIP, fases del ciclo de vida, continuidad, programación de recursos). La consistencia semántica con el Reglamento (arts. 36–38, 77–81) y la Ley (arts. 36–38) se asegura mediante fichas de datos que documenten fuente, método de cálculo, periodicidad y responsable (DGIP/Unidades).

2.3.4. Integración con el flujo SNIP–PNPIP

El score AHP resultante se integra en el flujo operativo establecido por el Reglamento: incorporación al PNPIP de proyectos con código SNIP y dictámenes por etapa (art. 77), priorización anual antes del 31 de mayo con prelación continuidad “nuevos” (art. 79), y coherencia con la programación presupuestaria (cap. V, arts. 80–81). Esta integración asegura que la priorización por AHP opere como herramienta normativa, no como sustituto del marco legal, y que sus resultados sean utilizables en la programación plurianual y anual.

2.4. Aprendizaje automático en la priorización de proyectos

El aprendizaje automático constituye un conjunto de técnicas orientadas a identificar patrones y estructuras en los datos a partir de la experiencia empírica. Su aplicación en contextos de política pública y, en particular, en la inversión, se vincula con la necesidad de gestionar grandes volúmenes de proyectos de manera objetiva, transparente y reproducible. En este trabajo, el interés recae en los enfoques de aprendizaje no supervisado, en tanto permiten descubrir agrupamientos naturales en los datos sin requerir variables objetivo predeterminadas, aportando así una perspectiva complementaria al marco normativo establecido.

2.4.1. Enfoques supervisados: alcances y limitaciones

Los métodos de aprendizaje supervisado requieren un conjunto de entrenamiento con variables explicativas y una variable dependiente conocida, lo cual resulta útil en problemas de clasificación o regresión. Sin embargo, en el caso de la priorización de proyectos públicos, no existe una “etiqueta” inequívoca que indique a priori qué proyectos deben considerarse prioritarios o no. La asignación de esta categoría suele ser resultado de procesos políticos, institucionales o normativos, y no de una característica inherente de los datos.

En consecuencia, la aplicación de técnicas supervisadas en este contexto presenta limitaciones. Por un lado, tienden a replicar los sesgos de las decisiones históricas, perpetuando patrones de priorización que no necesariamente responden a criterios objetivos. Por otro, restringen la capacidad exploratoria, ya que obligan a definir ex ante una variable dependiente que en este caso es debatible.

En el contexto actual, la única variable dependiente capaz de responder a los objetivos de este trabajo es precisamente el score AHP. En la fase inicial del diseño metodológico, se planteó el uso del score AHP como variable dependiente para entrenar modelos de Learning to Rank (LTR). Sin embargo, esta propuesta induce una duplicidad de procesos debido a que el modelo solo estaría estimando casi a la perfección lo obtenido en el método AHP, además de replicar sesgos históricos y limitar la exploración relaciones contractuales. Por ello, el presente trabajo traslada el uso de modelos supervisados a un diseño experimental incluido en Anexo 1, y se concentra en los enfoques no supervisados, más coherentes con la necesidad de descubrir tipologías y patrones latentes en la cartera del PNPIP.

2.4.2. Aprendizaje no supervisado: fundamentos y técnicas de clustering

En contraste, el aprendizaje no supervisado se orienta a descubrir estructuras ocultas en los datos sin necesidad de variables dependientes. Dentro de este enfoque, el clustering es la técnica más extendida y se define como el proceso de agrupar objetos de manera que aquellos dentro de un mismo grupo sean más similares entre sí que con respecto a objetos de otros grupos (De Smet & Montano Guzmán, 2004).

El método más representativo es el K-Means, que busca particionar un conjunto de n observaciones en k clusters minimizando la suma de distancias cuadradas dentro de cada grupo (Lücke & Forster, 2019):

$$\min \sum_{i=1}^n \|x_i - \mu_{c(i)}\|^2$$

Donde x_i representa un proyecto, $c(i)$ su asignación de cluster y $\mu_{c(i)}$ el centroide correspondiente. La formulación del algoritmo EM (Dempster et al., 1977) mostró cómo el K-Means puede interpretarse como un caso particular de las mezclas gaussianas (GMM), ampliando así su fundamentación probabilística.

El clustering jerárquico, por su parte, construye una estructura en forma de árbol (dendrograma) donde las observaciones o clusters se van fusionando progresivamente

según su grado de proximidad. Sus ventajas radican en que no requiere definir a priori el número de clusters y permite explorar múltiples niveles de agrupamiento (Goldberger & Roweis, 2004). No obstante, su costo computacional aumenta en grandes conjuntos de datos, lo que ha dado lugar a variantes escalables como el Hierarchical Model-based clustering (Posse, 2001).

Por su parte, los métodos basados en densidad, como HDBSCAN, identifican regiones de alta densidad y permiten clasificar proyectos como ruido o outliers, lo que los hace especialmente útiles en bases heterogéneas y con presencia de casos atípicos (McInnes & Healy, 2017). Finalmente, el multicriteria clustering (De Smet & Montano Guzmán, 2004) extiende el K-Means integrando ponderaciones y funciones de valor en la medida de similitud, lo que abre la posibilidad de conectar directamente la lógica normativa del SNIP con técnicas de agrupamiento. La comparación de estos métodos se sintetiza en la Tabla 1:

Tabla 1. Comparativa de modelos de machine learning no supervisados

Método	Fundamento teórico	Parámetros clave	Ventajas	Limitaciones	Validación típica	Fuente
K-means	Minimiza la suma de distancias cuadradas a los centroides (variacional/EM)	Número de clusters K, inicialización, métrica (Euclídea)	Simple, eficiente, ampliamente usado	Sensible a inicialización, clusters esféricos, requiere K	Silhouette, Gap, Davies-Bouldin	Lücke & Forster (2019); Dempster et al. (1977)
GMM (Gaussian Mixture Models)	Estima distribución como mezcla de gaussianas, ajustada por EM	Número de componentes, inicialización, covarianzas	Flexibilidad probabilística, soft clustering	Supone normalidad, sensible a outliers	BIC/AIC, log-likelihood	Dempster et al. (1977); Chassagnol et al. (2023)
Clustering jerárquico aglomerativo	Construcción de dendrograma por proximidad sucesiva	Criterio de enlace (single, complete, average), métrica	No requiere definir K a priori, visualización jerárquica	Costoso en grandes datasets, sensible a ruido	Cophenetic correlation, Silhouette	Goldberger & Roweis (2004); Posse (2001)
HDBSCAN	Extiende DBSCAN a jerarquía de densidad	Distancia mínima (ϵ), minPts	Detecta ruido/outliers, no requiere K, soft clustering	Elección de parámetros compleja, menor interpretabilidad	Persistencia de clusters, estabilidad	McInnes et al. (2017)
Model-based hierarchical clustering	Ajuste jerárquico de mezclas gaussianas con MST	Número de componentes, criterios de fusión	Escalable a grandes datasets, combina jerarquía y modelos	Complejo de implementar, dependiente de supuestos	BIC/AIC, estabilidad bootstrap	Posse (2001)
Multicriteria clustering	Extiende K-means con múltiples métricas de preferencia	K, funciones de valor/ponderaciones	Integra decisiones multicriterio, aplicable a contextos normativos	Mayor complejidad en definición de criterios	Silhouette, medidas de distancia ponderada	De Smet & Montano Guzmán (2004)

La Tabla 1 evidencia que, aunque el K-means continúa siendo el método de referencia por su simplicidad, sus limitaciones justifican la exploración de enfoques más flexibles como GMM, clustering jerárquico o HDBSCAN. Al mismo tiempo, la propuesta de multicriteria clustering resulta especialmente pertinente para este trabajo, ya que

posibilita la incorporación de ponderaciones normativas en el proceso de agrupamiento, reforzando la coherencia entre los resultados empíricos y las exigencias institucionales del SNIP.

2.4.3. Validación y calidad de los agrupamientos

El valor analítico de los algoritmos de clustering depende de la validez y calidad de los agrupamientos generados, por lo que resulta imprescindible aplicar medidas objetivas de evaluación. Estas métricas permiten verificar la cohesión interna de cada grupo, la separación entre clusters, la estabilidad frente a perturbaciones de la muestra y la robustez numérica de los modelos.

Entre los indicadores clásicos, el índice de Silhouette (Rousseeuw, 1987) mide para cada observación la diferencia entre la distancia promedio a los elementos de su propio cluster y la distancia al cluster vecino más próximo. Su rango oscila entre -1 y 1 : valores cercanos a 1 reflejan asignaciones adecuadas, mientras que valores negativos sugieren posibles errores de agrupamiento.

La medida de Davies–Bouldin (Davies & Bouldin, 1979) evalúa la similitud entre clusters en función de su dispersión interna y la distancia entre centroides; cuanto menor es este índice, mayor es la separación relativa y la compacidad de los grupos. De forma complementaria, el índice de Calinski–Harabasz (Caliński & JA, 1974) compara la varianza inter-cluster con la intra-cluster, favoreciendo soluciones con alta separación entre grupos y baja dispersión interna.

En escenarios donde se aplican modelos probabilísticos, como las Mezclas Gaussianas (GMM), se recurre además a criterios de información como el Akaike Information Criterion (AIC) y el Bayesian Information Criterion (BIC), que penalizan la complejidad del modelo en función del número de parámetros (Dempster et al., 1977). Extensiones recientes proponen también el Integrated Completed Likelihood (ICL), que incorpora la entropía de las asignaciones probabilísticas como medida de certeza (Biernacki et al., 2000).

En métodos de clustering denso como DBSCAN/HDBSCAN, la validación se apoya en el índice DBCV (Density-Based Clustering Validation), diseñado para capturar la coherencia estructural de los clusters de densidad frente al ruido y a la detección de outliers (Moulavi et al., 2014).

Adicionalmente, se emplean enfoques de validación por remuestreo (bootstrap y subsampling) para estimar la estabilidad de los clusters, es decir, la consistencia de las asignaciones cuando se perturba el conjunto de datos o se varían las condiciones de inicialización. Este criterio, propuesto por Yu, Liang y Hu (2019), resulta fundamental en contextos aplicados, ya que un agrupamiento inestable pierde valor predictivo y explicativo.

Finalmente, los estudios metodológicos recientes recomiendan incluir pruebas de estrés en la validación de modelos de clustering, evaluando la sensibilidad frente a la aparición de micro-clusters (con tamaños muy reducidos), el desbalance excesivo entre grupos y

la proporción de observaciones clasificadas como ruido (McInnes & Healy, 2017). Estas medidas complementarias permiten garantizar que la estructura obtenida no sea producto de artefactos estadísticos, sino de patrones robustos y útiles para la interpretación en políticas públicas.

2.4.4. Interpretabilidad y transparencia en la inversión pública

La interpretabilidad constituye un requisito esencial para la adopción de técnicas de inteligencia artificial en el sector público. Un modelo que clasifica proyectos sin ofrecer explicaciones claras sobre las razones de esa clasificación pierde legitimidad frente a los actores institucionales y territoriales involucrados.

En este sentido, la literatura resalta la importancia de vincular los clusters con dimensiones observables y comprensibles por los decisores, como tipologías sectoriales, niveles de financiamiento o características territoriales (Mariani et al., 2023). El análisis de perfiles de clusters, representaciones gráficas (radar plots) y medidas de contribución relativa facilitan esta tarea, permitiendo traducir resultados técnicos en evidencia comprensible.

2.5. Integración y evaluación del enfoque híbrido

2.5.1. Integración normativo–empírico

El AHP proporciona un marco normativo sólido, derivado de la Ley 498-06 y del Reglamento 493-07, al permitir la construcción de ponderaciones explícitas para los criterios institucionales C1–C7. Sin embargo, aplicado de manera aislada, puede limitar la detección de patrones subyacentes en la cartera de proyectos, al homogeneizar resultados en función de comparaciones normativas predefinidas.

Por su parte, el clustering permite identificar agrupamientos empíricos de proyectos en función de sus atributos técnicos, territoriales o institucionales, revelando tipologías que no necesariamente emergen de los criterios normativos. De esta forma, los clusters pueden actuar como complemento al AHP, mostrando coincidencias o divergencias entre la priorización formal y las estructuras empíricas de la cartera.

2.5.2. Evaluación normativa

El modelo híbrido debe ser evaluado en función de su capacidad para cumplir con los principios y objetivos establecidos en el marco del SNIP y la Estrategia Nacional de Desarrollo 2030. Entre los criterios normativos más relevantes se encuentran:

- **Equidad territorial:** distribución equilibrada de los recursos entre regiones, evitando la concentración excesiva de inversión en determinados territorios.
- **Diversidad institucional:** participación de diferentes sectores y entidades ejecutoras en la cartera priorizada, lo que refuerza la representatividad del sistema.
- **Alineación estratégica:** coherencia de los proyectos priorizados con los ejes de la Estrategia Nacional de Desarrollo, los planes sectoriales y los compromisos de política pública vigentes.

2.5.3. Evaluación técnica

La robustez de los agrupamientos se valida con métricas de clustering. A continuación, se describen las métricas que se emplean (Tabla 2):

Tabla 2. Métricas de validación en clustering

Métrica	Definición	Interpretación de valores	Fuente
<i>Silhouette</i>	Diferencia entre la cohesión interna (distancia promedio al cluster propio) y la separación (distancia al cluster vecino más cercano).	Valores cercanos a 1 indican asignación adecuada; valores cercanos a 0 indican solapamiento; valores negativos indican mala asignación.	Rousseeuw (1987)
<i>Davies–Bouldin</i>	Relación entre la dispersión interna de cada cluster y la distancia entre centroides.	Valores bajos reflejan mejor separación y mayor compacidad de los clusters.	Davies & Bouldin (1979)
<i>Calinski–Harabasz</i>	Cociente entre la varianza entre clusters y la varianza dentro de clusters.	Valores altos indican clusters bien separados y compactos.	Calinski & Harabasz (1974)

Capítulo 3

3. Metodología general

3.1. Etapa 1: Preparación de datos

3.1.1. Fuentes de datos utilizadas

La fuente primaria de información utilizada en esta investigación corresponde al conjunto de datos oficial de proyectos de inversión pública, compilado en el archivo Datos_TFM.xlsx. Este dataset fue proporcionado por la Dirección General de Inversión Pública (DGIP) del Ministerio de Economía, Planificación y Desarrollo (MEPyD) y constituye la base empírica sobre la cual se construye el modelo de priorización.

El archivo recoge la cartera de proyectos considerados para el Plan Nacional Plurianual de Inversión Pública (PNPIP) 2025–2028, integrando tanto proyectos de arrastre como nuevas iniciativas sometidas por las instituciones ejecutoras. La información se encuentra estructurada a nivel de proyecto y contiene variables normativamente definidas en el marco del Sistema Nacional de Inversión Pública (SNIP), tales como características institucionales, técnicas, territoriales y financieras.

El dataset presenta un total de 1.947 proyectos registrados, cada uno identificado con un código único SNIP y acompañado de metadatos como el nombre de la institución ejecutora, su sigla y la denominación del proyecto. En el inventario preliminar se observa que las variables categóricas (por ejemplo, Sección, Institución o Sigla) no presentan valores faltantes, mientras que los campos numéricos muestran una alta diversidad de registros, coherente con la heterogeneidad de la cartera. Esta exploración inicial confirma la completitud del archivo y permite establecer la base para las siguientes etapas de limpieza, transformación y análisis. A continuación, se detalla una breve descripción (Tabla 3):

Tabla 3. Diccionario de variables

VARIABLE	DESCRIPCIÓN	TIPO
SNIP	Código único asignado por el Sistema Nacional de Inversión Pública (SNIP) a cada proyecto. Permite la trazabilidad y control en el ciclo de vida del proyecto.	Texto
Proyecto	Denominación oficial del proyecto según el registro en el SNIP. Resume su objetivo o finalidad principal.	Texto
Institución / Sigla	Nombre de la institución ejecutora responsable del proyecto (ministerio, entidad descentralizada o ayuntamiento).	Catagórica nominal
Sección	Clasificación del proyecto según el sector de política pública al que pertenece (ej. transporte, salud, educación).	Catagórica nominal
Funcional	Clasificador de responsabilidad funcional del SNIP.	
Provincia	Localización administrativa principal del proyecto, asociada a las 32 provincias del país.	Catagórica nominal
Monto total (RD\$)	Presupuesto total estimado del proyecto en pesos dominicanos.	Numérica continua (monetaria)
Monto 2025 (RD\$)	Asignación presupuestaria programada para el año 2025.	Numérica continua (monetaria)
Monto 2026 (RD\$)	Asignación presupuestaria programada para el año 2026.	Numérica continua (monetaria)
Monto 2027 (RD\$)	Asignación presupuestaria programada para el año 2027.	Numérica continua (monetaria)
Monto 2028 (RD\$)	Asignación presupuestaria programada para el año 2028.	Numérica continua (monetaria)
Condición Presupuestaria	Clasificación de si el proyecto es de <i>arrastre</i> (ya iniciado) o <i>nuevo</i> (sin ejecución previa), según PNPIP.	Catagórica nominal
Tipología	Tipo de inversión: capital fijo, capital humano o creación de conocimiento.	Catagórica nominal
Grado De Avance Financiero (C1)	Porcentaje de recursos financieros disponibles respecto al monto total del proyecto. Refleja la proximidad a su finalización.	Numérica porcentual
Proporción De Financiamiento Externo (C2)	Proporción de financiamiento externo (créditos o donaciones) en el monto total del proyecto.	Numérica porcentual
Densidad Poblacional (C3)	Densidad poblacional del territorio de influencia del proyecto (habitantes por km ²).	Numérica continua
Nivel socioeconómico (C4)	Proporción de hogares de nivel socioeconómico familiar "Muy bajo".	Numérica porcentual
Inversión Histórica Promedio Per Cápita (C5)	Criterio normativo: monto promedio per cápita de inversión recibida históricamente por provincia.	Numérica continua (monetaria)
Contribución Al PNPS (C6)	Criterio normativo: contribución declarada del proyecto a los Resultados de Desarrollo del PNPS vigente.	Catagórica ordinal
Demanda Territorial (C7)	Criterio normativo: grado de respaldo del proyecto por demandas priorizadas en Consejos Provinciales de Desarrollo.	Catagórica ordinal

3.1.2. Procesos de limpieza e imputación

Con el propósito de garantizar la calidad del insumo empírico, se ejecutó un protocolo de verificación de datos orientado a unicidad, completitud, tipificación, rangos válidos y coherencias internas. De acuerdo con las trazas de control implementadas en el entorno de trabajo, no se detectaron registros duplicados ni valores perdidos. Asimismo, no se identificaron observaciones atípicas que comprometieran la validez de la base, con la excepción de ciertos valores extremos en variables territoriales y financieras. Estos casos no se corrigieron en esta fase, pues serán tratados en los próximos apartados mediante procedimientos de ingeniería de características que permitirán discretizar y conceptualizar dichos atributos en consonancia con los criterios normativos del PNPIP. De este modo, se mantuvo la integridad del dataset maestro sin aplicar imputaciones ni modificaciones irreversibles.

El protocolo comprendió las siguientes comprobaciones:

- **Unicidad de identificadores:** se verificó la unicidad del ID_SNIP a nivel de fila, confirmándose la ausencia de duplicados.

