

Evaluación de desempeño logístico mediante el agrupamiento de K-means y AHP difuso. Una perspectiva internacional.

Realizado por:

Santiago Pérez Santiago

Tutelado por:

Ana Gessa Perera

A thesis submitted in conformity with the requirements for the MSc in
Economics, Finance and Computer Science

University of Huelva & International University of Andalusia

uhu.es

un
i Universidad
Internacional
de Andalucía
A

Resumen

Partiendo del indicador de referencia del desempeño logístico a nivel mundial (Índice de Desarrollo Logístico) publicado periódicamente por el Banco Mundial, en este trabajo se propone una alternativa para evaluar el rendimiento de los sistemas logísticos de los países, aprovechando el potencial de la combinación de técnicas de aprendizaje automático y los métodos de decisión multicriterio. Aplicamos el método del proceso de jerarquía analítica difuso para construir un modelo de evaluación del rendimiento logístico de los países, con la ayuda del método de análisis de clustering (k-medias) y así determinar la ponderación de los diferentes indicadores que componen el Índice de Desarrollo Logístico. Para ello se ha empleado datos representativos de los flujos de mercancías y servicios entre países. De esta manera, el trabajo desarrollado propone una alternativa que permita un análisis comparativo, además de actualizado, más objetivo e insesgado de los sistemas logísticos de los países, convirtiéndose en un soporte eficaz para la toma de decisiones.

Palabra clave: Desempeño logístico, Índice de Rendimiento Logístico, Métodos de decisión multicriterio, AHP difuso, K-means, clúster.

Abstract

Based on the World Bank's regularly published global logistics performance benchmark (Logistics Development Index), this paper proposes an alternative way of assessing the performance of countries' logistics systems by exploiting the potential of combining machine learning techniques and multi-criteria decision methods. We apply the fuzzy analytic hierarchy process method to build a model for evaluating the logistics performance of countries, with the help of the clustering analysis method (k-means) and thus determine the weighting of the different indicators that make up the Logistics Development Index. For this purpose, representative data on the flow of goods and services between countries has been used. In this way, the work developed proposes an alternative to that proposed by the World Bank that allows for a comparative analysis, as well as being up-to-date, more objective and unbiased, of the countries' logistics systems, thus becoming an effective support for decision-making.

Key words: Logistic performance, Logistic Performance Index, Multi-Criteria Decision Methods, Fuzzy AHP, K-means, clustering

Índice

1. Introducción.....	5
2. Revisión de la literatura	6
2.1. Alternativas para determinar el desempeño logístico.	6
2.2. Aplicación de métodos multicriterio de decisión en el ámbito del desempeño logístico.....	8
2.3. Aplicación del clustering en el ámbito del desempeño logístico.....	10
3. Metodología.....	11
3.1. Datos	11
3.2. K-means (Clustering).....	13
3.3.AHP difuso.....	16
4. Resultados.....	19
5. Conclusiones.....	30
Referencias	32

Índice de tablas

Tabla 1. Componentes del Índice de Desempeño Logístico (IDL).....	12
Tabla 2. Variables de las transacciones comerciales	13
Tabla 3. Ponderaciones de los criterios de las variables lingüísticas	17
Tabla 4. Países agrupados en tres clústeres	20
Tabla 5. Distribución de los clústeres de países por variables y clúster.....	23
Tabla 6. Puntuación resultante de aplicar el algoritmo Relief.	24
Tabla 7. Matriz difusa triangular de comparaciones pareadas	25
Tabla 8. Valor de la extensión sintética.	26
Tabla 9. Pesos difusos triangulares.	26
Tabla 10. Criterio de peso promedio (M_i) y criterio de peso normalizado (N_i). 27	
Tabla 11. Resultados IDL-AHP	28
Tabla 12. Clasificación de los 10 primeros países del IDL-AHP.....	29
Tabla 13. Clasificación de los 10 últimos países del IDL-AHP	29

Índice de gráficos

Gráfico 1. Número óptimo de clústeres usando ancho de Silhouette y diagrama de clúster.....	15
Gráfico 2. Distribución Normal del IDL	21
Gráfico 3. Diagrama de caja del IDL.....	21
Gráfico 4. Diagrama de caja para cada componente del IDL	22

Índice de figuras

Figura 1. Estructura del modelo jerárquico.	16
Figura 2. Variables lingüísticas para el peso de importancia de cada criterio ...	19

1. Introducción

Debido al impacto de la globalización y a un panorama cada vez más competitivo, la logística se ha convertido en un elemento fundamental del comercio internacional (Ulutaş y Karaköy, 2019). Los servicios logísticos facilitan el movimiento de los productos y garantiza la seguridad, la rapidez y la rentabilidad del comercio internacional (Martí et al., 2014). Para realizar de manera efectiva las actividades logísticas, se deben considerar diversas funciones clave, como las infraestructuras y las regulaciones aduaneras gubernamentales (Yildirim y Mercangoz, 2020). Garantizar las condiciones óptimas en estas áreas es crucial para mantener la competitividad en las operaciones logísticas. Por lo tanto, la logística no tiene relevancia solo para los gobiernos, sino también para las organizaciones empresariales privadas (Yildirim y Mercangoz, 2020). En la actualidad, el sector de los servicios logísticos es una parte muy importante en la economía de un país, lo que significa una fuerte conexión entre las condiciones económicas y la demanda de los servicios logísticos (Kawa y Anholcer, 2019). La gestión de la cadena de suministro y la logística han contribuido notablemente a la ventaja competitiva sostenible (Çakır, 2017). Gestionar eficientemente los desafíos del transporte, almacenamiento, y embalaje es crucial, especialmente para mejorar la competitividad de empresarial y nacional. Sin embargo, unos servicios logísticos ineficientes tienen un impacto negativo tanto en los países como en las empresas, ya que aumentan los gastos operativos y disminuye el volumen de comercio (Martí et al., 2014).

Con el auge del comercio electrónico, la importancia de los servicios logísticos sigue creciendo. Los avances tecnológicos han elevado aún más la importancia de la logística tanto para los países como para las empresas privadas. Por esta razón evaluar el desempeño logístico es crucial para que los países evalúen su competitividad comercial internacional. El Índice de Desempeño Logístico (IDL) (Logistics Performance Index, LPI), publicado por el Banco Mundial (BM) cada dos años desde el 2007, sirve como herramienta de evaluación comparativa, ayudado a los países a identificar oportunidades y desafíos dentro de su desempeño logístico comercial (Çakır, 2017).

A pesar de que el último IDL publicado (2023) ofrece un análisis comparativo que integra, además de variables de carácter cualitativo, otras variables cuantitativas (movimiento de contenedores marítimos, cargas aéreas y envíos postales según las rutas comerciales y puntos de entradas), su cálculo sigue adoleciendo de simplicidad y subjetividad (Yildirim

y Adiguzel, 2020). Por ello, en este trabajo, que se encuadra en la línea propuesta por Beysenbaev y Dus (2020), con el fin de mejorar su evaluación, se propone una alternativa para evaluar el desempeño de los sistemas logísticos de los países, aprovechando el potencial de la combinación de técnicas de aprendizaje automático y los métodos de decisión multicriterio. Aplicamos el método del proceso de jerarquía analítica difuso para construir un modelo de evaluación del rendimiento logístico de los países, con la ayuda del método de análisis de clustering k-medias para determinar la ponderación de los diferentes indicadores que componen el Índice de Desarrollo Logístico.

Para ello, este trabajo se organiza de la siguiente manera. La primera parte es la propia introducción que nos ocupa, seguida de la sección de la revisión de la literatura. A continuación, se examina la metodología aplicada y se presentan los resultados obtenidos, finalizando con las principales conclusiones.

2. Revisión de la literatura

2.1. Alternativas para determinar el desempeño logístico.

La evaluación y medición del rendimiento logístico son aspectos críticos en el mundo empresarial y económico actual. Un instrumento fundamental para abordar esta evaluación es el Índice de Desempeño Logístico (IDL). Éste es una herramienta global que se utiliza para analizar y comparar la eficiencia de las cadenas de suministro y la logística en diversos países. En este contexto, el IDL no solo ofrece una visión integral del desempeño logístico a nivel macro, sino que también se ha convertido en un indicador clave para evaluar la competitividad y el desarrollo económico de los países. Prueba de ello, es la utilización del mismo como un referente en la literatura de los últimos años en el intento de proponer alternativas de evaluación con diferentes aplicaciones y objetivos, así como en distintos contextos geográficos

En este sentido, Martí et al. (2017), utilizan el Análisis Envolvente de Datos (Data Envelopment Analysis, DEA) para comparar e identificar aquellos países que son eficientes en términos de logística y aquellos que tienen margen de mejora. Rashidi et al. (2019) aplicaron la misma metodología, pero con el objetivo de evaluar la sostenibilidad del rendimiento logístico en los países de la OCDE y compararlo con la media nacional del IDL. En este caso el DEA se aplica para obtener una puntuación de eficiencia que refleje el grado en que se logra el desempeño logístico operativo sostenible. El estudio

busca identificar las naciones que están logrando un desempeño logístico eficiente y sostenible, al tiempo que ofrece una herramienta para evaluar y mejorar la gestión logística. Recientemente, este último autor, junto con Yu (2023), han propuesto un enfoque alternativo para evaluar el desempeño logístico a nivel global utilizando el método “*Common Weight Hybrid-Data Envelopment Analysis (H-DEA)*”. En este último caso, los autores sustituyen la asignación de pesos fijos a los criterios de evaluación logística por una asignación automática, en función de los datos disponibles.

Por otro lado, Beysenbaev y Dus (2020) proponen una forma de mejorar el IDL a través de un índice modificado que represente cualitativa y cuantitativamente una visión objetiva de los sistemas y subsistemas logísticos de 159 países, basados en datos estadísticos internacionales. Asimismo, Martí et al. (2014) en su trabajo, pretende analizar el impacto que tiene cada uno de los indicadores que componen el IDL en el comercio de los países emergentes empleando un modelo gravitacional, incluyendo factores tales como la distancia geográfica, el Producto Interior Bruto (PIB) per cápita, los costes de transporte y las barreras comerciales. Además, el estudio pretende analizar los posibles avances logísticos en los países, comparando los datos publicados del IDL en 2007 con los datos publicados en 2012.

Otros autores han utilizado modelos de regresión para determinar el desempeño logístico. De esta manera, d'Aleo (2015) ha utilizado un modelo de regresión lineal para analizar el papel mediador del IDL en la relación entre Índice de Competitividad Global y el Producto Interior Bruto (PIB) desde 2007 hasta 2014 en Europa (UE 28). Çemberci (2015) utilizó un modelo de regresión jerárquica con el objetivo de medir el efecto moderador del Índice de Competitividad Global sobre cada uno de los indicadores del IDL.

Asimismo, Cakir (2017) llevó a cabo una evaluación de la importancia de los criterios a través de la correlación intercriterio (Criteria Importance Through Intercriteria Correlation, CRITIC), la ponderación aditiva simple (Simple Additive Weighting, SAW) y los métodos de regresión lineal difusa de Peters para los países de la OCDE con los datos del IDL del 2014.

El método del indicador de distancia compuesto (Composite I-distance indicator, CIDI) fue utilizado por Petrovic et al. (2017) con el objetivo de obtener una nueva clasificación

del IDL de los países para los años 2016, 2014 y 2012, y posteriormente realizar una clasificación de los países y compararla con la clasificación original del Banco Mundial.

Finalmente, en otra línea de investigación, también basada en el IDL, Kálmán y Tóth (2021) examinan las relaciones entre la competitividad económica y el rendimiento logístico en los países del Grupo de Visegrád (Polonia, Hungría, República Checa y Eslovaquia) durante el período comprendido entre 2007 y 2018, destacando la importancia de una logística efectiva en el crecimiento económico y la competitividad a nivel regional.

2.2. Aplicación de métodos multicriterio de decisión en el ámbito del desempeño logístico.

Son cada vez más los trabajos que aplican métodos de decisión multicriterio para abordar el desempeño logístico de los países. Entre ellos destaca el artículo de Ulutaş y Karaköy (2019) que da una nueva ponderación a los diferentes indicadores del IDL. Para ello integra un método de combinación subjetiva (Step-Wise Weight Assessment Ratio Analysis, SWARA) y objetiva (Criteria for Ranks Incomplete Information Technique, CRITIC) con el objetivo de determinar los pesos de los criterios. La conclusión obtenida fue que el cambio en los pesos de los criterios afecta parcialmente a la clasificación de los países. Mešić et al. (2022) también emplearon el método multicriterio CRITIC para calcular los valores de ponderación de los seis indicadores del IDL. A continuación, se aplicó el método de Alternativas de medición y clasificación según la solución de compromiso (Measurement Alternatives and Ranking according to Compromise Solution, MARCOS) para los países de los Balcanes Occidentales. Este método es un enfoque de toma de decisiones utilizado en el campo del análisis de decisiones multicriterio. Es una técnica que ayuda a evaluar y clasificar alternativas en función de múltiples criterios.

Por otra parte, Işik et al. (2020) emplearon la técnica de toma de decisiones basada en múltiples criterios: MUlti-criteria ABility-ACcuracy (MABAC) y la Varianza Estadística (Statistical Variance, SV) para formar un modelo de toma de decisiones. El método SV se utilizó para ponderar los criterios IDL, mientras que el método MABAC se utiliza para evaluar y clasificar el rendimiento logístico de los países de Europa Central y Oriental. Los resultados que obtuvieron a partir del método SV demuestran que la puntualidad y la infraestructura son los criterios de rendimiento más y menos significativo,

respectivamente. A través del método MABAC los países que ocuparon los tres primeros puestos son la República Checa, Polonia y Hungría, respectivamente.

En otras investigaciones se emplean otros tipos de análisis de decisión multicriterio, como Rezaei et al. (2018) que utilizó la técnica Best-Worst para evaluar la importancia relativa o el ranking de los criterios IDL, en función de las preferencias de los encuestados. Esta investigación realizó una encuesta a 107 expertos en distintos países con el objetivo de asignar diferentes ponderaciones a los seis componentes del IDL y posteriormente aplicar estas puntuaciones a todos los países del estudio y hacer una comparación con el ranking original del Banco Mundial. Los resultados demostraron que las infraestructuras se consideran el componente más importante para el rendimiento logístico, considerando el seguimiento y la localización el factor menos importante.

Al igual que en este trabajo, existen varios autores que han aplicado la metodología AHP difusa en sus investigaciones para determinar las puntuaciones de ponderación de los seis indicadores. Entre ellos se incluye el estudio de Yildirim y Adiguzel (2020), en el cual utilizan el rendimiento logístico de los países de la OCDE entre los años 2010 y 2018 y lo comparan con las clasificaciones del IDL existente. Una vez calculados las ponderaciones a través del AHP difuso, emplean el método de evaluación de la relación aditiva gris (Grey Additive Ratio Assessment, ARAS-G) para evaluar los resultados logísticos de los países de la OCDE.

Otros autores, como Çalık et al. (2023), utilizan métodos híbridos de MCDM para realizar un análisis comparativo y evaluar el rendimiento logístico del IDL de 160 países. En este estudio se examinaron los métodos AHP-TOPSIS (Technique for Order of Preference by Similarity to Ideal Solution, TOPSIS), AHP-VIKOR (Ranking Multicriterio de Compromiso, VIKOR) y AHP-CODAS (Complex Proportional Assessment with Grey Relations, CODAS), tanto en entornos clásicos como en difusos. Además, se propone un enfoque de toma de decisión en grupo (Group Decision Making, GDM) para obtener las clasificaciones finales del IDL. Los resultados muestran que los países con las clasificaciones más altas son Alemania, Suecia, Países Bajos, Japón, Austria, Bélgica, Singapur, Reino Unido, Dinamarca, Hong Kong RAE (Región Administrativa Especial) y China, respectivamente. Por otro lado, el IDL es más bajo en Haití, Libia, Irak, Zimbabue, Eritrea, Burundi, Níger, Sierra Leona, Angola y Afganistán.