- **Complejidad:** se inspeccionó la presencia de valores nulos por variable, constatándose complejidad total en las columnas requeridas para el análisis.
- **Tipificación y dominios:** no se detectaron valores negativos en variables monetarias, porcentajes fuera de rango en los indicadores proporcionales ni clasificaciones erróneas en las variables categóricas nominales u ordinales.
- **Coherencias financieras plurianuales:** se validó la consistencia entre los montos anuales y la programación plurianual declarada, sin incongruencias reportadas.
- **Detección de atípicos:** se aplicaron reglas de detección estándar (percentiles/IQR) sobre montos y densidades, identificándose valores extremos en variables territoriales y financieras. Estos se conservaron en el dataset maestro, ya que en el apartado 3.1.3 se explicará cómo se transformarán en variables discretas para su uso analítico.
- **Preservación del dato maestro:** dado que no se registraron incidencias, no se realizaron transformaciones destructivas sobre el dataset.

3.1.3. Ingeniería de características

En esta etapa se procedió a la construcción de las variables conceptuales que permiten operacionalizar los criterios normativos C1–C7 definidos en el marco del PNPIP y el SNIP. Estas variables derivadas convierten información financiera, territorial y socioeconómica en escalas ordinales homogéneas, compatibles con el método AHP, preservando siempre la integridad de los datos.

A continuación, se detalla una breve descripción de estas variables:

- **C1. Grado de avance del proyecto:** se calculó el cociente entre el presupuesto disponible y el monto total del proyecto. Este indicador se discretizó en cuatro intervalos siguiendo la escala de Saaty (1, 3, 5, 7), asignando mayor puntuación a los proyectos más próximos a su finalización.
- **C2. Proporción de financiamiento externo:** se definió como la proporción de créditos externos y donaciones respecto al monto total del proyecto. Los proyectos con mayor financiamiento externo recibieron mayor calificación, conforme a la lógica de sostenibilidad fiscal.
- **C3. Densidad poblacional:** se utilizó la densidad poblacional a nivel provincial. Para evitar distorsiones, se calcularon los cuartiles tras excluir valores extremos mediante el rango intercuartílico (IQR). Los proyectos en provincias con menor densidad obtuvieron mayor puntuación, en línea con el principio de equidad territorial.
- **C4. Nivel socioeconómico del territorio:** se empleó la proporción de hogares en condición socioeconómica familiar “Muy bajo”. El criterio se segmentó mediante cuartiles robustos, priorizando provincias con mayores niveles de vulnerabilidad social.

- C5. Inversión histórica promedio per cápita: se derivó a partir de la inversión promedio ejecutada en el territorio (nivel provincial) en el quinquenio anterior. El criterio asigna mayor puntuación a provincias históricamente menos beneficiadas, corrigiendo la concentración territorial de la inversión pública.
- Los criterios C6 (Contribución con el PNPSP) y C7 (Demanda territorial priorizada) conservaron su codificación ordinal original, ya que fueron definidos institucionalmente y no requieren transformaciones adicionales.

3.1.4. Variables seleccionadas

Tras la fase de ingeniería de características, se procedió a definir el conjunto de variables que se utilizará en los modelos de priorización. La selección se realizó con un doble criterio: por un lado, la trazabilidad normativa con los criterios del Sistema Nacional de Inversión Pública (SNIP), y por otro, la pertinencia metodológica para la aplicación del método AHP y de los algoritmos de aprendizaje automático.

El núcleo analítico del modelo está conformado exclusivamente por los siete criterios normativos (C1–C7). Estas variables se encuentran estandarizadas en escalas ordinales homogéneas (1, 3, 5, 7), lo que garantiza su comparabilidad y coherencia con la metodología AHP. Asimismo, constituyen los insumos directos para el cálculo del score normativo y para la identificación de tipologías mediante clustering.

De manera complementaria, se conservaron variables institucionales, territoriales y financieras (SNIP, Proyecto, Institución, Sección, Funcional, Provincia, Condición presupuestaria, Monto total y Programación plurianual) con fines exclusivamente descriptivos y de control. Estas variables permiten caracterizar la cartera, analizar su distribución por territorios e instituciones, y evaluar la equidad de los resultados, pero no intervienen directamente en los algoritmos de priorización.

3.1.5. Análisis descriptivo

3.1.5.1. Variables financieras

El grado de avance financiero presenta un valor medio de 0.70, lo que indica que, en promedio, los proyectos aún deben ejecutar un 70% de su presupuesto. La mediana (0.81) y el primer cuartil (0.52) confirman que al menos una cuarta parte de los proyectos ya ha ejecutado cerca de la mitad de sus recursos, mientras que otro 25% apenas ha ejecutado menos del 20%. Este comportamiento evidencia la relevancia del criterio C1, que permite discriminar entre proyectos en fases iniciales y aquellos próximos a concluir.

Por su parte, la proporción de financiamiento externo muestra una media de 0.03, con mediana y cuartil superior prácticamente nula, lo que refleja que la mayoría de los proyectos carece de financiamiento externo. No obstante, la presencia de valores máximos de hasta 1.00 y una desviación estándar de 0.16 revelan un subconjunto reducido de proyectos con alta dependencia de fuentes externas, aspecto crítico para el criterio C2.

3.1.5.2. Variables de perfil del proyecto

Las variables de perfil institucional, sectorial y territorial permiten comprender la estructura de la cartera y los sesgos que podrían influir en la asignación de recursos.

- Sección institucional: predominan los proyectos del Gobierno Central (84,3%), mientras que empresas estatales y entidades descentralizadas representan el 12,2% y 3,2%, respectivamente. Esta centralización sugiere la necesidad de incorporar criterios de equidad interinstitucional.
- Institución ejecutora: aunque participan 60 entidades, dos instituciones concentran el 65% de los proyectos, lo que plantea riesgos de concentración presupuestaria.
- Condición presupuestaria: el 76,2% de los proyectos son de arrastre y 23,8% nuevos, proporción coherente con la Ley 498-06 (art. 17), que otorga prioridad a la continuidad.
- Clasificación funcional: se observa una concentración en Educación (34,2%), Transporte (18,9%) y Vivienda (13,5%), mientras que sectores como Energía, Medioambiente y Protección social muestran menor representación.
- Área de influencia: predominan los proyectos de escala municipal (79,6%), lo que refleja fragmentación territorial y limitada articulación interregional.
- Distribución regional y provincial: la región Metropolitana concentra el 16,1% de los proyectos, seguida de regiones del Cibao, mientras que zonas del sur presentan menor participación. A nivel provincial, Santo Domingo (10%), Santiago (7,4%) y el Distrito Nacional (5,1%) lideran en número de proyectos.

3.1.5.3. Variables sociodemográficas

Las variables sociodemográficas muestran desigualdades estructurales en la distribución de la población y en la trayectoria histórica de inversión:

- Densidad poblacional: la media (1,043 hab/km²) contrasta con una mediana de 189 hab/km², confirmando la existencia de outliers extremos y justificando la inclusión del criterio C3 para priorizar proyectos en zonas menos densas.
- Nivel socioeconómico del territorio: el promedio de 24,1% y la mediana de 21,5% muestran que al menos la mitad de los proyectos se ubican en zonas con una quinta parte de hogares en alta vulnerabilidad, lo que valida el criterio C4.
- Inversión histórica promedio per cápita: la media (4,43 mil millones de RD\$) y la mediana (1,76 mil millones) evidencian una distribución asimétrica, donde algunos territorios concentran la mayoría de la inversión, mientras otros permanecen rezagados.

3.2. Etapa 2: Construcción del score normativo (AHP)

3.2.1. Definición de criterios

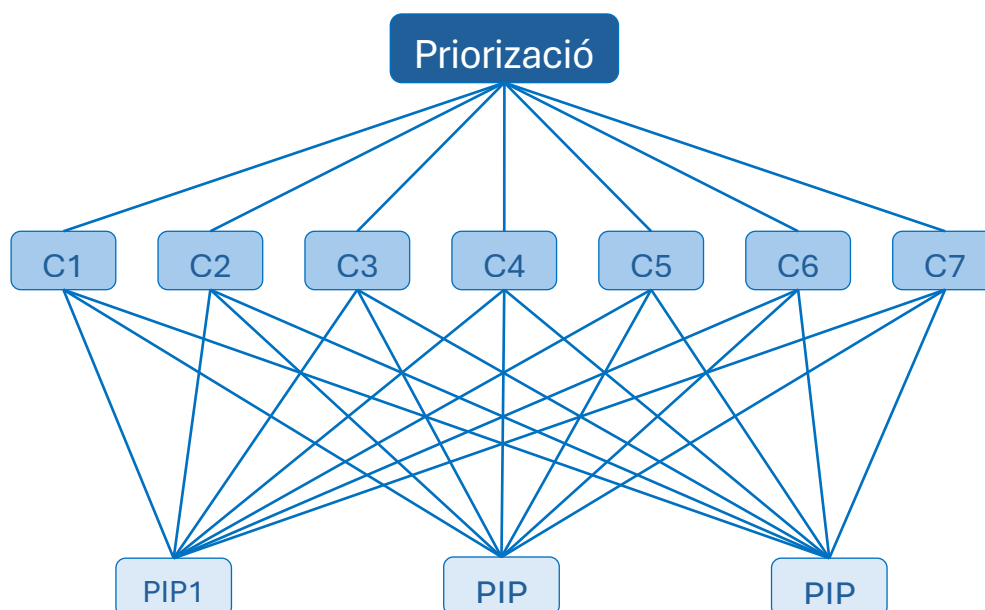
El modelo de priorización se apoya en el Proceso Analítico Jerárquico (AHP), propuesto por Saaty en la década de 1970 como técnica multicriterio para estructurar problemas complejos en jerarquías de objetivos, criterios y alternativas (Bertolini et al., 2006; Corrales et al., 2024). En este trabajo se definieron siete criterios (C1–C7), derivados de

la normativa del SNIP y del PNPIP, los cuales reflejan dimensiones financieras, territoriales y estratégicas de los proyectos de inversión.

El esquema jerárquico del modelo se organiza en tres niveles (Figura 1) :

- i. Objetivo global: priorización de proyectos de inversión pública bajo restricción presupuestaria.
- ii. Criterios normativos (C1–C7): grado de avance, financiamiento externo, densidad poblacional, nivel socioeconómico, inversión histórica, alineación con el PNPSP, demandas territoriales.
- iii. Alternativas: proyectos individuales (*PIP*) de la cartera PNPIP.

Figura 1. Esquema jerárquico de criterios



3.2.2. Matriz de comparación de pares

La comparación relativa entre criterios se realizó mediante la escala fundamental de Saaty (1–9), que permite expresar la importancia de un criterio respecto a otro. Se incluyeron también valores intermedios (2, 4, 6, 8) y recíprocos (1/2, 1/3, ..., 1/9) para matizar la intensidad del juicio (Tabla 4).

Tabla 4. Escala de Saaty de importancia relativa

Valor	Interpretación
1	Igual importancia
2	Importancia entre igual y moderada
3	Moderada importancia
4	Moderada a fuerte
5	Fuerte importancia
6	Fuerte a muy fuerte
7	Muy fuerte importancia
8	Muy fuerte a extrema
9	Importancia extrema

La matriz resultante tiene dimensión 7×7 y contiene 21 comparaciones únicas. A continuación, se presenta la tabla de juicios iniciales utilizados (valores normativos y técnicos interpretados a partir de la Ley 498-06, Decreto 493-07 y PNPIP 2025-2028):

Tabla 5. Comparación de criterios

Comparación	Valor asignado	Justificación normativa y técnica
C1 vs C2	7	La ley 498-06 resalta que el cumplimiento de las metas planificadas depende del “grado de avance del proyecto (C1)” para ser incluidos en el PNPIP. La proporción de financiamiento externo (C2) es importante, pero sugiere disponibilidad financiera exógena, no madurez del proyecto.
C1 vs C3	5	La densidad poblacional (C3) permite identificar impactos territoriales, pero si el proyecto no ha avanzado, no se puede ejecutar. El grado de avance tiene prioridad técnica y presupuestaria (Art. 30, Ley 498-06).
C1 vs C4	3	El nivel socioeconómico del territorio (C4) es un principio de equidad (Ley END, art. 17), pero no debe primar sobre la factibilidad técnica. Ambos criterios deben equilibrarse, pero el avance técnico prevalece.
C1 vs C5	2	La equidad territorial es importante según la END, pero si un proyecto ya inició y otro no, se recomienda continuar el que está en curso (Art. 3, literal h, Ley 498-06).
C1 vs C6	3	La contribución al PNSP (C6) es relevante, pero debe ejecutarse sobre proyectos técnicamente viables y avanzados, conforme a Art. 36 de Ley 498-06.
C1 vs C7	2	Las demandas territoriales priorizadas (C7) son importantes (Decreto 493-07, Art. 97), pero subordinadas a la viabilidad del proyecto (Art. 3.d, Ley 498-06).
C2 vs C3	2	Un financiamiento externo no implica por sí mismo mayor impacto. En zonas densamente pobladas, la inversión puede generar mayor efecto redistributivo.
C2 vs C4	0.5 (equivalente a 1/2)	Aunque el nivel socioeconómico es clave para focalización del gasto, el acceso a financiamiento externo facilita la ejecución en el corto plazo. Pero no justifica priorizarlo per se.
C2 vs C5	0.333 (1/3)	Invertir en provincias históricamente rezagadas es mandato normativo (Ley END, Art. 17), y debe tener mayor peso que la oportunidad de financiamiento externo.
C2 vs C6	0.5 (1/2)	Financiamiento disponible facilita la ejecución, pero la contribución al PNSP tiene mayor relevancia para cumplir metas nacionales (Art. 30, Ley 498-06).
C2 vs C7	0.5 (1/2)	Las demandas del territorio tienen un peso normativo fuerte (Decreto 493-07), en contraste con el financiamiento que es una condición instrumental.
C3 vs C4	1	Ambos abordan desigualdades territoriales desde distintas ópticas: concentración demográfica vs. vulnerabilidad estructural. Deben ponderarse en igualdad.
C3 vs C5	0.5 (1/2)	A igual densidad, se debe favorecer a provincias menos beneficiadas históricamente (END Art. 17, Ley 498-06).
C3 vs C6	0.333 (1/3)	Aunque la densidad permite mayor impacto, el alineamiento con el PNSP es criterio técnico prioritario para la inclusión en el PNPIP (Ley 498-06, Art. 36).
C3 vs C7	0.333 (1/3)	Las prioridades expresadas por la comunidad deben ser más consideradas que la densidad como factor aislado (Ley 498-06, Art. 3.j).
C4 vs C5	0.5 (1/2)	El nivel socioeconómico (C4) se correlaciona con C5, pero C5 apunta explícitamente a cerrar brechas de inversión histórica (mayor peso según END y Ley 498-06).
C4 vs C6	0.333 (1/3)	Aunque C4 refleja equidad, los planes nacionales son directrices jerárquicas que orientan la ejecución de proyectos (Art. 1, Ley 498-06).
C4 vs C7	0.5 (1/2)	Ambos buscan justicia territorial. La demanda local da contenido a las carencias, pero el nivel socioeconómico es un proxy más estructural.
C5 vs C6	1	Equilibrio. Una provincia poco beneficiada puede tener proyectos que no se alinean con el PNSP, y viceversa. Ambas deben ser consideradas.
C5 vs C7	3	Las demandas priorizadas localmente deben pasar por una evaluación técnica. C5 implica una injusticia histórica más objetiva.
C6 vs C7	2	Contribuir al PNSP responde a metas nacionales de largo plazo; aunque las demandas del territorio (C7) tienen legitimidad social, deben estar alineadas con las políticas nacionales (Ley 498-06, Art. 3.a).

Con estas comparaciones se construyó la matriz A :

$$A = \begin{bmatrix} 1 & a_{12} & a_{13} & \dots & a_{17} \\ 1/a_{12} & 1 & a_{23} & \dots & a_{27} \\ 1/a_{13} & 1/a_{23} & 1 & \dots & a_{37} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1/a_{17} & 1/a_{27} & 1/a_{37} & \dots & 1 \end{bmatrix} = \begin{bmatrix} 1 & 7 & 5 & \dots & 2 \\ 1/7 & 1 & 2 & \dots & 1/2 \\ 1/5 & 1/2 & 1 & \dots & 1/3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1/2 & 2 & 3 & \dots & 1 \end{bmatrix}$$

Donde a_{ij} representa la importancia relativa del criterio C_i sobre C_j .

3.2.3. Cálculo de pesos normalizados

El vector de prioridades w se obtuvo aplicando el método del autovector principal. El procedimiento seguido fue:

- i. Normalización por columnas:

$$a'_{ij} = \frac{a_{ij}}{\sum_{i=1}^n a_{ij}}$$

- ii. Promedio por filas:

$$w_i = \frac{1}{n} \sum_{j=1}^n a'_{ij}$$

- iii. Vector de prioridades:

$$W = (w_1, w_2, \dots, w_n)$$

Con estos pasos se derivan los pesos relativos de cada criterio, garantizando que $\sum w_i = 1$.

3.2.4. Validación del Índice de consistencia

Para comprobar la coherencia de los juicios, se calculó el índice de consistencia (CI) y la Razón de consistencia (CR):

$$CI = \frac{\lambda_{max} - n}{n - 1}, \quad CR = \frac{CI}{RI}$$

Donde λ_{max} es el valor propio máximo de la matriz, $n = 7$, y $RI = 1.32$, que corresponde al índice aleatorio esperado para matrices de tamaño 7 (Saaty, 1987).

En este caso, se obtuvo un $CR = 0.0481 < 0.1$. En el apartado de resultados se abordará más al respecto.

3.2.5. Asignación del score a cada proyecto

Una vez obtenidos los pesos, se procedió a calcular el score normativo de cada proyecto como suma ponderada de los valores en C1-C7:

$$Score\ AHP_k = \sum_{i=1}^7 w_i \cdot C_{ik}$$

Donde C_{ik} es la calificación del proyecto k en el criterio i , y w_i el peso normalizado.

Para facilitar la interpretación, el score se escaló a un rango porcentual 0-100%:

$$Score\ AHP_{ptge}_k = \frac{Score\ AHP_k - min}{max - min} \times 100$$

De esta forma, cada proyecto obtuvo una puntuación relativa que permite establecer un ranking normativo de priorización.

3.3. Etapa 3: Modelado no supervisado

El aprendizaje no supervisado se incorporó como segunda dimensión metodológica del estudio, con el objetivo de identificar agrupamientos naturales entre los proyectos del PNPIP a partir de los criterios C1–C7. Estos clusters permiten complementar el score normativo AHP con tipologías empíricas, aportando evidencia adicional sobre patrones de madurez técnica, equidad territorial y diversidad institucional. Para garantizar la solidez del análisis, se abordaron de forma sistemática cuatro enfoques: K-Means, clustering jerárquico (Ward), mezclas gaussianas (GMM) y clustering denso (HDBSCAN). Cada modelo se sometió a pruebas de estrés diseñadas para validar la robustez de la solución y evitar interpretaciones espurias.

3.3.1. Preprocesamiento de datos para modelado

Previo a la aplicación de algoritmos de aprendizaje no supervisado, se llevó a cabo un protocolo de preprocesamiento de datos orientado a garantizar la homogeneidad de escalas, la trazabilidad de transformaciones y la adecuación de las variables normativas a un espacio de características compatible con técnicas de clustering.