El método de evaluación proporcional compleja de alternativas (Complex Proportional Assessment of Alternatives, COPRAS-G) fue utilizado por Adıgüzel et al. (2020) para medir el rendimiento logístico de los países, proporcionando nuevas puntuaciones a los países seleccionados. El estudio se realizó sobre los 28 Estados miembros de la Unión Europea (UE) más 5 países candidatos (Albania, Macedonia, Montenegro, Serbia y Turquía). Además, aplicaron el método AHP difuso para calcular los pesos de los 6 indicadores tomando opiniones de expertos. Los resultados mostraron que el conjunto de datos obtenidos mediante el uso de COPRAS-G y AHP difuso representa unas puntuaciones muy próximas con la clasificación publicada por el Banco Mundial (BM), aunque existen pequeñas diferencias.

Recientemente, Miškić et al. (2023) evalúan el Índice de Desarrollo Logístico en los países de la Unión Europea (UE) mediante el uso también de un modelo de decisiones multicriterio, enfatizando la importancia de reevaluar los indicadores del desempeño logístico para una más objetiva y eficaz valoración.

2.3. Aplicación del clustering en el ámbito del desempeño logístico.

Las técnicas clustering es una herramienta estadística ampliamente utilizada para determinar un conjunto de datos que agrupa elementos que están lo suficientemente cerca entre sí y lo suficientemente lejos de otros elementos (Rokach y Maimon, 2005), es decir, tiene la mayor similitud intraclase y mínima similitud entre clases.

En el campo del desempeño logístico la técnica clustering también ha tenido sus aplicaciones, en diferentes contextos y objetivos. Así pues, Ulkhaq (2023) utiliza tres algoritmos de clasificación (k-means, k-medoids y clustering large applications) para agrupar los países en función del IDL. El análisis busca identificar patrones y similitudes que pueden ser relevantes para la planificación logística y la toma de decisiones a nivel global.

Por su parte, Gocer y Sener (2021) integran un algoritmo de agrupación K-means ponderado junto con una herramienta de toma de decisión multicriterio (MCDM). El método MCDM que se utiliza en este trabajo es el AHP, con el objetivo de medir la ponderación de los criterios para las regiones de Turquía. Tras aplicar ambas metodologías se deduce que Estambul, Ankara y Konya son las regiones más idóneas de Turquía en el ámbito logístico. Por el contrario, Kütahya es la peor alternativa.

Polat et al. (2022) utilizó los datos del Índice de Rendimiento Logístico de 2018 y los datos de las emisiones de CO₂ per cápita de 150 países para explicar la relación significativa entre el desempeño logístico de los países y sus emisiones de CO₂. El estudio se formuló en tres metodologías. La primera se realizó un análisis de conglomerados jerárquicos con los desempeños logísticos de los países y las emisiones de CO₂ per cápita. La segunda metodología se incluyó un análisis envolvente de datos para determinar la eficiencia de emisiones de CO₂ per cápita basadas en los subindicadores de desempeño logístico de los países. En la tercera etapa, se realizó un análisis de conglomerados no jerárquicos con las variables desempeño logístico y eficiencia de emisión de CO₂ per cápita de los países.

Otros de los autores que han incluido el clustering en sus investigaciones son Faria et al. (2015). En este caso es para evaluar el Índice de Desempeño Logístico (IDL) de Brasil en relación con sus principales competidores en el comercio internacional. Para ello, aplicaron técnicas estadísticas como el análisis de conglomerados y pruebas de comparación múltiple de medias para analizar los datos. Los resultados muestran que Brasil ocupa la posición 26 de los 39 competidores analizados, detrás de Sudáfrica, Kuwait, y Arabia Saudita. Los mejores son los principales países exportadores e importadores del mundo, es decir, Alemania, EE.UU, Japón, y los Países Bajos.

Por último, Alyoubi (2021) centra su análisis en el rendimiento logístico en Arabia Saudita, también mediante un enfoque de agrupamiento para identificar patrones y tendencias en el rendimiento logístico, proponiendo la implementación de las tecnologías 4.0 (blockchain, internet de las cosas, etc.) como alternativas para optimizar la gestión de la cadena de suministro y rastrear activos en dicho país.

3. Metodología

3.1. Datos

La primera fase de este trabajo aborda la preparación y extracción de los datos necesarios para su realización. Este trabajo utiliza los datos del Índice de Desempeño Logístico disponibles en línea en el Banco Mundial de manera periódica desde 2007. Este se calcula en base a la información recopilada, de una parte, a través de encuestas dirigidas a operadores logísticos internacionales sobre la calidad de las interacciones logísticas de los países con los que realizan transacciones comerciales; y, de otra parte, de los datos reales sobre el movimiento de contenedores marítimos, cargas aéreas y envíos postales

según las rutas comerciales y puntos de entradas. Los datos utilizados en este trabajo corresponden al último IDL publicado en 2023, que presenta la visión más reciente sobre el rendimiento de la logística comercial en 139 países (4.090 evaluaciones de 652 profesionales logísticos).

En una valoración de 1 a 5, donde la calificación más alta (5) refleja un mejor desempeño logístico y la calificación más baja (1) es un deficiente desempeño logístico, el IDL evalúa seis dimensiones agrupadas en diferentes indicadores cuya descripción se muestra en la Tabla 1.

Criterios	Indicadores	Definición
C1	Calidad y competencia logística	Calidad de los servicios logísticos proporcionados por operadores de transportes y agentes aduaneros.
C2	Puntualidad	Cumplimiento preciso de los plazos de entrega de los envíos.
C3	Infraestructuras	Calidad de las instalaciones de transporte marítimo, terrestre, ferroviario y aéreo. La percepción de los encuestados acerca de este indicador se evalúa en función de los métodos de transporte, así como del almacenamiento y movimiento de mercancías.
C4	Seguimiento y localización	Capacidad para rastrear y localizar envíos. Determinar la ubicación exacta y la trayectoria de cada bien resulta esencial hasta su entrega final. Este componente involucra a todos los actores en la Cadena de suministro, por lo que la trazabilidad es el resultado de una colaboración global.
C5	Aduanas	Agilidad de los procedimientos de despacho en términos de la rapidez, simplicidad y previsibilidad de los trámites realizados por los órganos de control aduanero
C6	Envíos internacionales	Facilidad para negociar tarifas competitivas en el transporte de mercancías.

Tabla 1. Componentes del Índice de Desempeño Logístico (IDL).
Fuente: Arvis et al. (2023).

Además, para la aplicación de la técnica de agrupamiento se han utilizado las variables recogidas en la Tabla 2.

Variable	Descripción	Categorías
Región	Localización geográfica de los diferentes países que participan el análisis	<ul style="list-style-type: none"> • Asia Oriental y Pacífico • Europa y Asia Central • América del Norte • Oriente Medio y Norte de África • África subsahariana • Asia Meridional • América Latina y Caribe
Importaciones de bienes y servicios (US\$ a precios constantes de 2010)	Valor de todos los bienes y otros servicios de mercado recibidos del resto del mundo.	<ul style="list-style-type: none"> • Low-imports (< 3.28¹¹) • Lower middle imports (3.28¹¹ - 6.6¹¹) • Middle imports (6.6¹¹ - 9.8¹¹) • Upper middle imports (9.8¹¹ - 1.3¹²) • Upper imports (≥ 1.3¹²)
Exportaciones de bienes y servicios (US\$ a precios constantes de 2010)	Valor de todos los bienes y demás servicios de mercado prestados al resto del mundo.	<ul style="list-style-type: none"> • Low-exports (<3.6¹¹) • Lower middle exports (3.6¹¹ - 7.2¹¹) • Upper middle exports (7.2¹¹ - 1.1¹²) • Upper exports (≥ 1.4¹²)
Volumen de ingresos	Ganancias que una entidad, individuo o país reciben durante un período de tiempo específico.	<ul style="list-style-type: none"> • Low income • Lower middle income • Upper middle income • High-income

Tabla 2. Variables de las transacciones comerciales.

Fuente: elaboración propia a partir del Banco Mundial (2022).

Toda la información recopilada fue integrada en una base de datos, para su posterior tratamiento estadístico con el software Orange 3.35.0.

3.2. K-means (Clustering)

La técnica de clustering es una técnica de aprendizaje automático no supervisado que se utiliza para identificar grupos de objetos de datos en función de características similares (Asensio et al., 2022; Nedyalkova et al., 2021; Rojas-Torres et al., 2022). Estas agrupaciones son útiles para explorar datos, identificar anomalías y hacer predicciones (Sun et al., 2008; Van Steenbergen y Mes, 2020).