El procedimiento se estructuró en cinco fases:

- i. Selección de variables base: en este proceso, se seleccionan las variables continuas base que corresponden a los criterios C1–C5 y las variables ordinales originales asociadas a los criterios C6 y C7. Esta decisión metodológica responde a dos consideraciones centrales: en primer lugar, los algoritmos de clustering requieren trabajar sobre magnitudes comparables y con información granular; en segundo lugar, las variables ordinales de origen reflejan escalas institucionales que deben ser transformadas para garantizar estabilidad discriminativa.
- ii. Tratamiento de variables ordinales: los criterios Contribución al PNPS (C6) y Demandas priorizadas por el territorio (C7) presentan naturaleza ordinal en escala discreta (1, 3, 5, 7). Su manejo requería una estrategia que preservara el orden jerárquico sin introducir supuestos de linealidad artificial. Por ello, en lugar de mantenerlos como enteros o aplicar One-Hot Encoding, se optó por un target-rank encoding supervisado por el score AHP:
 - Para cada categoría se calculó la media de score_AHP.
 - Se ordenaron las categorías según su impacto medio.
 - Se asignaron valores en el rango [0, 1] preservando la jerarquía empírica observada.
- iii. Escalado y estandarización: todas las variables de entrada al clustering fueron sometidas a un doble escalado:
 - RobustScaler, que reduce la influencia de valores atípicos mediante normalización sobre la mediana y el rango intercuartílico.
 - StandardScaler, que centra las variables en media cero y varianza unitaria, requisito habitual para algoritmos basados en distancias euclidianas.
- iv. Construcción de la matriz de características: el conjunto final de features incluyó las cinco continuas y las dos variables ordinales codificadas (C6 y C7). Como

resultado, se obtuvo un dataframe escalado (df_features) y su matriz NumPy equivalente (X_clu), de dimensión [n_proyectos x 7], que constituye la base analítica para el modelado no supervisado.

- v. Reproducibilidad: se documentaron los objetos generados para garantizar trazabilidad:
 - df_features: dataset listo para clustering.
 - X_clu: matriz de entrada a algoritmos.
 - col_features: lista de variables seleccionadas.
 - preprocess_clu: pipeline de preprocesamiento reutilizable.
 - df_clu_base: dataset con variables originales y codificaciones.

3.3.2. Clustering con K-Means

3.3.2.1. Construcción del modelo K-Means

El algoritmo K-Means se utilizó como línea base por su capacidad para particionar observaciones en k grupos disjuntos optimizando la inercia intra-cluster. Dados los features estandarizados de 3.3.1, el problema se plantea como:

$$\min_{\{C_1, \dots, C_k\}} \sum_{j=1}^k \sum_{x_i \in C_j} \|x_i - \mu_j\|_2^2 \quad \text{donde } \mu_j = \frac{1}{|C_j|} \sum_{x_i \in C_j} x_i$$

Se parte del siguiente procedimiento:

- i. Búsqueda de k (grid 2-12): para cada $k \in \{2, \dots, 12\}$ se ajustó un modelo K-Means (n_clusters= k , n_init="auto", random_state=42).
- ii. Evaluación multi-métrica (calidad interna): se calcularon métricas estándar con distancia euclídea:
 - Silhouette (+): cohesión/separación relativa.
 - Davies–Bouldin (-): promedio de “semejanzas” entre clusters (menor es mejor).
 - Calinski–Harabasz (+): varianza inter vs. intra.
- iii. Estabilidad por remuestreo (robustez): para cada k , se ejecutaron 30 remuestreos estratificados al 80% (SAMPLE_FRAC=0.80, semillas controladas). En cada submuestra s , se volvió a ajustar K-Means y se compararon las etiquetas con las del ajuste base mediante índice de Jaccard (promedio y desviación). Este bloque estima la estabilidad estructural del particionado ante perturbaciones de la muestra.
- iv. Selección de k por ranking combinado: se construyó una tabla resumen por k y se aplicaron ranks por criterio:
 - Silhouette (desc), Davies-Bouldin (asc), Calinski–Harabasz (desc) y Estabilidad (desc).
 - Rank promedio = media de los cuatro ranks.

3.3.2.2. Pruebas de estrés para selección de k dominante en K-Means

Una vez identificado el rango de candidatos a k mediante el análisis multi-métrica inicial, se realizaron pruebas de estrés con el objetivo de validar la robustez de la elección y

garantizar que el número de clusters seleccionado no dependiera de condiciones particulares de la muestra o de la inicialización del algoritmo.

El procedimiento consistió en evaluar los valores $k = \{7,8\}$ —y opcionalmente $k = 9$ — bajo un esquema reforzado de validación:

- Se ejecutaron 100 remuestreos bootstrap al 90% de la muestra para estimar la estabilidad de Jaccard con mayor precisión (Yu et al., 2019) .
- Se verificó que ningún cluster tuviera un tamaño inferior al 5% de la muestra, criterio que penaliza soluciones con particiones marginales y poco interpretables.
- Se calcularon las métricas internas globales: Silhouette, Davies–Bouldin y Calinski–Harabasz, asegurando consistencia con la etapa previa.
- Se analizaron los gráficos de Silhouette por cluster, que permiten visualizar la cohesión y separación de cada grupo (Bock, 2008).
- Se recuperaron los centroides en escala original para interpretar las características medias de cada agrupamiento en términos de las variables C1–C7.

El resultado de estas pruebas se sintetizó en un ranking combinado que integra métricas de calidad interna, estabilidad y penalización por clusters pequeños. Este esquema permite identificar el k dominante, es decir, aquel que se mantiene robusto incluso ante perturbaciones intensivas en la muestra.

3.3.3. Clustering Jerárquico (Ward)

El clustering jerárquico aglomerativo con enlace Ward se incorporó como alternativa a K-Means para contrastar supuestos geométricos y evaluar estructuras de partición que minimizan la varianza intra-cluster de forma greedy. A diferencia de K-Means, Ward no requiere inicialización aleatoria ni convexidad estricta de clústeres, y permite inspeccionar relaciones jerárquicas mediante dendogramas, lo que aporta trazabilidad estructural al proceso.

3.3.3.1. Construcción del modelo Ward

El procedimiento abordado se estructuró en tres fases:

- i. Dendrograma (diagnóstico estructural):
 - Se construyó el dendrograma con `linkage(..., method="ward")`.
 - Para conjuntos con $n > 1500$, se tomó una muestra aleatoria fija (semilla 42) para asegurar legibilidad y reproducibilidad.
 - Se estimó el coeficiente cofenético (muestra del dendrograma) como medida de fidelidad entre distancias originales y ultramétricas del árbol (cuando fue computable).
- ii. Barrido de $k(2 - 12)$ con métricas internas y estabilidad.
 - Para cada k se ajustó `AgglomerativeClustering(n_clusters=k, linkage="ward")` y se calcularon Silhouette, Davies–Bouldin y Calinski–Harabasz, manteniendo distancia euclídea coherente con el preprocesamiento y la naturaleza de Ward.

- Se estimó estabilidad por remuestreo (30 réplicas al 80%): una función `fit_predict_fn` re-ajusta Ward en cada submuestra y se calcula Jaccard frente a las etiquetas base, obteniendo media y desviación.
- iii. Selección de k por ranking multi-métrica.
 - Se asignaron ranks por métrica igual que en el modelo anterior y se promedió para obtener un rank combinado.
 - k es el de menor rank promedio, evitando decisiones basadas en una sola métrica.

3.3.3.2. Pruebas de estrés para selección de k dominante en Ward

Para confirmar un k dominante bajo Ward, se diseñó un bloque de pruebas de estrés sobre candidatos específicos $k = 5,7$, partiendo de lo siguiente:

- Tamaño del cluster: penalizaciones por baja interpretabilidad (<5%).
- Métricas globales: Silhouette, Davies–Bouldin, Calinski–Harabasz.
- Estabilidad reforzada (100 remuestreos al 90%): Jaccard medio y desviación.
- Silhouette por clúster (plot): inspección visual de cohesión y separación intra/inter-grupo.
- Centroides en escala original: medias de cada clúster en el espacio desescalado.

El k dominante se definió como el de rank promedio mínimo + penalización mínima, privilegiando robustez y parsimonia.

3.3.4. Mezclas Gaussianas (GMM)

El modelo de Mezclas Gaussianas (GMM) permite particiones blandas (probabilísticas) y clústeres con forma y escala más flexibles que K-Means/Ward. Se evaluó un grid de número de componentes $k = 5 \dots 7$ y tipos de covarianza (full, tied, diag, spherical), midiendo ajuste penalizado e índices geométricos/estabilidad.

3.3.4.1. Construcción del modelo GMM

El procedimiento abordado se estructuró en dos fases:

- i. Ajuste por grid (k, cov): para cada combinación se ajustó GaussianMixture con `reg_covar=1e-6` (estabilidad numérica) y se obtuvieron:
 - BIC y AIC del ajuste.
 - Etiquetas duras predict, responsabilidades predict_proba.
 - Métricas globas: Sil, DB, CH.
 - Estabilidad por remuestreo (30 réplicas al 80%): Jaccard medio y desviación.
- ii. Selección por BIC y por ranking multi-métrica:
 - Mejor BIC: (k, cov) que minimiza BIC.
 - Mejor ranking: ranks por BIC (asc), Sil (desc), DB (asc), CH (desc), Estab (desc).

3.3.4.2. Pruebas de estrés para selección de k dominante en GMM

Además de BIC/AIC, se incorporó un bloque de estrés específico para mezclas: ICL, certeza de clasificación, condición de covarianzas, pesos mínimos por componente y estabilidad reforzada. El objetivo es evitar configuraciones

numéricamente inestables, con componentes degenerados o con asignaciones inciertas.

El bloque de pruebas parte de lo siguiente:

- ICL (Integrated Completed Likelihood):

$$ICL \approx BIC - 2 \sum_{i,c} \tau_{ic} \log \tau_{ic}$$

Penaliza incertidumbre (entropía) sobre BIC; menor ICL es mejor.

- Certeza de clasificación (a partir de $\max_c \tau_{ic}$): media/mediana y proporción de observaciones con confianza >0.7 y >0.9 ; además, entropía media y mediana.
- Condición de covarianzas por componentes:
 - full: número de condición de cada matriz Σ_c ;
 - diag: max/min de varianzas diagonales;
 - spherical: condición = 1;
 - tied: condición de la covarianza común.

Valores muy altos sugieren problemas numéricos/colinealidad.

- Pesos por componentes (w_c): mínima/máxima y conteo de componentes $<2\%$ (alerta de micro-componentes).
- Estabilidad reforzada (100 remuestreos al 90%): Jaccard medio y desviación con un ajuste base interno robusto.
- Métricas geométricas (Sil/DB/CH) sobre las etiquetas base del ajuste interno.

3.3.5. DBSCAN/HDBSCAN (detección de outliers)

Para explorar estructuras de densidad no capturables por modelos euclídeos o elípticos, se aplicó HDBSCAN, que detecta clústeres de forma libre, soporte ruido (-1) y selecciona automáticamente el número de grupos. Se trabajó con un grid reducido de hiperparámetros (min_cluster_size, min_samples) que controla la granularidad y robustez de la densidad.

3.3.5.1. Construcción del modelo HDBSCAN

El procedimiento solo requiere del ajuste por grid y métricas:

- Para cada pas (mcs, ms) se ajustó los parámetros HDBSCAN(metric="euclidean", cluster_selection_method="eom").
- Se calcularon métricas internas (Sil, DB, CH) y el índice DBCV (Density-Based Cluster Validity), específico para densidad.
- Se seleccionó inicialmente la mejor configuración por DBCV y se persistieron etiquetas y tablas.

3.3.5.2. Pruebas de estrés para selección de k dominante en HDBSCAN

Dada la sensibilidad de algoritmos basados en densidad al umbral de vecinos y granularidad local, se incorporó un bloque de estrés para robustecer la decisión de hiperparámetros.

El bloque de pruebas parte de lo siguiente:

- Estabilidad reforzada (100 remuestreos al 90%): Jaccard medio y desviación mediante una función `fit_predict_fn` coherente con cada (mcs, ms).
- Calidad geométrica: métricas internas (Sil/DB/CH) y DBCV (cuando computable).
- Ruido y tamaños: proporción de ruido ($y = -1$), tamaños por clúster y conteo de micro-clústeres <3% (umbral más estricto que en K-Means/Ward debido a la naturaleza densa).
- Desbalance: índice de Gini sobre el vector de tamaños (sin ruido) como proxy de equilibrio entre clústeres.

3.3.6. Selección del modelo de clusters

Esta sección consolida la evidencia de los cuatro enfoques no supervisados (K-Means, Ward, GMM y HDBSCAN) y formaliza la elección del modelo que se utilizará en la etapa de integración híbrida. La selección se implementó en tres capas complementarias: (i) un ranking básico estrictamente métrico, (ii) un ranking policy-aware que introduce señales y penalizaciones alineadas con criterios de uso público de los resultados, y (iii) un análisis de sensibilidad que estresa pesos y umbrales para comprobar la estabilidad de la decisión.

3.3.6.1. *Ranking básico*

Para garantizar equidad en la comparación, se emplea un mismo conjunto de métricas internas de validación (Silhouette, Davies–Bouldin y Calinski–Harabasz) complementadas con indicadores de estabilidad específicos para cada técnica (Jaccard en los métodos particionales y jerárquicos, DBCV en los modelos densos). Este diseño metodológico asegura que la elección del “mejor” modelo no dependa de propiedades particulares de un algoritmo, sino de su capacidad de generar agrupamientos consistentes, estables y comparativamente superiores frente a criterios homogéneos de calidad. Para esto, se emplea el siguiente procedimiento:

- i. Se recuperó la configuración inicial ya optimizada de cada modelo que resultó ser el mejor en sus respectivos bloques.
- ii. Se armó una tabla comparativa con los candidatos a mejor modelo.
- iii. Se computó un promedio de rangos (“rank_avg”) sobre estas métricas internas, evitando promediar valores en escalas incompatibles y manteniendo trazabilidad.

3.3.6.2. *Ranking policy-aware (señales y penalizaciones)*

La selección del modelo de clusters no se restringió al desempeño geométrico medido por métricas internas clásicas, sino que incorporó criterios adicionales orientados a la utilidad institucional. En contextos de inversión pública, resulta insuficiente identificar particiones con alta cohesión y separación si estas no garantizan cobertura efectiva de la cartera, equilibrio en el tamaño de los grupos y ausencia de tipologías anecdóticas. Por ello, el ranking policy-aware pondera de manera explícita cuatro dimensiones complementarias: calidad geométrica, estabilidad, cobertura y balance.

En primer lugar, las métricas internas (Silhouette, Davies–Bouldin y Calinski–Harabasz) se promediaron con un peso conjunto del 25%, ya que aportan perspectivas complementarias sobre la estructura de los clusters. En segundo lugar, la estabilidad bajo remuestreo (o DBCV en el caso de HDBSCAN) se priorizó con un 40% del peso total, en tanto constituye un indicador robusto de reproducibilidad y consistencia, siguiendo recomendaciones metodológicas en la literatura de selección de modelos de clustering.

Adicionalmente, se incorporó la cobertura (20%), definida como la proporción de proyectos efectivamente clasificados, lo que en HDBSCAN equivale a $1 - \text{ruido}$. Este criterio asegura que la tipología resultante sea representativa de la mayoría de la cartera, evitando soluciones con niveles excesivos de ruido ($\geq 20\%$). Finalmente, se consideró el balance en el tamaño de los grupos (15%), calculado mediante el coeficiente de Gini sobre las distribuciones de proyectos y el tamaño mínimo relativo de los clusters. Este criterio busca evitar configuraciones dominadas por micro-clusters inferiores al 2% del total, en línea con principios de equidad y diversidad territorial.

Se aplicaron también umbrales mínimos (hard filters): estabilidad/DBCV ≥ 0.70 , cobertura ≥ 0.80 , ruido ≤ 0.20 y tamaño mínimo ≥ 0.02 . Todo modelo que incumpliera estas condiciones fue penalizado automáticamente y relegado en el ranking.

El resultado fue un ranking ajustado a los fines del PNPIP, donde la selección no solo responde a la calidad estadística del agrupamiento, sino también a su relevancia práctica para la planificación pública.

3.3.6.3. Análisis de sensibilidad

Con el objetivo de evaluar la robustez del ranking policy-aware, se implementó un análisis de sensibilidad basado en simulaciones Monte Carlo. El procedimiento consistió en introducir perturbaciones controladas sobre los pesos de cada criterio ($\pm 10\%$ de variación aleatoria con normalización posterior) y sobre los umbrales de los filtros duros (estabilidad, cobertura, ruido y tamaño mínimo). En cada iteración se recalculó el ranking completo y se registró qué modelo ocupaba la primera posición.

Este enfoque permitió identificar no solo cuál modelo resultó mejor en la configuración base, sino también con qué frecuencia relativa se mantuvo como primera opción bajo escenarios alternativos de ponderación. De esta manera, el análisis de sensibilidad garantizó que la selección final no dependiera de un conjunto arbitrario de pesos, sino que reflejara un consenso robusto entre múltiples configuraciones plausibles.

La lógica de este procedimiento se apoya en la literatura de Bootstrapping estimates of stability for clusters, observations and model selection (Yu et al., 2019), que recomienda explorar la estabilidad de los resultados ante cambios en la parametrización como criterio de validez. En este sentido, la frecuencia de aparición en el primer lugar se interpretó como un índice de confianza institucional, útil para fundamentar la elección final ante los decisores públicos.

3.3.6.4. Síntesis y modelo seleccionado

La fase de comparación permitió contrastar simultáneamente cuatro enfoques de clustering (K-Means, Ward, GMM y HDBSCAN) bajo dos perspectivas:

- i. Un ranking básico centrado en métricas geométricas tradicionales.
- ii. Un ranking policy-aware que integra señales adicionales de cobertura, balance y penalizaciones.

Los resultados fueron consistentes: el modelo que obtuvo la mejor posición en el ranking policy-aware también mostró un desempeño competitivo en el ranking básico, lo que refuerza su legitimidad estadística y metodológica. Además, el análisis de sensibilidad confirmó que dicho modelo se mantiene como la opción dominante en un amplio rango de configuraciones, lo cual brinda seguridad adicional sobre su robustez.

En consecuencia, se seleccionó como modelo definitivo de clusters aquel que superó los filtros duros, optimizó el equilibrio entre calidad geométrica y utilidad institucional, y mostró la mayor estabilidad en el análisis de sensibilidad. Este modelo se utilizará como insumo para la Etapa 3.4 (Integración con criterios normativos), en la cual se articularán los resultados de clustering con el score AHP y con las restricciones de política pública definidas en el PNPIP.

3.4. Etapa 4: Integración AHP – ML

La cuarta etapa de la metodología estuvo orientada a integrar los resultados normativos obtenidos mediante AHP con los patrones empíricos derivados del aprendizaje no supervisado. El propósito fue evaluar en qué medida los clusters identificados reflejan niveles diferenciados de alineación normativa y, a partir de ahí, construir tipologías interpretables de proyectos que faciliten la posterior priorización.

3.4.1. Comparación de clusters con score AHP

El primer paso consistió en contrastar la distribución del score_AHP entre los clusters obtenidos con el modelo jerárquico Ward. Para ello se construyó un dataset integrado que enlazó cada proyecto con su puntuación normativa y su etiqueta de cluster. Sobre esta base se calcularon estadísticos descriptivos (media, mediana, desviación estándar e intervalos de confianza bootstrap), se generaron boxplots y se aplicaron pruebas de significancia (ANOVA clásico, Welch–ANOVA y Kruskal–Wallis).