En este artículo se realiza una clasificación de los países a nivel internacional para evaluar el desempeño logístico a partir de un conjunto de datos analizados. Se ha aplicado la agrupación tradicional K-means, propuesto por primera vez por MacQueen (1967). El

agrupamiento se realiza minimizando la suma de distancias entre cada objeto y el centroide de su grupo o clúster. Se suele usar la distancia cuadrática.

El proceso estándar de aplicación del algoritmo se compone de tres etapas (MacKay, 2003):

1. Inicialización: Una vez que se ha determinado el número de grupos, denotado como k , se posiciona k centroides en el espacio de datos. Esto se logra, por ejemplo, seleccionando dichos centroides de manera aleatoria.
2. Asignación de objetos al centroide: cada elemento del conjunto de datos es asignado al centroide más próximo en función de su distancia.
3. Actualización de centroides: las posiciones de los centroides de cada grupo se actualizan recalculando su posición como el promedio de las ubicaciones de los elementos pertenecientes a dicho grupo.

Los dos últimos pasos, asignación y actualización, se repiten iterativamente hasta que los centroides dejan de moverse o sus desplazamientos quedan por debajo de un umbral de distancia establecido.

Los elementos son representados a través de vectores numéricos en un espacio con d dimensiones (x_1, x_2, \dots, x_n) , y el algoritmo *k-means* forma k clústeres con el propósito de reducir al mínimo la suma de las distancias entre los elementos y sus respectivos centroides dentro de cada clúster $S = \{S_1, S_2, \dots, S_k\}$. La cuestión puede ser planteada de la siguiente manera:

$$\min E(\mu_i) = \min_S \sum_{i=1}^k \sum_{x_j \in S_i} \|x_j - \mu_i\|^2 \quad (1)$$

Donde S hace referencia al conjunto de datos en el cual los elementos son los objetos, x_j representados mediante vectores, en los que cada componente denota una característica o atributo. De esta manera, se obtendrán k grupos o clústeres junto a sus respectivos centroides μ_i .

En cada iteración de la actualización de los centroides, desde una perspectiva matemática, aplicamos la condición esencial de extremo a la función $E(\mu_i)$, que en el caso de la función cuadrática (1) se expresa como:

$$\frac{\partial E}{\partial \mu_i} = 0 \rightarrow \mu_i^{(t+1)} = \frac{1}{|S_1^{(t)}|} \sum_{x_j \in S_1^t} x_j \quad (2)$$

y se toma el promedio de los elementos de cada grupo como nuevo centroide.

El número óptimo de grupos se selecciona mediante el método del ancho de la silueta (Silhouette score), que mide la similitud entre cada punto de un grupo y lo compara con el punto más cercano del grupo vecino. El valor k que corresponde al ancho de silueta más grande indica la cantidad óptima de grupos para el conjunto de datos dado. El Gráfico 1 muestra que segmentar el conjunto de datos en 3 grupos maximiza la similitud dentro de los grupos y la diferencia entre los grupos.

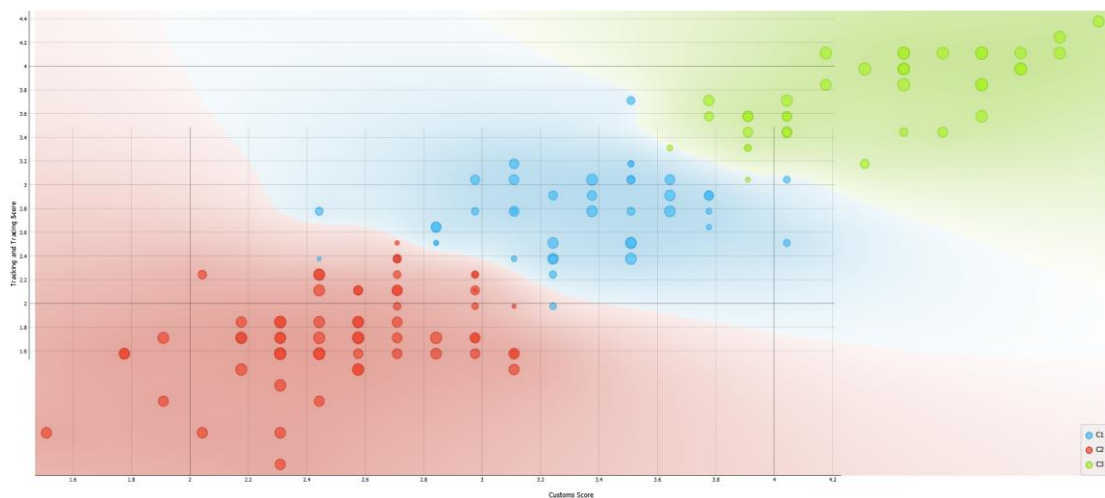


Gráfico 1. Número óptimo de clústeres usando ancho de Silhouette y diagrama de clúster.

Fuente: elaboración propia a partir de los datos del Banco Mundial.

Además, para determinar la importancia de cada uno de los 6 componentes del IDL, sobre la base del resultado de aplicar la técnica de agrupamiento, se ha utilizado el algoritmo seleccionado Relief. Este algoritmo fue desarrollado por Kira y Rendell (1992). Relief es un método de aprendizaje automático diseñado para la selección de características y la clasificación de datos. Este algoritmo recorre m instancias de entrenamiento aleatorio (R_i), escogidas sin repetición, donde m es un valor un parámetro definido por el usuario (Urbanowicz et al., 2018). En cada entrenamiento se calcula la distancia entra la instancia de referencia u objetivo y todas las demás instancias (Urbanowicz et al., 2018). Relief

identifica las dos instancias vecinas más cercanas a la referencia: una que pertenece a la misma clase, denominada el impacto más cercano (H), y la otra de la clase opuesta, denominada el error más cercano (M).

3.3. AHP difuso

El Analytic Hierarchy Process (AHP), conocido como Proceso Analítico Jerárquico en español, es un destacado método dentro del campo de toma de decisión multicriterio, desarrollado por primera vez por Saaty en la década de 1980 (Saaty, 1987) y empleado en una variedad de contexto dentro de los procedimientos de toma de decisión. El enfoque difuso de AHP puede ser considerado como una evolución avanzada del AHP convencional. A diferencia del AHP tradicional, el AHP difuso incorpora la incertidumbre y la imprecisión en el proceso de toma de decisiones. Mientras que AHP se basa en valores numéricos precisos para las comparaciones y priorizaciones, AHP difuso utiliza conjuntos difusos para representar la incertidumbre y permite una representación más flexible y realista de las preferencias y relaciones (Kabir & Hasin, 2011).

El método AHP difuso se emplea para deducir los pesos definitivos de los criterios, también a través de la comparación efectuada por un panel de expertos. En este trabajo vamos a sustituir la opinión subjetiva de los paneles de expertos, la cual nos proporciona la importancia de un indicador sobre otro, por el resultado que nos dé al aplicar el algoritmo k-means.

Para aplicar el AHP difuso deben emplearse los siguientes pasos (Chang, 1996):

1°. Determinación de la estructura del modelo jerárquico conforme al objetivo establecido, los criterios seleccionados (dimensiones del IDL) y las alternativas a evaluar (países) (Figura 1).

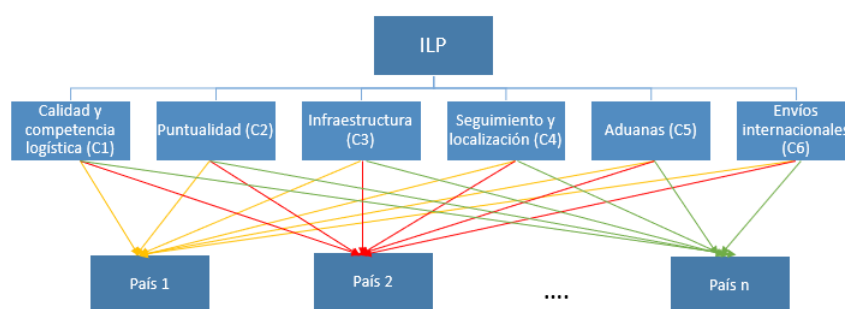


Figura 1. Estructura del modelo jerárquico.
Fuente: elaboración propia.

2°. Obtención de la matriz de comparación por pares usando números difusos (\widetilde{A}).

La matriz de comparación por pares se puede expresar de la siguiente manera:

$$\widetilde{A} = \begin{bmatrix} 1 & \widetilde{a}_{12} & \cdots & \widetilde{a}_{1n} \\ \widetilde{a}_{21} & 1 & \cdots & \widetilde{a}_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \widetilde{a}_{n1} & \widetilde{a}_{n2} & \cdots & 1 \end{bmatrix} \quad (3)$$

Para calcular los números difusos utilizaremos los resultados de haber aplicado el algoritmo Relief, explicado en el apartado anterior. Un número difuso se puede denotar como: $M = (l, m, u)$, donde $l \leq m \leq u$. Los parámetros $l; m; u$; indican el valor mas pequeño, el valor promedio y el valor más grande, respectivamente (Kabir & Hasin, 2011). Las variables lingüísticas que describen las ponderaciones de los criterios se muestran en la Tabla 3. Una variable lingüística se refiere a una variable en la cual sus valores se representan mediante términos lingüísticos. La noción de variable lingüística resulta altamente beneficiosa al abordar situaciones que resultan excesivamente intrincadas o que carecen de una definición precisa para ser adecuadamente descritas utilizando enunciados cuantitativos tradicionales (Kabir & Hasin, 2011).

Estado	Números difusos triangulares
Extremadamente fuerte	(7, 9, 9)
Muy fuerte	(5, 7, 9)
Fuertemente	(3, 5, 7)
Moderadamente	(1, 3, 5)
Igualmente	(1, 1, 1)

Tabla 3. Ponderaciones de los criterios de las variables lingüísticas.
Fuente: Yildirim y Mercangoz (2020).

3°. Cálculo del valor de la extensión sintética (S_i) para cada fila de la matriz de comparación por pares, mediante la siguiente fórmula:

$$S_i = \sum_{j=i}^m M_{gi}^j * \left[\sum_{i=1}^n \sum_{j=1}^m M_{gi}^j \right]^2 \quad (4)$$

Donde i representa el número de filas y j denota el número de columnas. M_{gi}^j son números difusos triangulares de la matriz de comparación pareada.

Lo valores de $\sum_{j=1}^m M_{gi}^j$, $\sum_{i=1}^n \sum_{j=1}^m M_{gi}^i$ y $[\sum_{i=1}^n \sum_{j=1}^m M_{gi}^i]^{-1}$ se puede calcular utilizando las siguientes formulas, respectivamente:

Para obtener $\sum_{j=1}^m M_{gi}^j$, utilizamos la siguiente formula:

$$\sum_{j=1}^m M_{gi}^j = \left(\sum_{j=1}^m l_j, \sum_{j=1}^m m_j, \sum_{j=1}^m u_j \right) \quad (5)$$

Para $\sum_{i=1}^n \sum_{j=1}^m M_{gi}^i$, utilizaremos la siguiente formula:

$$\sum_{i=1}^n \sum_{j=1}^m M_{gi}^i = \left(\sum_{i=1}^n l_i, \sum_{i=1}^n m_i, \sum_{i=1}^n u_i \right) \quad (6)$$

Finalmente, para obtener $[\sum_{i=1}^n \sum_{j=1}^m M_{gi}^i]^{-1}$, utilizaremos la siguiente formula:

$$\left[\sum_{i=1}^n \sum_{j=1}^m M_{gi}^j \right]^{-1} \left(\frac{1}{\sum_{i=1}^n u_i}, \frac{1}{\sum_{i=1}^n m_i}, \frac{1}{\sum_{i=1}^n l_i} \right) \quad (7)$$

Donde l_i, m_i, u_i indican el valor más pequeño, el valor promedio y el valor más grande, respectivamente, de los números difusos.

4°. Determinación de la magnitud de S_i uno con respecto al otro.

En general, si $M_1 = (l_1, m_1, u_1)$ y $M_2 = (l_2, m_2, u_2)$ son dos números difusos triangulares, como se muestra en la Figura 2, la magnitud de M_1 con respecto a M_2 pueden ser definidos como:

$$\begin{aligned} V(M_2 \geq M_1) &= hgt(M_1 \cap M_2) = \mu_{M_2}(d) \\ &= \begin{cases} 1, & \text{if } m_2 \geq m_1 \\ 0, & \text{if } l_1 \geq u_2 \\ \frac{l_1 - u_2}{(m_2 - u_2) - (m_1 - l_1)}, & \text{otherwise} \end{cases} \quad (8) \end{aligned}$$

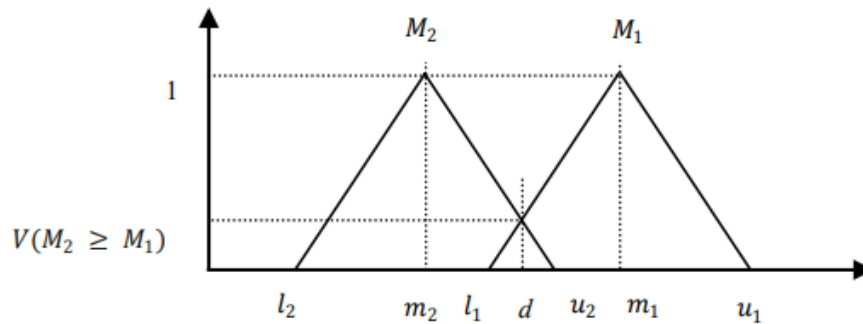


Figura 2. Variables lingüísticas para el peso de importancia de cada criterio.
Fuente: Aghdaie y Behzadian (2010).

Además, es posible calcular la dimensión de un número difuso triangular al considerar k como otro número difuso triangular, utilizando la siguiente fórmula:

$$V(M \geq M_1, M_2, \dots, M_k) = V[(M \geq M_1) \text{ and } (M \geq M_2) \text{ and } \dots (M \geq M_{1k})] = \min V(M \geq M_i) \quad i = 1, 2, 3, \dots, k \quad (9)$$

5°. Cálculo de los pesos de los criterios y alternativas en la matriz de comparación pareada.

La siguiente formula se puede utilizar para este fin:

$$d'(A_i) = \min V(S_i \geq S_k) \quad k = 1, 2, \dots, n, \quad k \neq i \quad (10)$$

Por lo tanto, el vector de pesos no normalizado se puede dar de la siguiente manera:

$$W' = (d'(A_1), d'(A_2), \dots, d'(A_n))^T \quad A_i (i = 1, 2, \dots, n) \quad (11)$$

6°. Calculo el vector de peso final

Para calcular el vector de peso final, se debe normalizar el vector de peso calculado en el paso anterior, por lo tanto:

$$W = (d(A_1), d(A_2), \dots, d(A_n))^T \quad (12)$$

4. Resultados

Los resultados de la aplicación del algoritmo k-mean se muestran en la Tabla 4, donde se recogen los países agrupados en los tres clústeres.

Clúster	Países			
C1	1. India	11. Slovak Republic	21. Egypt, Arab Rep.	31. Colombia
	2. Lithuania	12. Slovenia	22. North Macedonia	32. Costa Rica
	3. Portugal	13. Vietnam	23. Panama	33. Honduras
	4. Saudi Arabia	14. Brazil	24. Bosnia and Herzegovina	34. Mexico
	5. Türkiye	15. Bulgaria	25. Chile	35. Namibia
	6. Croatia	16. Cyprus	26. Indonesia	36. Argentina
	7. Czech Republic	17. Hungary	27. Peru	37. Montenegro
	8. Malta	18. Kuwait	28. Uruguay	38. Rwanda
	9. Oman	19. Romania	29. Antigua and Barbuda	39. Serbia
	10. Philippines	20. Botswana	30. Benin	40. Solomon Islands
C2	42. The Bahamas	56. Mali	71. Moldova	86. Burkina Faso
	43. Belarus	57. Nigeria	72. Mongolia	87. Fiji
	44. Djibouti	58. Russian Federation	73. Nicaragua	88. The Gambia
	45. El Salvador	59. Uzbekistan	74. Tajikistan	89. Ira.
	46. Georgia	60. Albania	75. Togo	90. Kyrgyz
	47. Kazakhstan	61. Algeria	76. Trinidad and Tobago	91. Madagascar
	48. Papua New Guinea	62. Armenia	77. Zimbabwe	92. Mauritania
	49. Paraguay	63. Bhutan	78. Bolivia	93. Syrian Arab Republic
	50. Ukraine	64. Central African Republic	79. Cambodia	94. Venezuela
	51. Bangladesh	65. Congo, Dem.Rep.	80. Gabon	95. Cuba
52. Congo, Rep.	66. Ghana	81. Guyana	96. Yemen, Rep.	
53. Dominican Republic	67. Grenada	82. Iraq	97. Angola	
54. Guatemala	68. Guinea	83. Lao PDR	98. Cameroon	
55. Guinea-Bissau	69. Jamaica	84. Liberia	99. Haiti	
	70. Mauritius	85. Sudan	100. Somalia	
			101. Afghanistan	
			102. Liby	
C3	103. Singapore	113. Sweden	123. Greece	133. Malaysia
	104. Finland	114. United Arab Emirates	124. Italy	134. New Zealand
	105. Denmark	115. France	125. Norway	135. Poland
	106. Germany	116. Japan	126. South Africa	136. Bahrain
	107. Netherlands	117. Spain	127. United Kingdom	137. Latvia
	108. Switzerland	118. Taiwan, China	128. Estonia	138. Qatar
	109. Austria	119. Korea, Rep.	129. Iceland	139. Thailand
	110. Belgium	120. United States	130. Ireland	
	111. Canada	121. Australia	131. Israel	
	112. Hong Kong SAR, China	122. China	132. Luxembourg	