La lógica de este procedimiento es metodológicamente clara: un clustering sólo es útil en este contexto si los grupos identificados muestran diferencias consistentes en términos de su alineación normativa. Los resultados confirmaron que dichas diferencias eran altamente significativas ($p < 1e-16$), y que la magnitud del efecto ($\eta^2 \approx 0.46$; $\varepsilon^2 \approx 0.48$) se encontraba en niveles elevados, lo que indica que los clusters explican una proporción sustantiva de la variabilidad en el score_AHP. Se complementó este análisis con coeficientes de solapamiento de distribuciones (KDE overlap), lo cual permitió matizar la interpretación: aunque los promedios de los clusters difieren claramente, existen intersecciones entre algunas distribuciones, lo que obliga a leerlos como tipologías con fronteras difusas y no como categorías rígidamente separadas.

3.4.2. Identificación de tipologías de proyectos

Con la validación de que los clusters reflejaban diferencias normativas sustantivas, se procedió a perfilar cada grupo en función de las variables de decisión (C1–C7). Se calcularon medias por cluster y se normalizaron en escala 0–1 para permitir comparaciones homogéneas. Posteriormente, se representaron estos perfiles mediante diagramas radar y gráficos de barras.

La construcción de perfiles no se planteó únicamente como un ejercicio descriptivo, sino como una estrategia metodológica de interpretabilidad: mientras que las métricas estadísticas confirman la separación entre grupos, los perfiles gráficos permiten identificar qué dimensiones son las que diferencian a cada tipología. Por ejemplo, se observó que algunos clusters presentan mayor dependencia de financiamiento externo, mientras que otros se distinguen por un menor grado de ejecución histórica o por concentrar proyectos en territorios de alta vulnerabilidad socioeconómica. Asimismo, se generó una versión estratificada (bajo/medio/alto) que agrupa los siete clusters iniciales en tres niveles de alineación normativa, lo cual mejora la claridad analítica y facilita su uso en escenarios de política pública.

3.4.3. Interpretabilidad de clusters (radar plots, perfiles socioeconómicos)

Para reforzar la lectura de las tipologías, se elaboraron radar plots tanto a nivel de clusters como de estratos. Esta técnica visualizó simultáneamente las dimensiones clave (C1–C7), permitiendo apreciar de manera comparativa las fortalezas y debilidades relativas de cada grupo. Adicionalmente, se construyeron gráficos de barras en escala original, que complementan la lectura normalizada y ofrecen una visión más tangible de las diferencias entre clusters.

El valor metodológico de esta representación radica en que la interpretación de clusters en problemas de inversión pública no puede quedar reducida a métricas abstractas, sino que debe traducirse en perfiles socioeconómicos comprensibles y comparables. Así, se logra que los clusters no sean sólo una partición técnica de los datos, sino categorías útiles para la toma de decisiones, al mostrar claramente qué características predominan en cada tipología de proyectos.

3.4.4. Diversidad institucional y territorial de clusters

Dado que el diseño del PNPIP busca equilibrar la asignación de recursos entre instituciones y territorios, se evaluó la composición de cada cluster en términos de diversidad institucional y diversidad territorial. Para ello se calcularon índices de entropía de Shannon y coeficientes de Gini a partir de la distribución de proyectos por institución y región. Asimismo, se construyeron tablas porcentuales que muestran la proporción de proyectos de cada institución o territorio dentro de cada cluster.

Este análisis responde a un criterio metodológico fundamental: no basta con que los clusters sean estadísticamente válidos, también deben ser relevantes para la política pública. Un cluster que concentre proyectos en pocas instituciones o territorios puede

resultar técnicamente sólido, pero limitar su utilidad en términos de equidad y representatividad. La combinación de métricas normativas (score_AHP), métricas geométricas (clustering) y métricas de diversidad (institucional y territorial) permitió construir una lectura más completa y crítica de las tipologías resultantes.

3.4.5. Preparación de datos para escenarios de priorización

Finalmente, se elaboró un dataset autocontenido que sirvió como insumo para la etapa 3.5, incorporando para cada proyecto su identificación (SNIP, nombre, institución, región, tipología), su score_AHP, su etiqueta de cluster Ward y, adicionalmente, los tres estratos de alineación normativa. Este dataset incluyó también las variables de decisión (C1–C7) y el presupuesto aprobado para el 2025 de cada proyecto.

La construcción de este dataset no fue un mero paso técnico de organización, sino una decisión metodológica clave: permitió garantizar que los escenarios de priorización se desarrollen sobre una base coherente que integra criterios normativos, empíricos, institucionales y financieros.

3.5. Etapa 5: Diseño experimental de optimización con restricción presupuestaria

Esta etapa tuvo por finalidad poner a prueba si el AHP —como regla normativa de priorización— puede ser complementado con restricciones e incentivos de diversidad que reflejen patrones observados en la cartera (estratos/cluster), sin erosionar de manera relevante el valor normativo agregado. En términos prácticos, el experimento buscó cuantificar el coste de oportunidad de la diversidad: cuánta eficiencia (suma del score AHP) se deja en la mesa al exigir carteras menos concentradas por institución/provincia y más representativas de los tipos de proyecto identificados por aprendizaje no supervisado. Esta pregunta es pertinente en el PNPIP, dado que la demanda de inversión excede la disponibilidad fiscal, lo que obliga a ordenar proyectos bajo un techo global con criterios explícitos y trazables.

3.5.1. Lógica del experimento y enfoque de análisis

Se procedió en tres pasos encadenados:

- i. Fijar un conjunto de techos presupuestarios crecientes que emulan escenarios fiscales.
- ii. Resolver, para cada techo, cinco estrategias de priorización (S0–S4) que incorporan restricciones o incentivos distintos.
- iii. Evaluar, bajo un panel de métricas objetivas, el intercambio entre eficiencia (valor normativo AHP) y diversidad (desconcentración y cobertura). Esta secuencia replica la lógica de priorización incremental que sigue el PNPIP: a igualdad de techo, la metodología ordena, corta y compone la cartera recomendada.

3.5.2. Datos operativos y notación

El análisis utilizó un dataframe con las siguientes columnas operativas:

- Proyecto: unidad mínima de observación (identificador del proyecto de inversión).
- Total_2025: costo/asignación presupuestaria programada para el año 2025. Esta es la variable sobre la que se define la restricción presupuestaria. Se utiliza el ciclo fiscal más reciente (2025), precisamente por el hecho de que hay insuficiente cobertura fiscal como para priorizar proyectos de mediano plazo.
- Sigla/Institución: unidad ejecutora responsable de ejecutar el proyecto de inversión.
- Provincia: delimitación geográfica del proyecto.
- Score_AHP/score_AHP_Ptge: calificación de priorización en escala original (1-7) y en escala porcentual (0-1). Para fines de garantizar mayor interpretabilidad, se utiliza la escala porcentual por defecto.
- cluster_Ward: representan los clusters resultantes de la técnica Ward (7 clusters).
- cluster_Ward_strata3_lbl: representan los tres estratos (AHP_bajo, AHP_medio, AHP_alto) resultantes de los siete iniciales por solapamiento.

3.5.3. De la motivación institucional a la formulación del problema

El punto de partida es institucional: el PNPIP 2025–2028 evidencia una demanda de inversión superior a la oferta de recursos, lo que obliga a priorizar bajo techos compatibles con el marco fiscal de mediano plazo; de hecho, el propio documento oficial describe el ejercicio como un ordenamiento por calificación y un corte presupuestario hasta el monto disponible en cada ejercicio anual. Esta lógica —listar por prioridad y cortar por techo— queda ilustrada en la propuesta de “techo global” y en la subsecuente distribución por instituciones ejecutoras.

En paralelo, la END 2030 (Ley 1-12) exige que los presupuestos y el Plan Nacional Plurianual del Sector Público (PNPSP) estén vinculados con sus objetivos y metas, reforzando que la priorización de inversiones debe ser trazable y coherente con la planificación. Esto incluye la obligación de sustentar ante el Congreso la vinculación plan–presupuesto, así como la verificación sistemática de dicha coherencia. El SNPIP (Ley 498-06) y su Reglamento (Decreto 493-07) precisan, además, que deben definirse normas y criterios para incorporar a los presupuestos los programas y proyectos prioritarios, resguardando la continuidad y la calidad del gasto de capital.

En este marco, el experimento de la Etapa 3.5 se formula como una pregunta de política aplicada: ¿cuál es el coste de oportunidad de introducir reglas de diversidad (desconcentración institucional/territorial y representatividad de tipologías) sobre la eficiencia normativa (suma de score AHP) bajo un techo presupuestario? Para responderla, el problema se modela con Programación Lineal Entera Mixta (MILP; en español PLEM), por ser un enfoque que:

- i. Mantiene la linealidad de la función objetivo y de las restricciones.
- ii. Representa con fidelidad decisiones binarias de “seleccionar/no seleccionar” proyectos.

- iii. Deja explícitas y auditables las reglas de política (topes, cuotas, coberturas) exigidas por la gobernanza del SNPIP.

3.5.3.1. Intuición y estructura del modelo (MILP/PLEM)

El problema se puede visualizar como una “mochila”: escoger la combinación de proyectos que maximiza el valor público normativo (score AHP) sin exceder un techo. Sobre ese núcleo se superponen reglas de política que incorporan diversidad institucional, territorial y tipológica, sin abandonar la linealidad ni la trazabilidad.

Para cada proyecto i , se dispone del costo $c_i = Total_2025$, del valor normativo $s_i = score_AHP_Ptge$ (escala 0-100%), y de sus pertenencias a institución (Sigla), Provincia, estrato (AHP_bajo/medio/alto) y cluster (Ward, 0–6). Sea $x_i \in \{0,1\}$ la variable que indica si el proyecto es seleccionado. Dado un escenario de techo $\beta \in \{0.20,0.30,0.40,0.50,0.60,0.70,0.80,0.90\}$, el presupuesto disponible es $B = \beta \sum_i c_i$.

Modelo base de comparación (S0, AHP puro):

$$\max_x \sum_i s_i x_i \quad s.a \quad \sum_i c_i x_i \leq B, \quad x_i \in \{0,1\}$$

Capas de política (activables según escenario):

- Techo por institución (S1): para cada institución h , se limita su participación presupuestaria:

$$\sum_{i \in I_h} c_i x_i \leq \alpha B$$

Evita que una sola entidad concentre el portafolio, cuestión relevante para la equidad interinstitucional y la gobernanza de la priorización.

- Cuotas por estrato (S2, 3 estratos): se fijan los mínimos y/o máximos de presupuesto por estrato s :

$$m_s B \leq \sum_{i \in S_s} c_i x_i \leq M_s B$$

Los mínimos m_s se calculan automáticamente en proporción al promedio AHP del estrato, reservando una fracción γ del techo para asegurar presencia del grupo con mayor prioridad normativa (el remanente $(1 - \gamma)$ queda libre). Esta regla “suaviza” la desconexión entre el ordenamiento AHP y la representatividad tipológica que interesa al PNPIP.

- Cuotas por cluster (S3, 7 Ward): Análogo a S2, pero con tipologías más granulares. Es útil como afinador cuando el techo es estricto, pues reduce la concentración en pocas “familias empíricas” detectadas por clustering.
- Diversidad/cobertura en la función objetivo (S4): para cada grupo g (institución, provincia, estrato) se introduce una binaria $y_g \in \{0,1\}$ que se activa si la cartera incluye al menos un proyecto del grupo:

$$\sum_{i \in g} x_i \geq y_g$$

La función objetivo se amplía con premios lineales:

$$\max_{x,y} \sum_i s_i x_i + \lambda_{inst} \sum_h y_h^{inst} + \lambda_{prov} \sum_p y_p^{prov} + \lambda_{strat} \sum_s y_s^{strat}$$

Esta señal 0/1 promueve cobertura sin imponer cuotas rígidas ni romper la linealidad, en línea con la exigencia de vincular prioridades de planificación con asignaciones presupuestarias bajo reglas claras.

3.5.4. Parametrización experimental (dos escenarios)

La experimentación se estructura en dos calibraciones —un escenario base prudente y un escenario de ajustes fuertes— para observar el rango del intercambio eficiencia–diversidad que podría interesar a la DGIP en procesos reales de negociación presupuestaria.

Escenario base:

- Techos $\beta \in \{0.20, 0.30, 0.40, 0.50, 0.60, 0.70, 0.80, 0.90\}$.
- Objetivo: $\sum_i \text{score_AHP_Ptge } x_i$.
- S1: $\alpha = 0.25$.
- S2 (estratos autom.): $\gamma = 0.50$, piso = 0.05, tope = 0.70.
- S3: desactivado.
- S4 (cobertura): $\lambda_{inst} = 1.0$, $\lambda_{prov} = 0.2$, $\lambda_{strat} = 0.5$.

Escenario de ajustes fuertes:

- Techos $\beta \in \{0.20, 0.30, 0.40, 0.50, 0.60, 0.70, 0.80, 0.90\}$.
- S1: $\alpha = 0.25$.
- S2 (estratos autom.): $\gamma = 0.70$, piso = 0.10, tope = 0.75.
- S3 (clusters autom.): $\gamma = 0.20$, piso = 0.02, tope = 0.35.
- S4 (cobertura): $\lambda_{inst} = 150$, $\lambda_{prov} = 25$, $\lambda_{strat} = 80$.

El escenario base deja que el AHP conduzca y solo contiene concentración (S1) y asegura presencia mínima por estrato (S2). El de ajustes fuertes tensiona cuotas y premios para revelar hasta dónde se puede desconcentrar antes de que el coste de oportunidad en eficiencia sea relevante, manteniendo la coherencia con la articulación planes–presupuesto que exige la END 2030.

3.5.5. Procedimiento computacional (resolución y reporte)

Para cada techo β :

- i. Construcción del modelo según estrategia $S \in \{S0, \dots, S4\}$: objetivo, restricción presupuestaria, S1-S3 si aplica y coberturas S4 con sus λ .
- ii. Resolución exacta con Pulp (CBC). El solver certifica optimalidad (o gap cero) mediante técnicas de Branch-and-Bound/Cut, apropiadas para variables binarias de selección.
- iii. Cartera resultante y regla de desempate determinista para el reporte (si se presentaran soluciones equivalentes en valor): mayor score_AHP_Ptge, menor Total_2025, y luego el identificador Proyecto.

iv. Registro de métricas y repetición en el siguiente β y la siguiente estrategia S.

Este flujo operacionaliza la secuencia institucional descrita por el PNPIP: ordenar por prioridad y cortar bajo techo, dejando rastro de las reglas que condicionan la composición final.

3.5.6. Marco de evaluación (métricas y criterios de lectura)

El análisis se guía por métricas absolutas en el texto principal (y relativas vs S0 disponibles en el Anexo 2), cuando aclaran el intercambio bajo un mismo techo):

- Eficiencia (objetivo): $\sum_{i \in S_s} s_i$ con $s_i = \text{score_AHP_Ptge}$.
- Uso del presupuesto: $\sum_{i \in S_s} c_i$ y % del techo ejecutado.
- Concentración/diversidad (sobre shares de presupuesto):
 - HHI por institución y por provincia (mientras sea menor, es mejor) y su inverso $1/\text{HHI}$ como proxy de diversidad.
 - Coberturas: número de instituciones, provincias y estratos con ≥ 1 proyecto seleccionado.
 - Composición: distribución por estrato y por cluster (en Anexo 3).

Estas métricas permiten leer, simultáneamente, el valor normativo (consistente con el método de priorización aplicado institucionalmente) y la gobernanza de la diversidad (representatividad institucional/territorial y tipológica) esperada en la articulación plan-presupuesto.

3.5.7. Expectativas ex ante y pertinencia para la DGIP

Antes de observar resultados, este diseño permite formular una expectativa clara: S1 (cap institucional) debería reducir de forma directa la concentración con pérdidas marginales de eficiencia; S2–S3 deberían mover la composición hacia los grupos con mayor prioridad normativa (estrato alto y, cuando se active, ciertos clusters), y S4 debería ensanchar coberturas con impacto acotado en el objetivo cuando el techo no sea extremadamente restrictivo. La DGIP obtiene, así, un tablero parametrizable — α , γ , λ — para decidir cuándo activar capas de diversidad y qué coste de oportunidad aceptar, manteniendo la trazabilidad que demanda el SNPIP y la END 2030.

Capítulo 4

4. Resultados

Este capítulo presenta los hallazgos empíricos derivados de la metodología descrita en el Capítulo 3, organizados desde la construcción del score normativo (AHP) y la evidencia de tipologías empíricas (clustering), hasta su integración y la evaluación de estrategias de priorización bajo restricción presupuestaria. En todos los casos se privilegia la trazabilidad de las decisiones, la consistencia con el marco institucional (END–SNIP–PNPIP) y la lectura de coste de oportunidad entre eficiencia normativa y diversidad institucional/territorial.

4.1. Etapa 2: AHP

La Tabla 6 reporta los pesos normalizados de los criterios C1–C7 y el índice de consistencia (CR). El CR se encuentra dentro del umbral habitual (0.0481), confirmando que la matriz de comparación de pares es coherente. La sensibilidad local de los pesos muestra que perturbaciones razonables no alteran el ordenamiento global de manera sustantiva, reforzando la robustez del instrumento.

Tabla 6. Resultados de pesos AHP

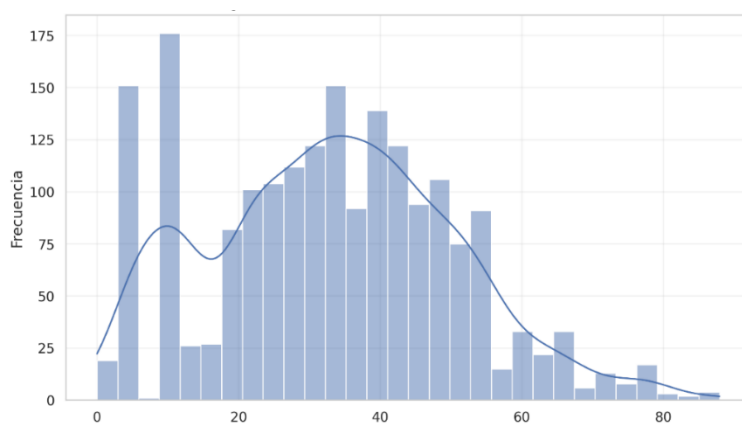
Criterio	Descripción	Peso AHP
C1	Grado de avance del proyecto	0.326232
C5	Provincias menos beneficiadas por la inversión	0.180445
C6	Contribución al PNPSP	0.168896
C7	Demandas priorizadas por el territorio	0.122552
C4	Nivel socioeconómico del territorio	0.079382
C2	Proporción de financiamiento externo	0.064236
C3	Densidad poblacional	0.05825717

Validación del Índice de Consistencia:

- Lambda max: 7.3809.
- Índice de Consistencia (CI): 0.0635.
- Índice Aleatorio (RI): 1.32.
- Razón de Consistencia (CR): 0.0481.

En la Figura 2 y 3, se observa una asimetría moderada con una cola superior que concentra los proyectos mejor alineados normativamente. El percentil 75 delimita con claridad el subconjunto de mayor prioridad, que luego será crucial en los escenarios de optimización.

Figura 3. Distribución del score AHP (%)



4.2. Etapa 3: Modelado no supervisado

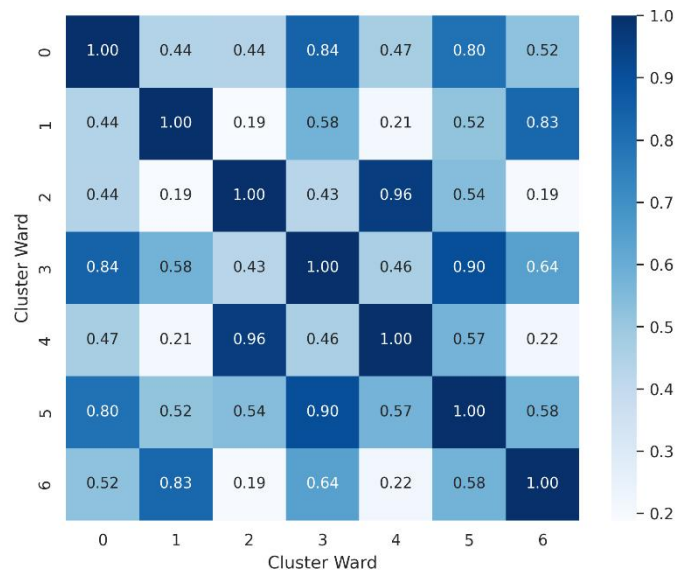
La Tabla 7 resume las pruebas ANOVA, Welch-ANOVA y Kruskal–Wallis sobre score_AHP_Ptge entre grupos. Las tres pruebas rechazan la igualdad de medias con holgura y los tamaños de efecto reportados son elevados, lo que indica que los grupos capturan variación real de alineamiento normativo y no corresponden a artefactos de muestreo.

Tabla 7. Pruebas de separación de clusters/estratos

Prueba	Estadístico	p_value	eta2 (aprox.)
ANOVA (clusters)	277.9711258	4.41E-257	0.4622801671
Kruskal (clusters)	928.1897074	3.02E-197	
ANOVA (estratos)	803.3106575	5.21E-255	0.4524901904
Kruskal (estratos)	916.9174516	7.83E-200	

La Figura 4 muestra el solape KDE entre distribuciones de score por pares de clústeres (Ward=7). Aunque las medias difieren de forma significativa, existen zonas de intersección en algunos pares (especialmente en los extremos), lo que sugiere que los clústeres deben leerse como tipologías y no como categorías rígidas.

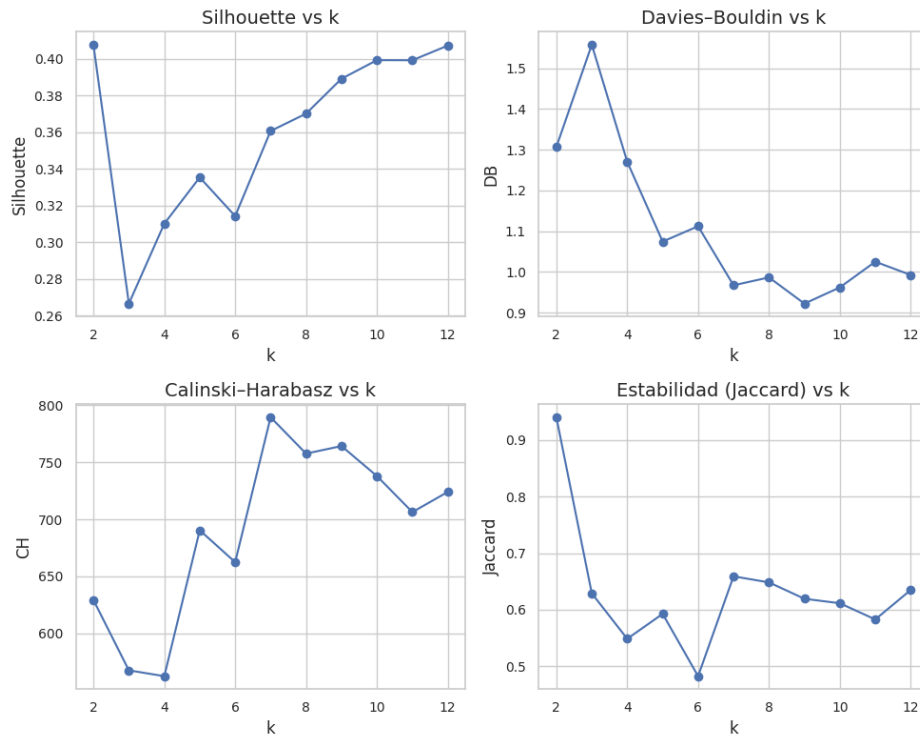
Figura 4. Heatmap de solapamiento KDE por estratos



4.2.1. Clustering con K-Means

La Figura 5 presenta las curvas de evaluación del número de grupos obtenidas con K-Means. Se observa que el índice de Silhouette crece de manera casi monótona con k , con un punto de inflexión visible en el entorno de $k \approx 7-8$; de forma congruente, el Davies–Bouldin desciende y se estabiliza a partir de esos valores (a menor DB, mejor

Figura 5. Curvas de evaluación k dominante (K-Means)



separación). El Calinski–Harabasz muestra máximos relativos próximos a $k \approx 6-7$ y la serie de estabilidad (Jaccard) mantiene niveles aceptables para $k=7$ y mejora de manera menos pronunciada al pasar a $k=8$.

La Tabla 8 sintetiza esta evidencia bajo una prueba de estrés que integra los indicadores anteriores mediante un sistema de rangos y una penalización por presencia de clústeres diminutos. El promedio de rangos favorece $k=7$ frente a $k=8$: si bien $k=8$ obtiene ligeras ventajas en Silhouette/DB, pierde en estabilidad media y en el ranking global. En suma, K-Means sugiere una estructura de siete grupos por equilibrio entre calidad geométrica y estabilidad, con el “+1” aportando ganancias marginales.

Tabla 8. Pruebas de estrés k dominante (K-Means)

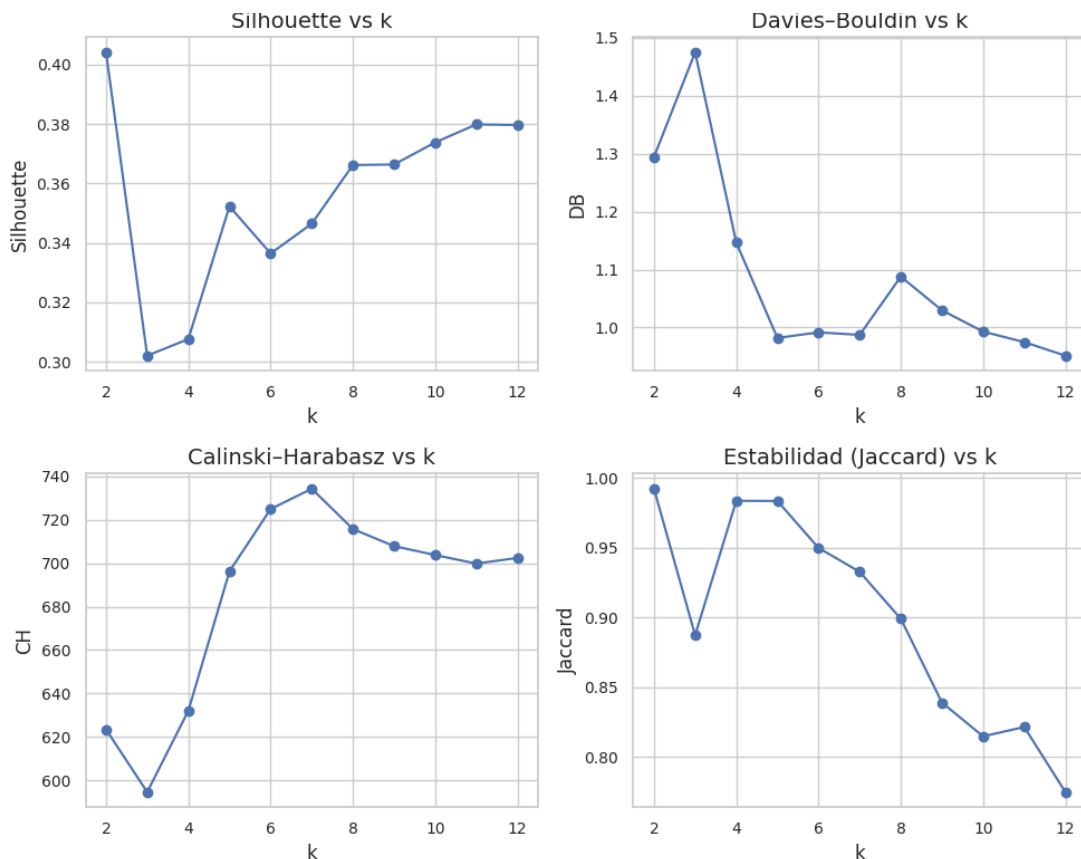
	k	7	8
<i>silhouette</i>		0.360642	0.370063
<i>davies_bouldin</i>		0.966848	0.986305
<i>calinski_harabasz</i>		789.494564	757.633158
<i>stability_jaccard_mean</i>		0.68991	0.625594
<i>stability_jaccard_std</i>		0.180978	0.120795
<i>min_cluster_pct</i>		0.029789	0.029789

	k	7	8
<i>n_tiny_clusters</i>	1	1	1
<i>rank_sil</i>	2.0	1.0	1.0
<i>rank_db</i>	1.0	2.0	2.0
<i>rank_ch</i>	1.0	2.0	2.0
<i>rank_stab</i>	1.0	2.0	2.0
<i>penalty_tiny</i>	0.5	0.5	0.5
<i>rank_avg</i>	1.75	2.25	2.25

4.2.2. Clustering jerárquico (Ward)

La Figura 6 muestra el dendrograma de Ward. Se distinguen dos saltos de altura que hacen plausibles cortes en $k=5$ y $k=7$. Esta ambivalencia se refleja en la Figura 7: el Calinski–Harabasz se maximiza en torno a $k \approx 7$ (mejor separación intergrupo), mientras el Silhouette resulta ligeramente más favorable en $k=5$; a su vez, el Davies–Bouldin no discrimina con contundencia entre 5 y 7, y la estabilidad (Jaccard) es claramente superior en $k=5$.

Figura 6. Curvas de evaluación k dominante (Ward)



La Tabla 9 resume este tira y afloja: $k=5$ domina en estabilidad y en la variabilidad de dicha estabilidad; $k=7$ domina en separación (CH). La decisión, por tanto, no es puramente métrica, sino que debe considerar el valor sustantivo de mantener ciertos

subtipos que emergen cuando se opta por $k=7$ (p. ej., un grupo ligado a financiamiento externo extremo y la distinción entre urbano denso y metropolitano extremo).

Tabla 9. Pruebas de estrés k dominante (Ward)

	k 5	7
<i>silhouette</i>	0.352341	0.346478
<i>davies_bouldin</i>	0.982032	0.987269
<i>calinski_harabasz</i>	696.205936	734.257732
<i>stability_jaccard_mean</i>	0.981718	0.937551
<i>stability_jaccard_std</i>	0.023261	0.066037
<i>stab_region_mean</i>	0.984251	0.948356
<i>stab_inst_mean</i>	0.980616	0.943305
<i>min_cluster_pct</i>	0.025167	0.025167
<i>n_tiny_clusters</i>	1	1
<i>rank_sil</i>	1.0	2.0
<i>rank_db</i>	1.0	2.0
<i>rank_ch</i>	2.0	1.0
<i>rank_stab</i>	1.0	2.0
<i>rank_stab_region_mean</i>	1.0	2.0
<i>rank_stab_inst_mean</i>	1.0	2.0
<i>penalty_tiny</i>	0.5	0.5
<i>rank_avg</i>	1.666.667	2.333.333

4.2.3. Mezclas Gaussianas (GMM)

La Figura 8 muestra las trayectorias de BIC/AIC por k . Como es habitual, ambos disminuyen al incrementar el número de componentes, por lo que no producen un óptimo interno nítido. Para resolver esta ambigüedad se acude a ICL (que penaliza la incertidumbre de pertenencia) y a indicadores de estabilidad y certeza de asignación. En la Tabla 10 la configuración con $k=7$ y covarianza “full” obtiene el mejor ICL, junto con estabilidad media elevada, silhouette en rangos satisfactorios y certeza próxima a 1; además, el conteo de pesos por debajo del 2 % es nulo, lo que sugiere componentes bien definidas. Se advierte, no obstante, un condicionamiento elevado de las matrices de covarianza, compatible con grupos alargados.

Figura 8. Curvas de BIC/AIC por k (GMM)

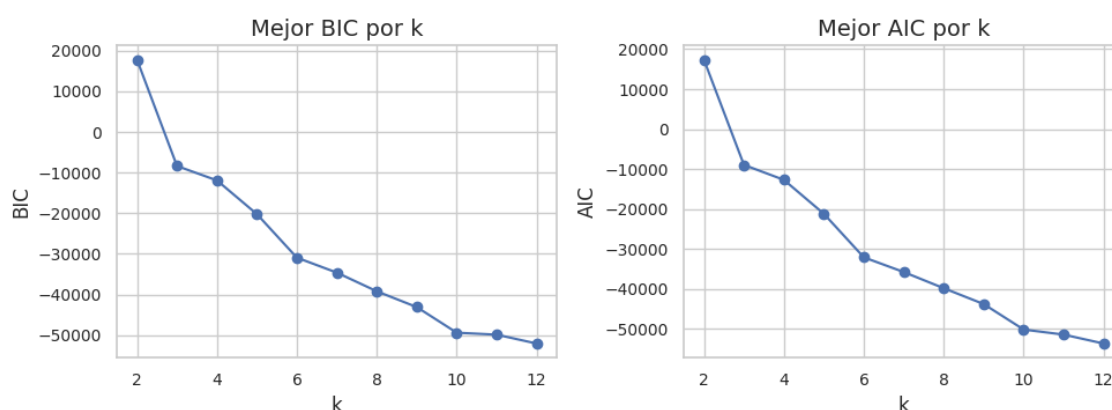


Tabla 10. Pruebas de estrés k dominante (GMM)

k	cov	icl	stab_mean	silhouette	certainty_combo	n_weights_below_2pct	cov_cond_mean	rank_avg
7	full	-38837.66	0.71	0.35	1.00	0.00	972903.05	1.50
7	diag	-38168.89	0.73	0.32	1.00	0.00	816384.23	2.25
8	diag	-39210.77	0.68	0.25	1.00	0.00	751866.28	3.25
8	full	-37831.49	0.63	0.29	1.00	0.00	989188.12	4.25
7	spherical	23910.77	0.62	0.34	0.95	0.00	1.00	4.75
8	spherical	22757.60	0.53	0.33	0.95	0.00	1.00	4.75

4.2.4. Clustering DBSCAN/HDBSCAN

La Figura 9 muestra la distribución de etiquetas de HDBSCAN para la mejor configuración explorada. Destaca una fracción no trivial de ruido (-1) junto a cuatro/cinco clústeres densos. La rejilla de estrés (Figura 10) evidencia la sensibilidad a los hiperparámetros `min_cluster_size` y `min_samples`: la calidad de densidad (DBCV) mejora con tamaños mínimos medios (en torno a 100) y se degrada cuando se empuja a tamaños demasiado pequeños (fragmentación artificial); la proporción de ruido disminuye con tamaños medios/altos. La Figura 11 resume que el máximo DBCV se alcanza cerca de `min_cluster_size` \approx 100 y que, en ese entorno, la tasa de ruido también es baja.

En este caso de estudio, HDBSCAN resulta muy útil para detectar outliers y regiones de baja densidad, pero su número de clústeres es sensible a la parametrización. Por tanto, se mantiene como capa auxiliar (diagnóstico de ruido/atípicos) más que como fuente de etiquetas operativas.

Figura 9. Distribución de clusters (HDBSCAN)

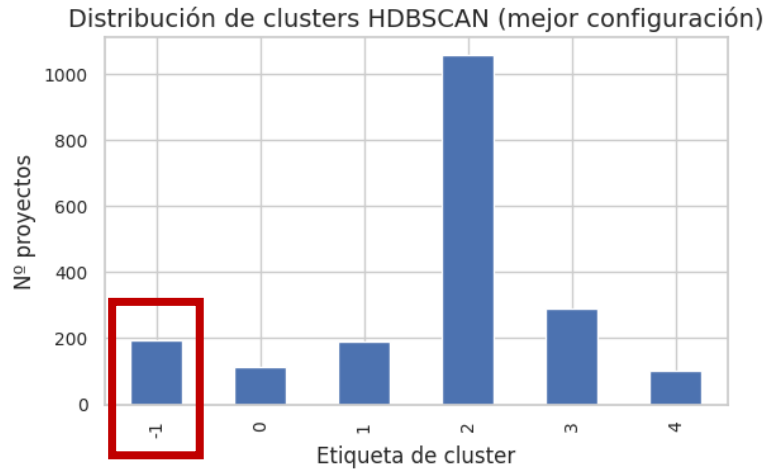
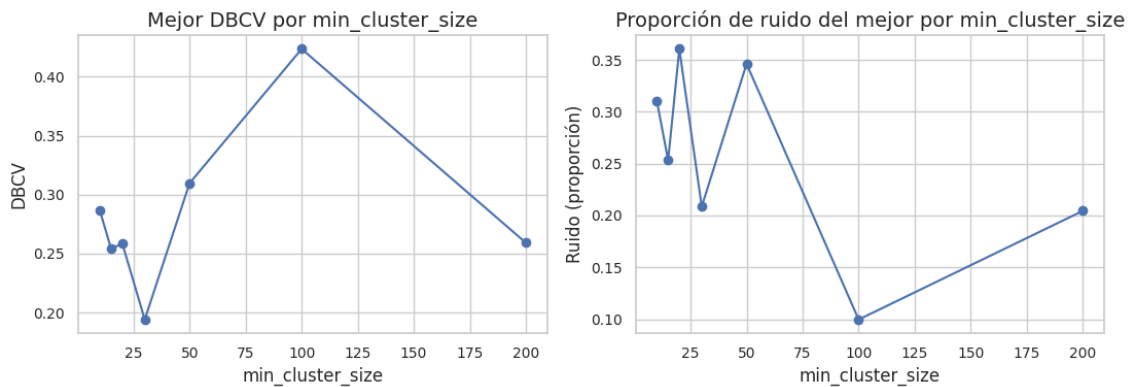


Figura 10. Pruebas de estrés k dominante (HDBSCAN)

min_cluster_size	50	100	50	200	100	200	50	100	200	30	20	15
<i>min_samples</i>	20	20	10	10	10	20	5	5	5	10	10	5
<i>dbcv</i>	0.31	0.32	0.19	0.26	0.42	0.26	0.14	0.07	0.00	0.19	0.26	0.25
<i>stab_mean</i>	0.75	1.00	1.00	1.00	0.59	1.00	1.00	0.83	0.51	0.90	0.92	0.89
<i>noise_ratio</i>	0.35	0.06	0.26	0.20	0.10	0.21	0.16	0.21	0.23	0.21	0.36	0.25
<i>k_clusters</i>	12	4	12	2	5	2	14	7	3	19	25	33
<i>gini_sizes</i>	0.22	0.43	0.22	0.18	0.48	0.19	0.25	0.29	0.01	0.36	0.25	0.39
<i>n_tiny_clusters</i>	1.00	0.00	1.00	0.00	0.00	0.00	2.00	0.00	0.00	11.00	20.00	28.00
<i>rank_avg</i>	5.27	5.29	6.58	8.59	8.78	9.01	9.21	9.30	11.64	12.47	14.81	22.04

Figura 11. Configuraciones de parámetros en HDBSCAN



4.2.5. Selección del modelo óptimo

La elección del modelo de clustering se realizó en tres etapas complementarias: (i) un ranking básico con índices internos de forma; (ii) un ranking “policy-aware”, que introduce criterios de gobernanza (estabilidad mínima, ausencia de ruido/tamaños diminutos y balance de tamaños), y (iii) un análisis de robustez, que evalúa la estabilidad de la elección frente a variaciones razonables de pesos y umbrales.