Tabla 4. Países agrupados en tres clústeres.

Fuente: elaboración propia a partir de los datos de Banco Mundial.

En el Gráfico 2 podemos ver como se distribuyen los clústeres en función de su puntuación IDL de 2023. Se puede observar que en el clúster 1 (C1) se localizan los países con niveles medios de IDL. En el clúster 2 se distribuyen aquellos países con bajas puntuaciones del IDL. Por el contrario, el clúster 3 (C3) son aquellos países con altas puntuaciones.

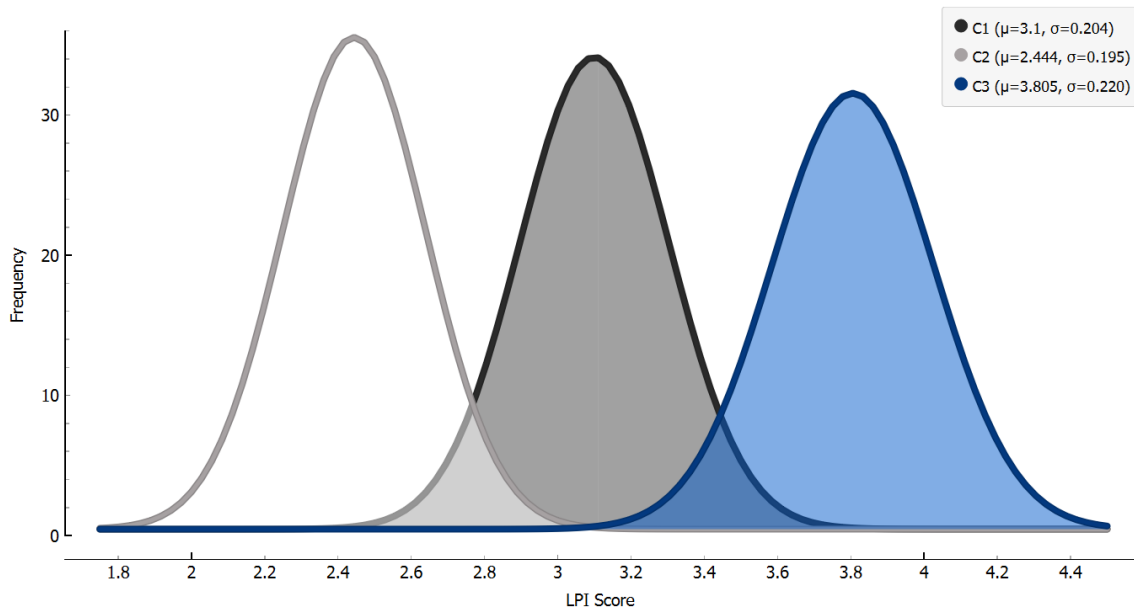


Gráfico 2. Distribución Normal del IDL.

Fuente: elaboración propia a partir de los datos de Banco Mundial .

A continuación, se presentan las principales características de los 3 grupos de países. Mediante un gráfico de caja (Gráfico 3), se puede observar los principales estadísticos descriptivos del IDL para los tres grupos de países identificados.

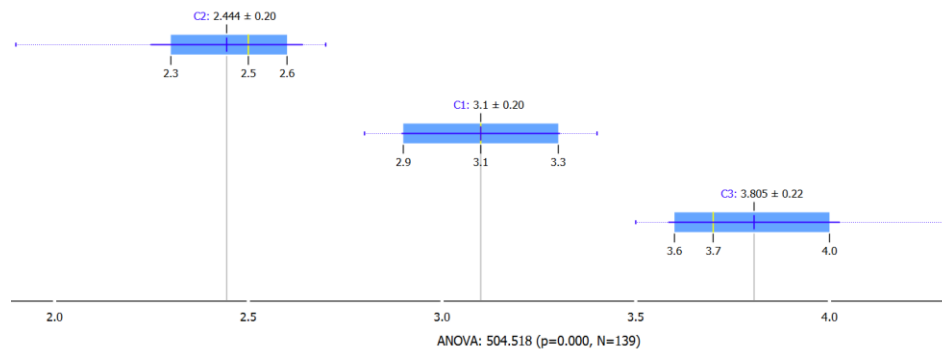


Gráfico 3. Diagrama de caja del IDL.

Fuente: elaboración propia a partir de los datos de Banco Mundial.

En el Gráfico 3 podemos observar que en el clúster 1 (C1), la media y la mediana coinciden. En el clúster 2 (C2), la media (2.4) es inferior a mediana (2.5). Y, por el contrario, en el clúster 3 (C3), la media (3.8) es superior a la mediana (3.7).

Con respecto a la dispersión, podemos ver que el clúster 2 (C2) tiene una mayor dispersión de los datos, concentrándose la mayor parte en el primer cuartil. Por el contrario, el clúster 1 (C1) es el que tiene menor dispersión concentrándose mayoritariamente en el segundo y tercer cuartil.

De la misma manera, en los Gráficos 4 a) b) c) d) e) y f), se visualizan las valoraciones de los diferentes componentes del IDL para los tres grupos de países.