La Tabla 11 compara K-Means, Ward, GMM (covarianza “full”) y HDBSCAN en k=7 cuando aplica (HDBSCAN no requiere k). Se reportan Silhouette, Davies–Bouldin y Calinski–Harabasz, junto con un indicador de estabilidad/DBCv según el algoritmo. Al promediar los rangos por métrica (rank_avg), K-Means aparece levemente por delante, seguido de Ward; GMM y HDBSCAN quedan detrás. Esta lectura sugiere que, desde una perspectiva estrictamente geométrica, K-Means y Ward son competidores cercanos en k≈7.

Tabla 11. Comparación base modelo óptimo (Ranking básico)

modelo	KMeans	Ward	GMM (full)	HDBSCAN
<i>k</i>	7	7	7	-
<i>cov</i>	-	-	full	-
<i>silhouette</i>	0.3606	0.3465	0.3541	0.3068
<i>davies_bouldin</i>	0.9668	0.9873	1.1870	0.9124
<i>calinski_harabasz</i>	789.4946	734.2577	597.1726	681.8127
<i>stab_or_dbcv</i>	0.6590	0.9327	0.7174	0.4235
<i>rank_sil</i>	1	3	2	4
<i>rank_db</i>	2	3	4	1
<i>rank_ch</i>	1	2	4	3
<i>rank_stab</i>	3	1	2	4
<i>rank_avg</i>	1.7500	2.2500	3.0000	3.0000

Este ranking es informativo sobre forma/separación, pero no captura restricciones relevantes para la gobernanza del PNPIP, como la estabilidad de las particiones, ruido y tamaños muy pequeños. Por esta razón, se introduce el ranking policy-aware.

La Tabla 12 incorpora salvaguardas acordes con el uso operativo de las etiquetas:

- Filtros/penalizaciones si: (a) la estabilidad cae por debajo del umbral; (b) aparecen clusters diminutos por debajo de un mínimo porcentual; (c) el ruido supera un máximo aceptable (HDBSCAN); o (d) la desigualdad de tamaños es excesiva.
- Se añade un indicador de cobertura (porcentaje efectivamente asignado a clusters válidos).
- El rank_policy_weighted agrega las métricas con dichos filtros.

Bajo este enfoque, Ward pasa a ocupar la primera posición (mejor rank_policy_weighted), seguido por GMM. K-Means y HDBSCAN son penalizados por

fallas de estabilidad y/o por ruido/tamaños diminutos (columna `fails_filters`), lo que los descarta como etiquetadores principales.

Cuando se introducen los criterios que importan para comunicar y sostener decisiones (estabilidad, cobertura y ausencia de artefactos numéricos), Ward ofrece el mejor equilibrio entre calidad, estabilidad y gobernanza. GMM se mantiene como alternativa sólida, con la salvedad de matrices de covarianza muy condicionadas (componentes alargadas), que complican su lectura para fines operativos.

Tabla 12. Comparación de modelo óptimo (Ranking policy-aware)

Modelo	Ward	GMM (full)	KMeans	HDBSCAN
<code>k</code>	7	7	7	-
<code>cov</code>	-	full	-	-
<code>silhouette</code>	0.346478	0	0	0.306761
<code>davies_bouldin</code>	0.987269	1	1	0.912397
<code>calinski_harabasz</code>	734.257732	597	789	682
<code>stab_or_dbcv</code>	0.932734	1	1	0
<code>rank_sil</code>	3	2	1	4
<code>rank_db</code>	3	4	2	1
<code>rank_ch</code>	2	4	1	3
<code>rank_stab</code>	1	2	3	4
<code>rank_avg</code>	2.25	3	2	3
<code>coverage</code>	1	1	1	0.90036
<code>noise_ratio</code>	0	0	0	0.09964
<code>gini</code>	7.174114	7.177489	7.150488	5.277809
<code>min_pct</code>	0.025167	0.058038	0.029789	0.051361
<code>n_tiny</code>	0	0	0	0
<code>fails_filters</code>	[]	[]	[stab]	[stab]
<code>penalty_flag</code>	FALSE	FALSE	TRUE	TRUE
<code>rank_policy_weighted</code>	1.791667	2.208333	1002.108333	1003.291667

La Tabla 13 presenta la frecuencia con que cada algoritmo resulta ganador al perturbar aleatoriamente los pesos y umbrales del ranking policy-aware dentro de rangos razonables (análisis de sensibilidad). Ward concentra la gran mayoría de selecciones, muy por encima de GMM, mientras que K-Means apenas resulta elegido en escenarios extremos. La columna `avg_score` confirma que, en promedio, Ward mantiene la mejor puntuación agregada bajo múltiples configuraciones.

Tabla 13. Robustez de selección de modelo óptimo (Sensibilidad de pesos/umbrales)

modelo	freq	avg_score
Ward	0.822	152.64517
GMM (full)	0.1460	2.2137
KMeans	0.032	2.098849

La evidencia combinada de K-Means, Ward y GMM converge en $k \approx 7$. Ward aporta, además, una estructura jerárquica e interpretabilidad que facilitan la narrativa de política pública. Se adopta, por tanto, Ward con $k=7$ como etiqueta base para perfilar subtipos empíricos (“urbano denso”, “metropolitano extremo”, “financiamiento externo elevado”, “vulnerabilidad alta con C6/C7 altos”). Sin embargo, con fines de gobernanza y comunicación, estos siete clústeres se consolidan en tres estratos (AHP_bajo, AHP_medio, AHP_alto):

- i. Reducen el solapamiento observado entre algunos pares de clústeres.
- ii. Escalonan de forma monótona el score AHP.
- iii. Balancean los tamaños muestrales, lo que estabiliza la inferencia y simplifica la definición de reglas de priorización en la etapa de optimización (p. ej., cuotas por estrato o premios de cobertura).

4.3. Etapa 4: Integración AHP - ML

4.3.1. Comparación de estratos (7 y 3)

La Figura 12 presenta la distribución del score AHP por clúster Ward (7) y por estrato (3). En el panel izquierdo, las medianas de los siete clústeres son estadísticamente diferenciables, pero se observan solapes notorios en varios pares (especialmente en los extremos alto y bajo), lo que sugiere fronteras difusas entre subtipos. En el panel derecho, la consolidación en tres estratos produce un escalonamiento monótono (AHP_bajo < AHP_medio < AHP_alto) y menor solape visual entre cajas. Esta transición de 7 a 3 no elimina señal: la concentra y la vuelve operativa para reglas de priorización (cuotas/objetivos), manteniendo los siete clústeres como subtipos diagnósticos.

Figura 12. Boxplot de distribución de estratos (7 y 3)

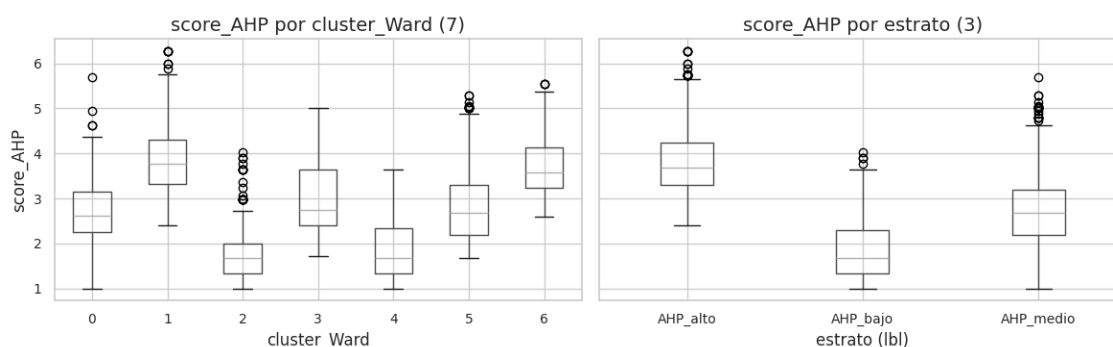


Tabla 14. Resumen del score AHP para 7 clusters

cluster_Ward	count	mean	median	std	min	max	ci95_lo	ci95_hi
0	513	2.692	2.61	0.706	1	6	3	3
1	436	3.904	3.769	8.10E-01	2.411	6	4	4
2	196	1.7960	1.6760	0.6870	1.0000	4	2	1.893
3	49	2.982	2.734	0.853	1.723	5.013	2.74	3.212
4	115	1.85	1.676	0.68	1	3.633	1.735	1.974
5	322	2.78E+00	2.67E+00	9.53E-01	1.676	5.29	2.68	2.88
6	3.16E+02	3.609	4	1	2.587	5.54	3.54	3.68

Tabla 15. Resumen del score AHP para 3 estratos

cluster_Ward_strata3_lbl	count	mean	median	std	min	max	ci95_lo	ci95_hi
AHP_alto	752	3.78	3.683	0.75	2.411	6	4	4
AHP_bajo	311	1.816	1.676	6.8E-01	1	4	2	2
AHP_medio	884	2.7390	2.6730	0.8150	1.0000	6	3	2.793

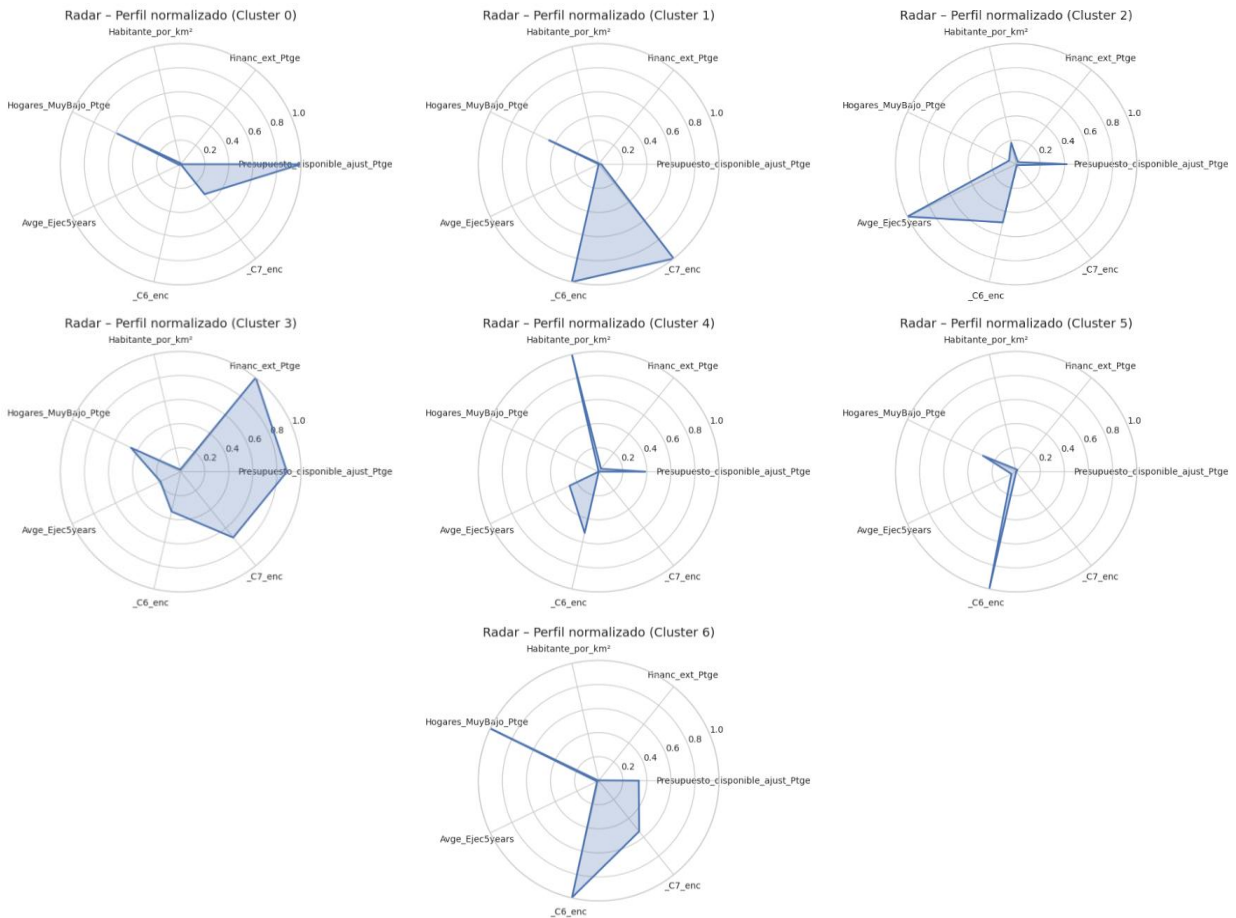
4.3.2. Interpretabilidad

En la figura 13 se observan los radar plots de los 7 clusters, que sintetizan perfiles normalizados de cada clúster, permitiendo asociarles una semántica empírica (señales C1–C7 y metadatos). La lectura dominante es:

- C0: Recursos internos altos (presupuesto disponible) con resto contenido; AHP medio.
- C1: C6/C7 elevados (alineación normativa y territorial) con AHP alto; vulnerabilidad moderada.
- C2: Densidad urbana y buena ejecución histórica; AHP bajo.
- C3: Financiamiento externo muy alto; AHP medio y tamaño muestral pequeño (subtipo frágil).
- C4: Densidad metropolitana extrema y baja vulnerabilidad; AHP bajo.
- C5: C6 elevado con resto contenido; AHP medio.
- C6: Vulnerabilidad alta + C6/C7 altos; AHP alto.

Se confirma lo anticipado en las pruebas globales: los clústeres capturan tipologías con sentido (urbano denso, alta vulnerabilidad, financiamiento externo, entre otros), aunque varios pares se solapan en el eje AHP.

Figura 13. Identificación de tipologías de estratos con radar plot (7 estratos)

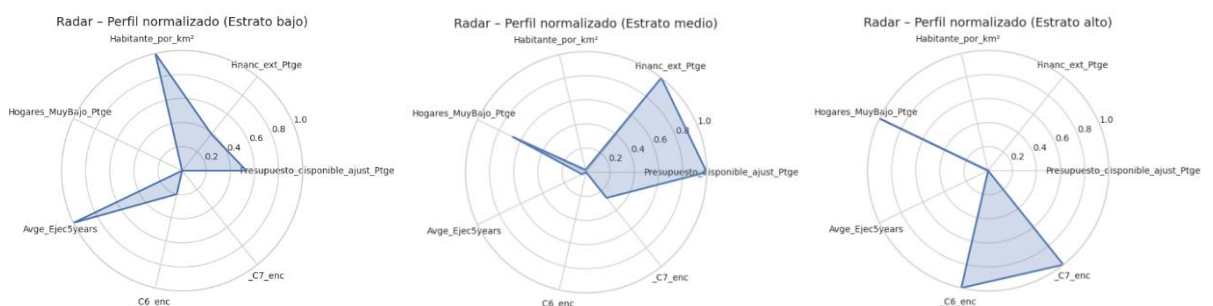


En cambio, la consolidación en tres estratos produce perfiles más sólidos (Figura 14):

- AHP_bajo: Alta densidad y baja vulnerabilidad; C6/C7 bajos.
- AHP_medio: Recursos internos altos y financiamiento externo; C6/C7 medios.
- AHP_alto: Vulnerabilidad alta con C6/C7 altos; baja densidad y financiamiento externo bajo.

La monotonidad respecto al score AHP y el equilibrio de tamaños por estrato mejoran la estabilidad inferencial y simplifican la comunicación con actores institucionales.

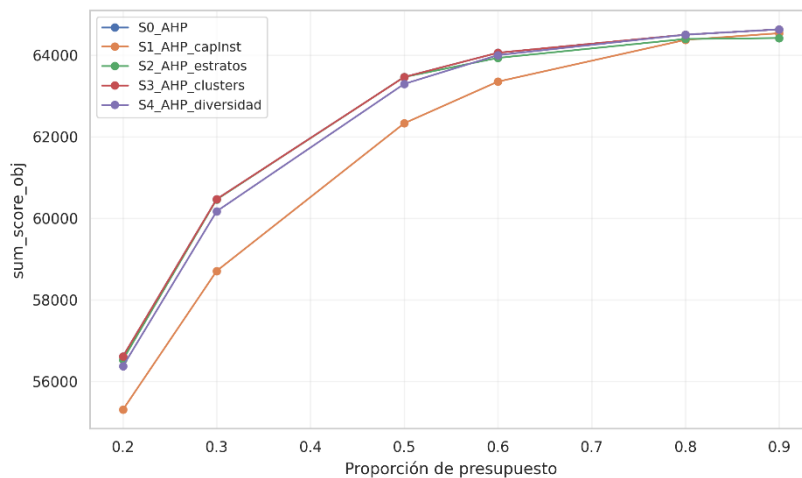
Figura 14. Identificación de tipologías de estratos con radar plot (3 estratos)



4.4. Etapa 5: Diseño experimental de optimización con restricción presupuestaria

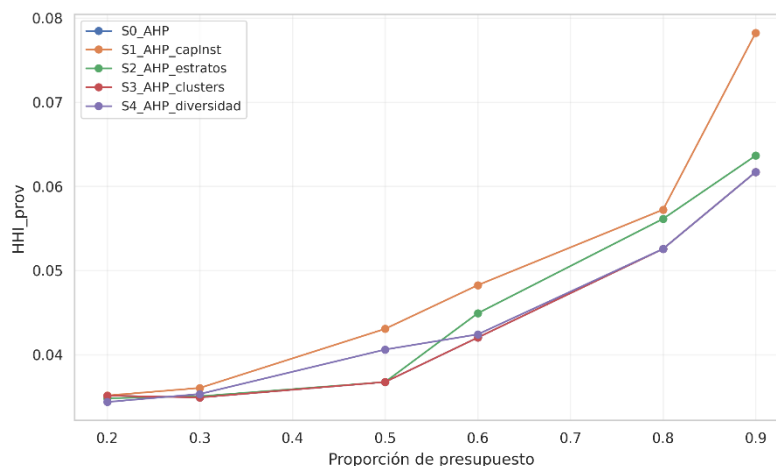
La Figura 15 compara, bajo “ajustes fuertes”, el valor agregado (\sum score AHP) alcanzado por S0–S4 a distintos techos. S0 marca la frontera por construcción. S2–S4 permanecen muy próximas a S0 para techos intermedios/altos, evidenciando que las reglas de cuotas por estrato/cluster y los premios de cobertura pueden activarse con pérdidas marginales de eficiencia. S1 (cap institucional) muestra menor desempeño a techos muy bajos (penaliza la selección “greedy” de proyectos con mayor score pertenecientes a pocas instituciones), pero converge a S0 cuando el techo se relaja.

Figura 15. Score AHP acumulado vs presupuesto (ajustes fuertes)



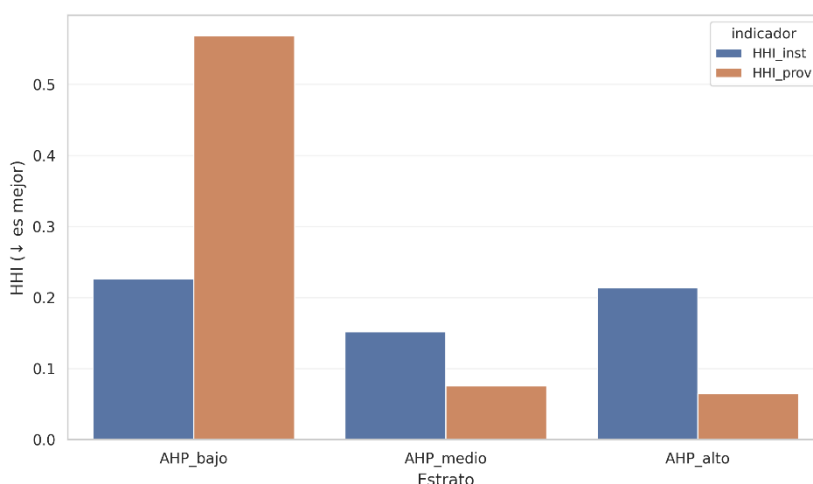
La Figura 16 muestra la concentración territorial (HHI de provincias) a lo largo de los techos. Todas las curvas crecen con el presupuesto (más gasto seleccionado en menos provincias relevantes), pero se observan diferencias por estrategia: S1 puede terminar con HHI_prov superior a S0 en techos altos, lo que sugiere un desplazamiento de la concentración desde la dimensión institucional (que S1 corrige) hacia la territorial (algunas instituciones tienden a concentrar en sus territorios “núcleo”). S2–S4 tienden a contener mejor el HHI provincial, coherente con su diseño: cuotas por estrato y premios de cobertura fuerzan mezcla territorial.

Figura 16. HHI provincial vs presupuesto (ajustes fuertes)



La Figura 17 reporta, por estrato, los HHI de instituciones y provincias dentro de los proyectos del grupo. AHP_bajo presenta la mayor concentración provincial, mientras AHP_alto exhibe la menor. La concentración institucional es moderada y más estable entre estratos, aunque puede aumentar en los extremos. Este patrón no contradice la narrativa de perfiles: el estrato bajo, vinculado a espacios metropolitanos densos, concentra inversión en pocas provincias; el alto, ligado a vulnerabilidad y C6/C7 altos, se distribuye mejor en el territorio.

Figura 17. HHI por estrato (Institución y provincia)



4.5. Discusión

La estrategia híbrida propuesta —AHP como señal normativa, clustering como capa empírica y MILP como articulador bajo restricción— permite observar la priorización desde dos planos complementarios: lo que debería priorizarse (alineamiento y brecha normativa) y lo que efectivamente se configura en la cartera (patrones y concentraciones). A partir de los resultados, pueden discutirse tres cuestiones.

¿Qué aporta el clustering respecto al AHP?

En primer lugar, el clustering revela estructura latente en la cartera que el AHP, por diseño, no busca modelar. Los siete grupos derivados (y su consolidación en tres estratos) organizan el portafolio en tipologías interpretables —urbano denso, alta vulnerabilidad con C6/C7 altos, financiamiento externo excepcional, etc.— que ayudan a explicar por qué determinados subconjuntos tienden a recibir puntuaciones normativas altas o bajas. En segundo lugar, la reducción de 7 a 3 aporta gobernabilidad: los estratos presentan escalonamiento monótono del score y tamaños más balanceados, lo que facilita reglas sencillas (cuotas por estrato, premios de cobertura) sin renunciar a los matices de los siete subtipos. En tercer lugar, al combinarse con MILP, los grupos permiten modular riesgos de concentración institucional y territorial, ofreciendo carteras más diversas con pérdidas de eficiencia acotadas respecto a la frontera AHP.

¿En qué casos divergen?

Las divergencias se concentran donde la realidad operativa empuja en sentido distinto a la señal normativa. Los clústeres con altísima densidad urbana y buena ejecución

histórica registran, en promedio, score AHP inferior por menor contribución en C6/C7 y menor vulnerabilidad; sin embargo, reúnen demanda y economías de escala. Algo análogo ocurre con el subtipo de financiamiento externo extremo: su existencia es informativamente relevante, pero no necesariamente mejora el alineamiento normativo. En las estrategias de optimización, el cap institucional (S1) mitiga acaparamientos por entidad, aunque puede desplazar la concentración hacia la dimensión territorial si no se equilibra con cuotas por estrato o incentivos de cobertura.

Estas discrepancias no desacreditan al AHP; muestran que, en la práctica, coexisten criterios de legitimidad normativa con criterios de gestión (densidad, logística, financiamiento) que la política pública debe transparentar y ponderar.

¿Cómo usar estos resultados para la toma de decisiones en la DGIP?

Se propone un protocolo de decisión en cuatro pasos:

- i. Referencia normativa: publicar siempre la cartera S0 (AHP puro) como frontera superior para cada techo presupuestario.
- ii. Diagnóstico empírico: describir la composición por estratos y monitorear concentración (HHI) y cobertura (institucional y territorial). Mantener los siete clústeres como subtipos para comprensión fina, no como reglas operativas.
- iii. Modulación con reglas simples: activar S2 (cuotas por estrato) y S4 (premios de cobertura) cuando la concentración supere umbrales o la cobertura caiga por debajo de mínimos definidos; usar S1 como correctivo puntual frente a acaparamientos iniciales.
- iv. Negociación transparente: escoger parámetros que limiten la pérdida de Σ score frente a S0 dentro de un umbral explícito (2–3%) y documenten el “precio de la diversidad” en términos de reducción de HHI y aumento de cobertura.

Capítulo 5

5. Conclusiones y líneas futuras de investigación

Este capítulo sintetiza lo aprendido a lo largo del estudio y lo traduce en implicaciones técnicas e institucionales. El hilo conductor ha sido claro: partir de un criterio normativo (AHP) que fija la referencia de priorización; incorporar evidencia empírica mediante clustering para comprender cómo se estructura de facto la cartera; y, finalmente, articular ambos planos en un esquema de optimización lineal entera mixta (MILP) que hace explícitos los compromisos entre eficiencia normativa y diversidad institucional/territorial bajo restricciones de presupuesto. Sobre esta base se presentan las conclusiones.

5.1. Conclusiones técnicas

En términos metodológicos, el trabajo demuestra que es posible construir un modelo híbrido robusto, trazable y replicable para la priorización de proyectos del PNPIP. El AHP

aporta una señal estable y reproducible que ordena los proyectos de acuerdo con los criterios del marco SNIP/PNPIP; el modelado no supervisado revela tipologías que se repiten en la cartera (urbano denso, alta vulnerabilidad con C6/C7 elevados, financiamiento externo excepcional, entre otras) y que explican parte de la variabilidad del score. La consolidación de siete clústeres en tres estratos (AHP_bajo, AHP_medio, AHP_alto) resulta especialmente útil: reduce el solapamiento observado, mantiene la monotonía del score y equilibra tamaños, con lo que habilita reglas simples de decisión.

La integración de estas capas en un MILP permite cuantificar de manera transparente el coste de oportunidad de la diversidad. Las comparaciones S0–S4 muestran que, en techos intermedios, estrategias con cuotas por estrato y premios de cobertura (S2 y S4) desconcentran y amplían la cobertura a un precio pequeño en términos de pérdida de Σ score frente a S0. El cap institucional (S1) es eficaz para frenar acaparamientos por entidad cuando el techo es reducido, pero puede trasladar concentración hacia lo territorial si se aplica aisladamente. En conjunto, los resultados confirman que la sinergia AHP–ML–MILP no sustituye la norma; la hace gobernable: cada cartera reportada es “óptima dado un conjunto público de reglas”, y su comparación contra S0 explicita qué se gana y qué se sacrifica.

Finalmente, la arquitectura propuesta es extensible. El mismo armazón admite variantes multiobjetivo, robustas o plurianuales sin reformular el sistema desde cero, lo cual la convierte en una base útil para la mejora incremental del proceso de priorización.

5.2. Conclusiones institucionales

Desde la perspectiva de política pública, el principal aporte es un protocolo de decisión claro y comunicable. Primero, publicar S0 como frontera normativa para cada techo presupuestario; segundo, diagnosticar concentración (HHI) y cobertura (instituciones y provincias) usando los tres estratos como unidad de lectura; tercero, modular la cartera con reglas simples: cuotas por estrato que favorezcan AHP_alto y premios de cobertura para asegurar representación institucional y territorial, dejando el cap como corrección puntual frente a acaparamientos iniciales; cuarto, negociar parámetros con un umbral explícito de pérdida de Σ score frente a S0, de modo que el “precio de la diversidad” quede justificado y documentado.

Este enfoque eleva la legitimidad del proceso: hace visibles las razones técnicas de cada decisión, distribuye mejor la atención pública y favorece la rendición de cuentas. A la vez, preserva el papel del AHP como eje rector, lo que facilita continuidad institucional.

5.3. Limitaciones del estudio

El ejercicio se realiza con un horizonte anual y un costo simple (2025), sin modelar compromisos plurianuales ni flujos intertemporales de ejecución. El contenido normativo del AHP —pesos y criterios— se asume como dado; cambios institucionales o mejoras de calidad de datos pueden alterar el ranking. En la capa empírica, los clústeres presentan solapes y existe al menos un subtipo pequeño (financiamiento externo extremo) cuya estabilidad es menor, lo que aconseja interpretarlos como tipologías y no

como categorías rígidas. La función objetivo se centra en el score AHP y coberturas lineales; no incorpora aún complementariedades entre proyectos, riesgo de ejecución ni beneficios económicos. Finalmente, ciertas restricciones (por ejemplo, el cap) pueden mejorar una dimensión de concentración y empeorar otra si no se combinan con reglas complementarias.

5.4. Recomendaciones de política pública y de investigación futura

Para la DGIP (operación y gobernanza):

Se recomienda institucionalizar un flujo en cuatro pasos:

- i. Difundir S0 como referencia.
- ii. Vigilar HHI y coberturas por estratos, instituciones y provincias.
- iii. Activar S2/S4 cuando la concentración supere umbrales ex ante o la cobertura caiga por debajo de mínimos.
- iv. Limitar la pérdida de Σ score frente a S0 a un máximo acordado y documentar la decisión. Conviene además disponer de un tablero de control con series históricas de concentración/cobertura y dejar registro público de los parámetros empleados en cada ciclo. Este esquema promueve transparencia, previsibilidad y aprendizaje institucional.

Para la agenda técnica (investigación y desarrollo):

Son prioritarias cinco extensiones:

- i. Una optimización plurianual con continuidad de proyectos y manejo de arrastres.
- ii. Un enfoque multiobjetivo (fronteras Pareto) y robusto ante incertidumbres de costos y scores.
- iii. La incorporación de riesgo y valor económico (probabilidad de ejecución, sobrecostos, NPV/beneficio esperado) en la función objetivo.
- iv. Datos enriquecidos —texto de expedientes (NLP), información geoespacial y variables de conectividad— para mejorar la capa empírica.
- v. Mecanismos human-in-the-loop que permitan simular techos y parámetros, visualizar el coste de oportunidad antes del cierre de cartera y capturar feedback experto. Finalmente, conviene emprender evaluaciones causales ex post que comparen las carteras elegidas frente a contrafactuales S0–S4 y midan impacto en concentración, cobertura y resultados de desarrollo.

Bibliografía

- Baffo, I., Leonardi, M., D'Alberti, V., & Petrillo, A. (2024). Optimizing public investments: A Sustainable Economic, Environmental, and Social Investment Multi-criteria Decision Model (SEESIM). *Regional Science Policy & Practice*, 16(11), 100140. <https://doi.org/10.1016/j.rspp.2024.100140>
- Bertolini, M., Braglia, M., & Carmignani, G. (2006). Application of the AHP methodology in making a proposal for a public work contract. *International Journal of Project Management*, 24(5), 422-430. <https://doi.org/10.1016/j.ijproman.2006.01.005>
- Biernacki, C., Celeux, G., & Govaert, G. (2000). Assessing a mixture model for clustering with the integrated completed likelihood. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(7), 719-725. <https://doi.org/10.1109/34.865189>
- Bock, H.-H. (2008). Origins and extensions of the k-means algorithm in cluster analysis. *Journal Électronique d'Histoire des Probabilités et de la Statistique [electronic only]*, 4.
- Boix Cots, D. (2023). *Sistema de evaluación y priorización de inversiones públicas* [Universitat Politècnica de Catalunya]. <https://doi.org/10.5821/dissertation-2117-405577>
- Caliński, T., & JA, H. (1974). A Dendrite Method for Cluster Analysis. *Communications in Statistics - Theory and Methods*, 3, 1-27. <https://doi.org/10.1080/03610927408827101>
- Congreso Nacional de la República Dominicana. (2011). *Ley 1-12 Estrategia Nacional de Desarrollo* (Ley nacional Nos. 1-12; p. 94). Congreso Nacional de la República Dominicana.

- Corrales, D. G. N., Achurra, V. E. P., Blanco, L., & A, H. R. Á. (2024). Aplicación del método AHP para la priorización de proyectos de inversión social en la Comarca NGÄBE-BUGLÉ. *I+D Tecnológico*, 20(1), 73-81. <https://doi.org/10.33412/idt.v20.1.3976>
- Darwin Marcelo, Cledan Mandri-Perrott, Schuyler House, & Jordan Z. Schwartz. (2016, abril 14). *An Alternative Approach to Project Selection: The Infrastructure Prioritization Framework*. <https://www.github.org/resources/publications/an-alternative-approach-to-project-selection-the-infrastructure-prioritisation-framework/>
- Davies, D. L., & Bouldin, D. W. (1979). A Cluster Separation Measure. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-1(2), 224-227. <https://doi.org/10.1109/TPAMI.1979.4766909>
- De Smet, Y., & Montano Guzmán, L. (2004). Towards multicriteria clustering: An extension of the *k*-means algorithm. *European Journal of Operational Research*, 158(2), 390-398. <https://doi.org/10.1016/j.ejor.2003.06.012>
- Decreto 493-07. (2007). *Reglamento de Aplicación No. 1 para la Ley No. 498-06* (Reglamento ejecutivo Nos. 493-07). Presidencia de la República Dominicana. <https://mepyd.gob.do/estrategia-nacional-de-desarrollo-2030/marco-legal/>
- Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum Likelihood from Incomplete Data via the EM Algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*, 39(1), 1-38.
- Goldberger, J., & Roweis, S. (2004). Hierarchical Clustering of a Mixture Model. En *Advances in Neural Information Processing Systems* (Vol. 17).

- Ley 498-06. (2006). *Ley No. 498-06 sobre Planificación e Inversión Pública* (Ley nacional Nos. 498-06). Congreso Nacional de la República Dominicana.
<https://mepyd.gob.do/estrategia-nacional-de-desarrollo-2030/marco-legal/>
- Lobo Flores, W. (2024). Evaluación multicriterio para la priorización de asignación de recursos económicos en proyectos públicos. Caso «Municipio de Achocalla», gestión 2023. *Fides et Ratio - Revista de Difusión cultural y científica de la Universidad La Salle en Bolivia*, 27(27), 227-248.
- Lücke, J., & Forster, D. (2019). K-means as a variational EM approximation of Gaussian mixture models. *Pattern Recognition Letters*, 125, 349-356.
<https://doi.org/10.1016/j.patrec.2019.04.001>
- Mariani, C., Navrotska, Y., & Mancini, M. (2023). Unsupervised machine learning for project stakeholder classification: Benefits and limitations. *Project Leadership and Society*, 4, 100093. <https://doi.org/10.1016/j.plas.2023.100093>
- McInnes, L., & Healy, J. (2017). Accelerated Hierarchical Density Clustering. *2017 IEEE International Conference on Data Mining Workshops (ICDMW)*, 33-42.
<https://doi.org/10.1109/ICDMW.2017.12>
- Moulavi, D., Andretta Jaskowiak, P., Campello, R., Zimek, A., & Sander, J. (2014). *Density-Based Clustering Validation*. <https://doi.org/10.1137/1.9781611973440.96>
- Posse, C. (2001). Hierarchical Model-Based Clustering for Large Datasets. *Journal of Computational and Graphical Statistics*, 10(3), 464-486.
- Ricardo Viana Vargas. (2010, octubre 12). *Using the analytic hierarchy process (ahp) to select and prioritize projects in a portfolio*. PMI® Global Congress 2010, North America, Washington, DC. Newtown Square.

https://www.pmi.org/learning/library/analytic-hierarchy-process-prioritize-projects-6608?utm_source=chatgpt.com

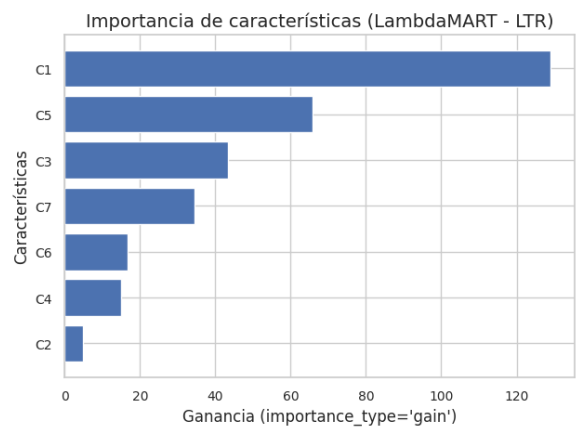
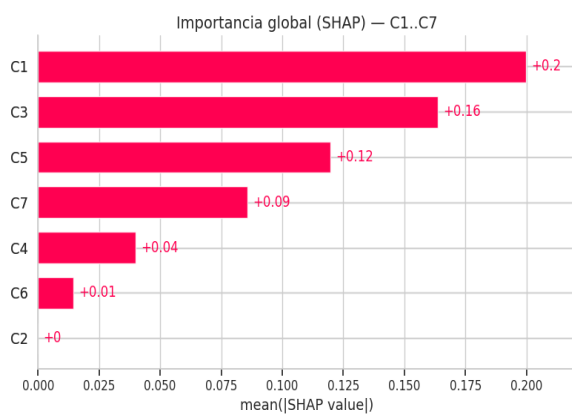
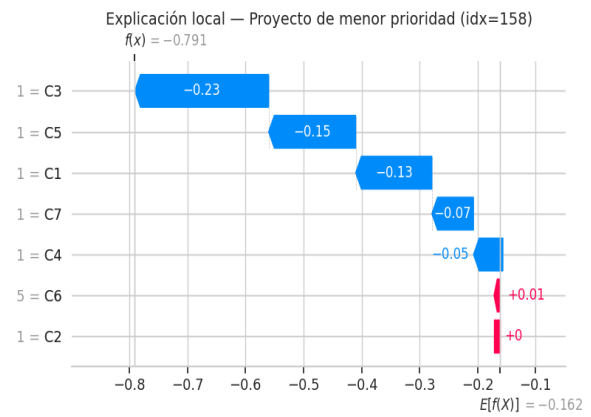
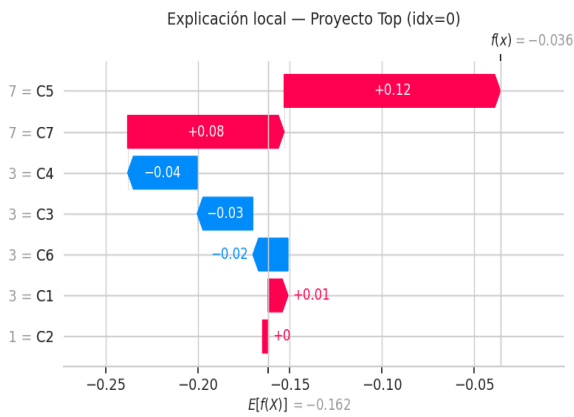
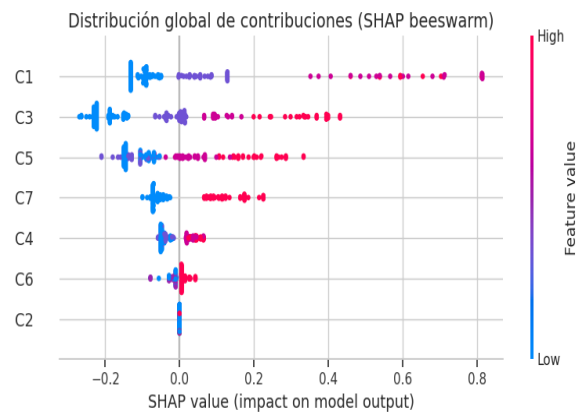
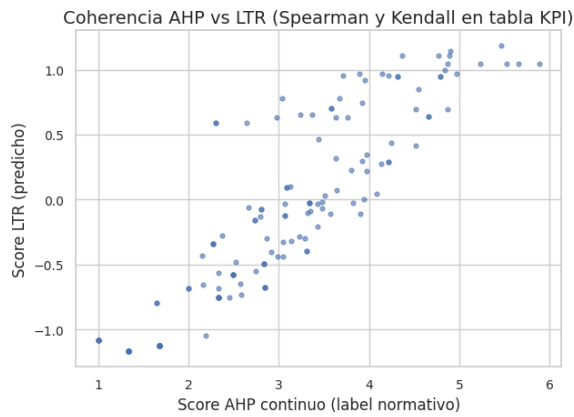
Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, 20, 53-65. [https://doi.org/10.1016/0377-0427\(87\)90125-7](https://doi.org/10.1016/0377-0427(87)90125-7)