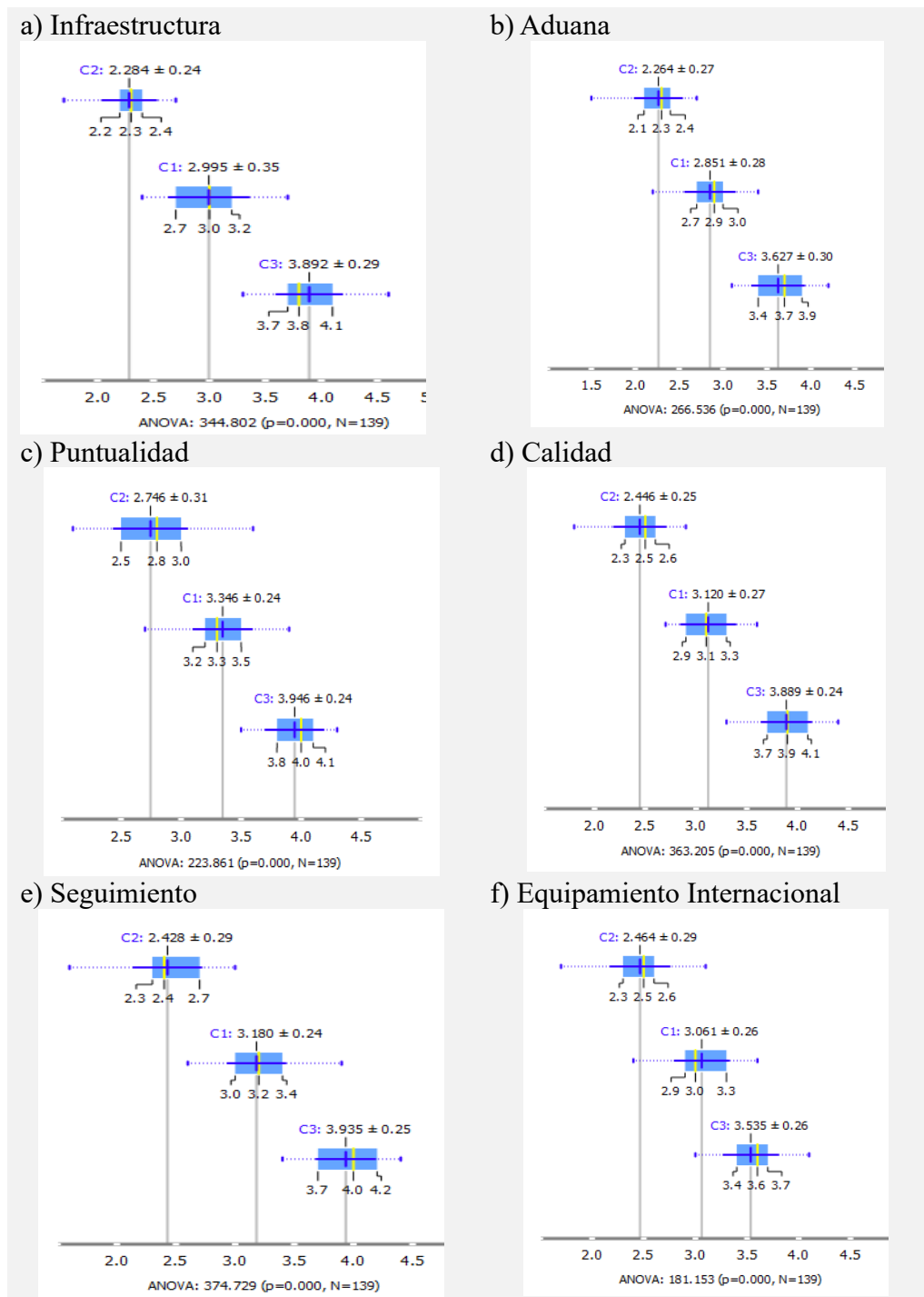


Gráfico 4. Diagrama de caja para cada componente del IDL.
 Fuente: elaboración propia a partir de los datos de Banco Mundial.

En todas las variables analizadas, los países del clúster 3 (C3) tienen una valoración promedia superior al resto de los otros grupos de países.

Posteriormente, representamos gráficamente, en el Tabla 5, las variables analizadas por clúster para encontrar diferencias entre ellas y poder reconocer variables discriminatorias.

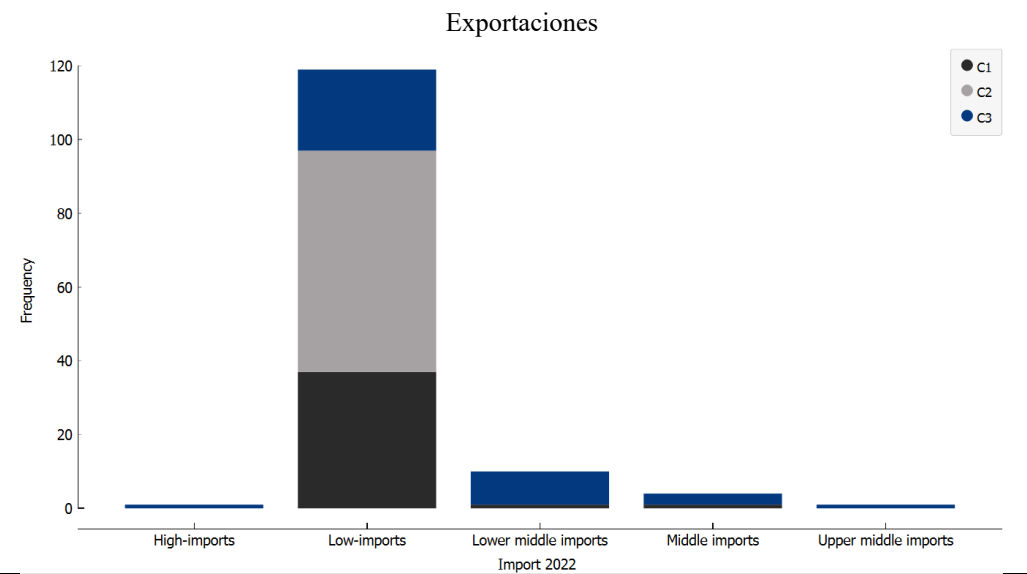
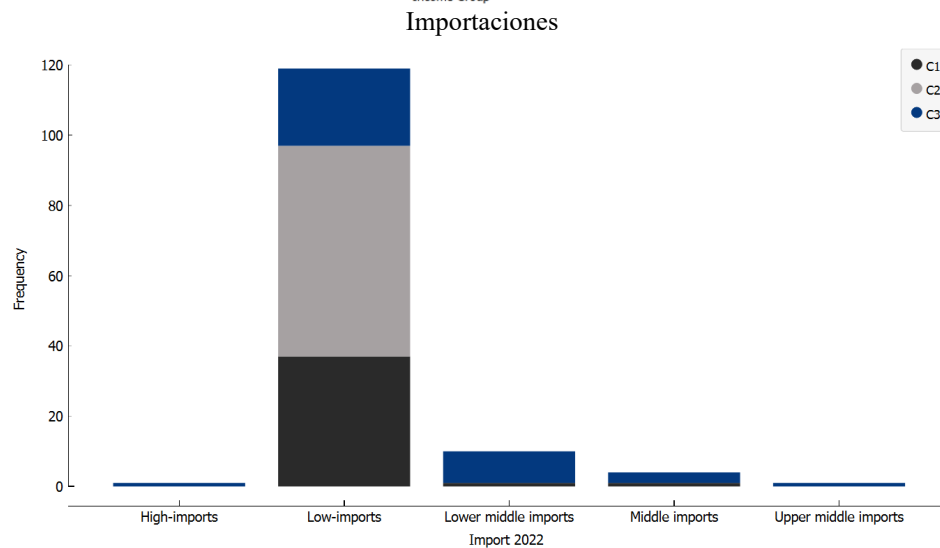
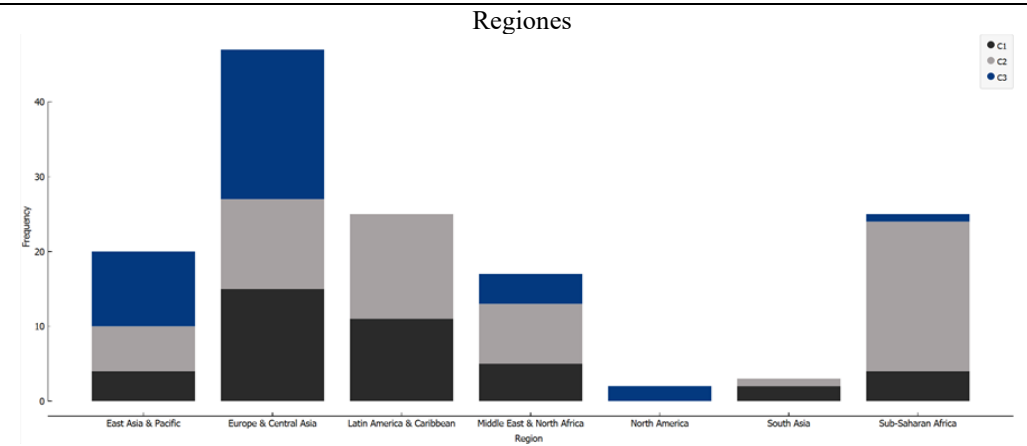
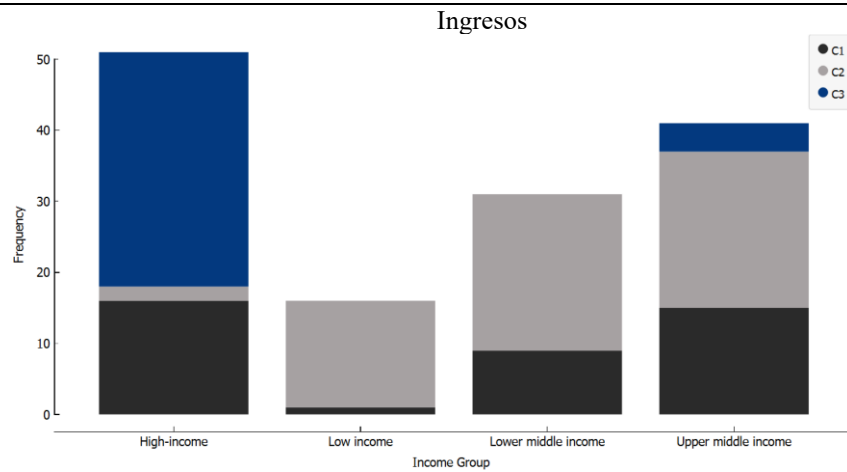


Tabla 5. Distribución de los clústeres de países por variables y clúster.
Fuente: elaboración propia a partir de los datos de Banco Mundial.