Saaty, R. W. (1987). The analytic hierarchy process—What it is and how it is used. *Mathematical Modelling*, 9(3), 161-176. [https://doi.org/10.1016/0270-0255\(87\)90473-8](https://doi.org/10.1016/0270-0255(87)90473-8)

Yu, H., Chapman, B., Di Florio, A., Eischen, E., Gotz, D., Jacob, M., & Blair, R. H. (2019). Bootstrapping estimates of stability for clusters, observations and model selection. *Computational Statistics*, 34(1), 349-372. <https://doi.org/10.1007/s00180-018-0830-y>

Anexos

Anexo 1. Diseño experimental con aprendizaje supervisado (LTR)



Anexo 2. Métricas relativas vs S0

budget_pct	strategy	sum_score_obj	loss_obj_pct	rel_HHI_pct	rel_ENT_pct	rel_cov_inst
0.2	S0_AHP	56.614.274	0	0	0	0
0.2	S1_AHP_capInst	55.309.128	-2.305	-45.664	25.311	4
0.2	S2_AHP_estratos	56.532.781	-144	-5.431	2.945	0
0.2	S3_AHP_clusters	56.614.274	0	0	0	0
0.2	S4_AHP_diversidad	56.372.981	-426	-8.009	7.656	6
0.3	S0_AHP	60.471.344	0	0	0	0
0.3	S1_AHP_capInst	58.710.513	-2.912	-50.314	21.955	0
0.3	S2_AHP_estratos	60.462.862	-14	-545	-65	0
0.3	S3_AHP_clusters	60.471.344	0	0	0	0
0.3	S4_AHP_diversidad	60.173.054	-493	-7.277	7.315	5
0.5	S0_AHP	63.464.021	0	0	0	0
0.5	S1_AHP_capInst	62.329.848	-1.787	-42.074	16.866	3
0.5	S2_AHP_estratos	63.464.021	0	0	0	0
0.5	S3_AHP_clusters	63.464.021	0	0	0	0
0.5	S4_AHP_diversidad	63.295.279	-266	-11.321	8.548	5
0.6	S0_AHP	64.059.725	0	0	0	0
0.6	S1_AHP_capInst	63.352.579	-1.104	-33.849	11.664	1
0.6	S2_AHP_estratos	63.933.626	-197	376	-2.368	-1
0.6	S3_AHP_clusters	64.059.725	0	0	0	0
0.6	S4_AHP_diversidad	64.002.454	-89	-3.429	2.678	2

Anexo 3. Distribución por estrato y por cluster

Estrato	min_share_estrato
AHP_bajo	87.500
AHP_medio	195.821
AHP_alto	416.679
Cluster	
0	21.357
1	50.240
2	16.775
3	28.259
4	16.775
5	23.382
6	43.211

Anexo 3. Bloques de códigos que sustentan este trabajo

https://colab.research.google.com/github/diegolassis/TFM---Diego-Lassis/blob/main/TFM_DiegoLassis.ipynb

Anexo 4. Tabla maestra de resultados del diseño experimental de optimización

budget_pct	strategy	sum_score_obj	HHI_inst	HHI_prov	coverage_inst	coverage_prov	coverage_strat	pct_strat_AHP_alto	loss_obj_pct	rel_HHI_inst_pct	rel_HHI_prov_pct	d_cov_inst	d_cov_prov
0.2	S0_AHP	56614.274	0.261	0.035	47	35	3	0.567	0.000	0.000	0.000	0	0
0.2	S1_AHP_caplnst	55309.128	0.142	0.035	51	35	3	0.501	-2.305	-45.664	0.052	4	0
0.2	S2_AHP_estratos	56614.274	0.261	0.035	47	35	3	0.567	0.000	0.000	0.000	0	0
0.2	S3_AHP_clusters	56614.274	0.261	0.035	47	35	3	0.567	0.000	0.000	0.000	0	0
0.2	S4_AHP_diversidad	56614.274	0.261	0.035	47	35	3	0.567	0.000	0.000	0.000	0	0
0.3	S0_AHP	60471.344	0.272	0.035	51	35	3	0.516	0.000	0.000	0.000	0	0
0.3	S1_AHP_caplnst	58710.513	0.135	0.036	51	35	3	0.487	-2.912	-50.314	3.240	0	0
0.3	S2_AHP_estratos	60471.344	0.272	0.035	51	35	3	0.516	0.000	0.000	0.000	0	0
0.3	S3_AHP_clusters	60471.344	0.272	0.035	51	35	3	0.516	0.000	0.000	0.000	0	0
0.3	S4_AHP_diversidad	60471.344	0.272	0.035	51	35	3	0.516	0.000	0.000	0.000	0	0
0.5	S0_AHP	63464.021	0.203	0.037	55	35	3	0.418	0.000	0.000	0.000	0	0
0.5	S1_AHP_caplnst	62329.848	0.118	0.043	58	35	3	0.442	-1.787	-42.074	17.175	3	0

budget_pct	strategy	sum_score_obj	HHI_inst	HHI_prov	coverage_inst	coverage_prov	coverage_strat	pct_strat_AHP_alto	loss_obj_pct	rel_HHI_inst_pct	rel_HHI_prov_pct	d_cov_inst	d_cov_prov
0.5	S2_AHP_estratos	63464.021	0.203	0.037	55	35	3	0.418	0.000	0.000	0.000	0	0
0.5	S3_AHP_clusters	63464.021	0.203	0.037	55	35	3	0.418	0.000	0.000	0.000	0	0
0.5	S4_AHP_diversidad	63464.021	0.202	0.037	56	35	3	0.418	0.000	-0.564	0.242	1	0
0.6	S0_AHP	64059.725	0.171	0.042	58	35	3	0.411	0.000	0.000	0.000	0	0
0.6	S1_AHP_caplnst	63352.579	0.113	0.048	59	35	3	0.390	-1.104	-33.849	14.804	1	0
0.6	S2_AHP_estratos	64059.725	0.171	0.042	58	35	3	0.411	0.000	0.000	0.000	0	0
0.6	S3_AHP_clusters	64059.725	0.171	0.042	58	35	3	0.411	0.000	0.000	0.000	0	0
0.6	S4_AHP_diversidad	64059.725	0.171	0.042	58	35	3	0.411	0.000	-0.001	0.003	0	0
0.7	S0_AHP	64335.915	0.146	0.051	59	35	3	0.353	0.000	0.000	0.000	0	0
0.7	S1_AHP_caplnst	63984.813	0.108	0.051	60	35	3	0.343	-0.546	-26.303	-0.822	1	0
0.7	S2_AHP_estratos	64335.915	0.146	0.051	59	35	3	0.353	0.000	0.000	0.000	0	0
0.7	S3_AHP_clusters	64335.915	0.146	0.051	59	35	3	0.353	0.000	0.000	0.000	0	0
0.7	S4_AHP_diversidad	64335.915	0.146	0.051	59	35	3	0.353	0.000	0.000	0.000	0	0
0.8	S0_AHP	64502.672	0.147	0.053	60	35	3	0.309	0.000	0.000	0.000	0	0

budget_pct	strategy	sum_score_obj	HHI_inst	HHI_prov	coverage_inst	coverage_prov	coverage_strat	pct_strat_AHP_alto	loss_obj_pct	rel_HHI_inst_pct	rel_HHI_prov_pct	d_cov_inst	d_cov_prov
0.8	S1_AHP_caplnst	64375.661	0.113	0.057	60	35	3	0.341	-0.197	-23.122	8.860	0	0
0.8	S2_AHP_estratos	64485.527	0.134	0.055	60	35	3	0.349	-0.027	-8.714	5.127	0	0
0.8	S3_AHP_clusters	64502.672	0.147	0.053	60	35	3	0.309	0.000	0.000	0.000	0	0
0.8	S4_AHP_diversidad	64502.672	0.147	0.053	60	35	3	0.309	0.000	0.000	0.000	0	0
0.9	S0_AHP	64633.653	0.143	0.062	60	35	3	0.311	0.000	0.000	0.000	0	0
0.9	S1_AHP_caplnst	64538.769	0.112	0.078	60	35	3	0.277	-0.147	-21.980	26.802	0	0
0.9	S2_AHP_estratos	64633.653	0.143	0.062	60	35	3	0.311	0.000	0.000	0.000	0	0
0.9	S3_AHP_clusters	64633.653	0.143	0.062	60	35	3	0.311	0.000	0.000	0.000	0	0
0.9	S4_AHP_diversidad	64633.653	0.143	0.062	60	35	3	0.311	0.000	0.000	0.000	0	0

Anexo 5. Glosario de términos de métricas implementadas en el análisis.

AHP (score y consistencia)

Métrica / Columna	Descripción	Cómo se debe interpretar
<i>score_AHP_Ptge</i>	Puntuación AHP del proyecto en escala 0–100 (suma ponderada de criterios C1–C7).	Mayor es mejor (mayor alineación normativa). Se usa para ordenar y como valor a maximizar en S0–S4.
<i>score_AHP</i>	Puntuación AHP en escala original (antes de llevar a 0–100).	Solo referencia; no cambia el orden respecto a <i>score_AHP_Ptge</i> .
<i>lambda_max</i>	Autovalor máximo de la matriz de comparación por pares.	En AHP, sirve para calcular CI. Valores cercanos a <i>n</i> (número de criterios) indican coherencia.
<i>CI (Índice de Consistencia)</i>	$CI = (\lambda_{max} - n) / (n - 1) CI$.	Se combina con RI para el CR. Menor es mejor.
<i>RI (Índice Aleatorio)</i>	Valor esperado de CI para matrices aleatorias (depende de <i>n</i>).	Dado por tablas estándar. No es “calidad”.
<i>CR (Razón de Consistencia)</i>	$CR = CI / RI$.	CR ≤ 0.10 (o ≤0.08, según escuela) indica juicios coherentes .

Métricas de evaluación de clustering (selección de k/modelo)

Métrica / Columna	Descripción	Cómo se debe interpretar
<i>silhouette</i>	Cohesión/separación media (–1 a 1).	Más alto es mejor. ≥0.5 muy bueno; 0.3–0.5 razonable.
<i>davies_bouldin (db)</i>	Media de ratios intra/inter-cluster (≥0).	Más bajo es mejor. 0 indica separación perfecta.
<i>calinski_harabasz (ch)</i>	Varianza inter/intra ponderada.	Más alto es mejor.
<i>stability_jaccard_mean / stab_mean</i>	Estabilidad media (Jaccard) de particiones re-muestreadas (0–1).	Más alto es mejor. >0.7 indica muy estable.
<i>stability_jaccard_std</i>	Desviación de la estabilidad.	Más bajo es mejor (menos variabilidad).
<i>dbcv (HDBSCAN)</i>	Densidad-between/within (–1 a 1).	Más alto es mejor.
<i>noise_ratio (HDBSCAN)</i>	Proporción de puntos etiquetados como ruido (–1).	Más bajo es mejor.
<i>k_clusters</i>	Nº de grupos obtenidos.	Informativo; no “mejor/peor” per se.
<i>gini / gini_sizes</i>	Desigualdad de tamaños de clusters (0–1).	Más bajo es mejor (tamaños más balanceados).
<i>min_pct</i>	Tamaño relativo del cluster más pequeño.	Más alto es mejor ; umbrales <1–2% suelen penalizarse.
<i>n_tiny / a_tiny_clusters</i>	Número de clusters diminutos (por debajo de un umbral).	Más bajo es mejor.
<i>cov (GMM)</i>	Estructura de covarianza (full, diag, spherical).	full es más flexible (captura anisotropías); puede ser inestable si condición es alta.
<i>icl (GMM)</i>	<i>Integrated Completed Likelihood</i> (likelihood penalizada por ambigüedad).	Más bajo es mejor. Favorece particiones “nítidas”.
<i>bic / aic (GMM)</i>	Criterios de información (penalizan complejidad).	Más bajo es mejor. Suelen decrecer con <i>k</i> ; no siempre dan óptimo interno.
<i>certainty_combo (GMM)</i>	Proporción de asignaciones con probabilidad alta.	Más alto es mejor (poca ambigüedad).

Métrica / Columna	Descripción	Cómo se debe interpretar
<i>n_weights_below_2pct (GMM)</i>	Nº de componentes con peso <2%.	Más bajo es mejor (evita componentes espurias).
<i>cov_cond_mean (GMM)</i>	Condicionamiento medio de matrices de covarianza.	Más bajo es mejor ; valores altos indican componentes muy alargadas.
<i>rank_* / rank_avg</i>	Ranking por métrica y promedio de rangos.	Más bajo es mejor (1 = mejor).
<i>rank_policy_weighted</i>	Ranking ponderado incorporando filtros “policy-aware” (estabilidad mínima, sin ruido/tiny, cobertura).	Más bajo es mejor . Determina el modelo elegido .
<i>fails_filters / penalty_flag</i>	Indicadores de fallas a umbrales (estabilidad, tiny, ruido).	TRUE implica penalización o descarte.

Optimización/cartera (S0-S4)

Métrica / Columna	Descripción	Cómo se debe interpretar
<i>budget_pct (β)</i>	Techo presupuestario como fracción del costo total 2025.	Escenarios típicos: 0.2, 0.3, ..., 0.9.
<i>strategy</i>	Estrategia aplicada: S0_AHP, S1_AHP_capIns, S2_AHP_estratos, S3_AHP_clusters, S4_AHP_diversidad.	Identifica el tipo de reglas/recompensas usadas.
<i>status</i>	Estado del solver (Optimal, Infeasible, etc.).	“Optimal” garantiza solución óptima bajo reglas.
<i>sum_score_obj</i>	Función objetivo alcanzada . En S0–S3 coincide con $\sum \text{score_AHP_Ptge}$; en S4 incluye el premio de coberturas .	Más alto es mejor . No es comparable entre S0–S3 y S4 si S4 incluye términos adicionales (se compara con su propia referencia).
<i>sum_score_AHP_Ptge</i>	Suma de score_AHP_Ptge de los proyectos seleccionados.	Mayor es mejor (eficiencia normativa agregada).
<i>sum_score_AHP</i>	Suma del score AHP en escala original.	Misma interpretación; útil si se requiere la escala “cruda”.
<i>budget_used</i>	Costo total seleccionado (moneda 2025).	Debe ser $\leq \text{budget_pct} \cdot \text{costo total}$; cercano al techo indica buen “relleno”.
<i>n_projects</i>	Nº de proyectos seleccionados.	Mayor implica cartera más granular; no garantiza diversidad institucional/territorial.
<i>HHI_inst</i>	Índice Herfindahl–Hirschman del gasto por institución . $\sum s_g^2 \sum s_g^2$.	Más bajo es mejor (menos concentración).
<i>HHI_prov</i>	HHI del gasto por provincia .	Más bajo es mejor .
<i>ENT_inst</i>	Entropía de Shannon del gasto por institución.	Más alto es mejor (más dispersión).
<i>GINI_inst</i>	Gini del gasto por institución (0–1).	Más bajo es mejor . Complementa HHI.
<i>coverage_inst</i>	Nº de instituciones con ≥ 1 proyecto seleccionado.	Más alto = mayor cobertura institucional.
<i>coverage_prov</i>	Nº de provincias con ≥ 1 proyecto seleccionado.	Más alto = mayor cobertura territorial.
<i>coverage_strat</i>	Nº de estratos (0–3) presentes en la cartera.	3 indica presencia de AHP_bajo/medio/alto.
<i>pct_strat_AHP_alto</i>	Proporción del gasto seleccionado asignado al estrato AHP_alto.	Más alto = mayor énfasis en prioridad normativa.

Métricas relativas vs S0 (coste de oportunidad/diversidad)

Métrica / Columna	Descripción	Cómo se debe interpretar
<i>loss_obj_pct</i>	Pérdida porcentual de <i>sum_score_obj</i> respecto a S0 en el mismo β : $100 \cdot (S - S0) / S0$	Valores entre 0 y -3% indican pérdida marginal aceptable.
<i>rel_HHI_inst_pct</i>	Cambio % de <i>HHI_inst</i> vs. S0: $100 \cdot (HHI / HHIS0 - 1)$.	Negativo = menos concentración institucional que S0.
<i>rel_HHI_prov_pct</i>	Cambio % de <i>HHI_prov</i> vs. S0.	Negativo = menos concentración territorial que S0.
<i>d_cov_inst</i>	Diferencia absoluta en <i>coverage_inst</i> vs. S0.	Positivo = más instituciones cubiertas.
<i>d_cov_prov</i>	Diferencia absoluta en <i>coverage_prov</i> vs. S0.	Positivo = más provincias cubiertas.
<i>dOBJ_vs_S0</i>	Diferencia absoluta de <i>sum_score_obj</i> vs. S0.	Negativa = pérdida; positiva = ganancia (poco habitual con restricciones).
<i>dHHI_inst_vs_S0 / dHHI_prov_vs_S0</i>	Diferencia absoluta de HHI vs. S0 (no en %).	Negativa = desconcentración .
<i>dENT_inst_vs_S0 / dGINI_inst_vs_S0</i>	Diferencia absoluta de entropía/Gini vs. S0.	Entropía positiva y Gini negativa = más diversidad.
<i>dCoverageInst_vs_S0</i>	Igual a <i>d_cov_inst</i> (nomenclatura alternativa).	Ver <i>d_cov_inst</i> .
<i>dPctStratAlto_vs_S0</i>	Diferencia en <i>pct_strat_AHP_alto</i> vs. S0 (en puntos porcentuales).	Positivo = mayor énfasis en <i>AHP_alto</i> que S0.

Parámetros de escenario (para lectura de tablas y anexos)

Parámetro	Descripción	Cómo se debe interpretar
<i>CAP_INST_FRAC (S1)</i>	Tope relativo por institución (fracción del techo asignable a una única entidad).	Menor = cap más estricto (más desconcentración, posible pérdida de Σ score).
<i>min_share_estrato / min_share_cluster (S2/S3)</i>	Cuota mínima por estrato/clúster (fracción del gasto seleccionado).	Mayor fuerza presencia mínima; mejora diversidad con posible coste en Σ score si se aleja del ranking AHP.
<i>LAMBDA_INST, LAMBDA_PROV, LAMBDA_STRAT (S4)</i>	Pesos de cobertura en la función objetivo.	Más alto premia más la diversidad; puede reducir Σ score si es excesivo.