En los ingresos podemos observar cómo se distribuyen los países por clúster y su nivel de ingresos. Los países con mayores niveles de ingresos y mayor puntuación del IDL se localizan mayoritariamente en clúster 3 (C3). Por el contrario, los países con niveles de ingresos medios (bajos y alto) y bajos se ubican tanto en el clúster 1 (C1), como el clúster 2 (C2).

El gráfico de las regiones clasifica los distintos clústeres por áreas. Los países del clúster 3 (C3) se concentra principalmente en Europa, Asia y Norteamérica.

Con respecto a las exportaciones e importaciones, podemos destacar que la mayor parte de los países poseen bajos niveles de exportaciones e importaciones.

A continuación, el resultado de aplicar el algoritmo Relief para ponderar las diferentes dimensiones del IDL se refleja en la Tabla 6. A partir de este ranking se aplica el método AHP difuso y se procede a construir la matriz de comparación pareada difusa de los diferentes criterios del IDL (Tabla 7), determinado la importancia que tiene un indicador sobre otro, en base a la escala de Saaty (ver Tabla 3).

Criterios	Indicadores	Relieff
C1	Calidad y competencia logística	0,227
C2	Puntualidad	0,213
C3	Infraestructura	0,212
C4	Seguimiento y localización	0,199
C5	Aduanas	0,198
C6	Envíos internacionales	0,159

Tabla 6. Puntuación resultante de aplicar el algoritmo Relief.
Fuente: elaboración propia a partir de los datos del Banco Mundial.

Criterios	C1			C2			C3			C4			C5			C6		
C1	1	1	1	2	3	4	2	3	4	4	5	6	4	5	6	9	9	9
C2	0,25	0,33	0,50	1	1	1	1	1	1	2	3	4	2	3	4	6	7	8
C3	0,25	0,33	0,50	1,00	1,00	1,00	1	1	1	1	1	1	2	3	4	6	7	8
C4	0,17	0,20	0,25	0,25	0,33	0,50	1,00	1,00	1,00	1	1	1	1	1	1	4	5	6
C5	0,17	0,20	0,25	0,25	0,33	0,50	0,25	0,33	0,50	1,00	1,00	1,00	1	1	1	4	5	6
C6	0,11	0,11	0,11	0,13	0,14	0,17	0,13	0,14	0,17	0,17	0,20	0,25	0,17	0,20	0,25	1	1	1

Tabla 7. Matriz difusa triangular de comparaciones pareadas.
Fuente: elaboración propia a partir de los datos de Banco Mundial.

De acuerdo con el proceso descrito en el anterior apartado, el valor de la extensión sintética (S_i) para cada fila de la matriz de comparación por pares se recoge en la Tabla 8.

Criterios	l_j	m_j	u_j
C1	2,884	3,557	4,160
C2	1,348	1,661	2,000
C3	1,201	1,383	1,587
C4	0,742	0,833	0,953
C5	0,589	0,693	0,849
C6	0,191	0,212	0,240
C7	0,000	0,000	0,000
Total	6,955	8,339	9,790
S_i	0,144	0,120	0,102

Tabla 8. Valor de la extensión sintética (S_i).

Fuente: elaboración propia a partir de los datos de Banco Mundial.

Después, se calcula los pesos difusos de los números triangulares.

CRI	W_i		
C1	0,254	0,395	0,599
C2	0,119	0,184	0,288
C3	0,106	0,154	0,229
C4	0,065	0,092	0,137
C5	0,052	0,077	0,122
C6	0,017	0,024	0,035

Tabla 9. Pesos difusos triangulares.

Fuente: elaboración propia a partir de los datos del Banco Mundial.

Por último, calculamos el vector de peso final (Mi) y lo normalizamos (Ni).

CRI	Mi	Ni	Rank
C1	0,416	0,423	1
C2	0,197	0,201	2
C3	0,163	0,165	3
C4	0,098	0,100	4
C5	0,084	0,085	5
C6	0,025	0,025	6

Tabla 10. Criterio de peso promedio (Mi) y criterio de peso normalizado (Ni).
Fuente: elaboración propia a partir de los datos del Banco Mundial.

Con las ponderaciones obtenidas anteriormente de cada componente, y conforme al modelo jerarquizado recogido en la Figura 1, se procede a evaluar el nuevo IDL de cada país. Los resultados se presentan en la Tabla 11.

En la nueva clasificación propuesta (IDL-AHP), observamos diferencias significativas respecto al IDL publicado por Banco Mundial. Un porcentaje reducido mantiene su posición en el nuevo ranking (14 %), el 37% suben al menos una posición y 48% restante bajan al menos una posición. En ambos indicadores el país que lidera el ranking (Singapur) y el que está a la cola se mantiene (Afganistan).

Podemos destacar que los que países que más posiciones han subido con el nuevo ranking son regiones menos desarrolladas como Guyana (23), Tajikistan (21), República Centro Africana (20), Bhutan (18) o Cuba (17). Además, estos países se encuentran en las últimas posiciones de ranking del Banco Mundial. Con respecto a los países que más bajan destacan Bahamas (24), Algeria (22), Granada (20), Guinea (19) o Mali (15). En el caso de España solamente bajaría un puesto, pasado de la posición 14 a la 15.

IDL		IDL Banco Mundial		Comparación de la clasificación	ITA	Italy	3,79	21	23	↑ 2
Código ISO	Países	Score	Clasificación	Clasificación	JAM	Jamaica	2,68	86	99	↑ 13
AFG	Afghanistan	1,91	139	139	JPN	Japan	4,02	13	13	⇒ 0
AGO	Angola	2,39	120	134	KAZ	Kazakhstan	2,67	87	84	↓ -3
ALB	Albania	2,63	94	102	KGZ	Kyrgyz Republic	2,29	129	126	↓ -3
ARE	United Arab Emirates	4,15	6	11	KHM	Cambodia	2,43	115	115	⇒ 0
ARG	Argentina	2,95	66	74	KOR	Korea, Rep.	3,88	16	18	↑ 2
ARM	Armenia	2,43	116	104	KWT	Kuwait	3,21	55	54	↓ -1
ATG	Antigua and Barbuda	2,85	74	68	LAO	Lao PDR	2,45	112	115	↑ 3
AUS	Australia	3,93	14	19	LBR	Liberia	2,36	123	117	↓ -6
AUT	Austria	3,87	17	11	LBY	Libya	1,92	138	138	⇒ 0
BEL	Belgium	4,13	7	7	LKA	Sri Lanka	2,79	78	77	↓ -1
BEN	Benin	2,87	72	72	LTU	Lithuania	3,42	42	38	↓ -4
BFA	Burkina Faso	2,34	124	128	LUX	Luxembourg	3,73	24	27	↑ 3
BGD	Bangladesh	2,64	91	97	LVA	Latvia	3,64	32	34	↑ 2
BGR	Bulgaria	3,33	47	52	MDA	Moldova	2,57	100	100	⇒ 0
BHR	Bahrain	3,49	39	37	MDG	Madagascar	2,18	133	132	↓ -1
BHS	The Bahamas	2,53	104	80	MEX	Mexico	2,88	70	65	↓ -5
BIH	Bosnia and Herzegovina	3,03	63	65	MKD	North Macedonia	3,27	50	58	↑ 8
BLR	Belarus	2,84	75	82	MLI	Mali	2,51	105	90	↓ -15
BOL	Bolivia	2,49	108	119	MLT	Malta	3,40	43	43	⇒ 0
BRA	Brazil	3,37	45	54	MNE	Montenegro	2,85	73	71	↓ -2
BTN	Bhutan	2,61	96	114	MNG	Mongolia	2,45	113	112	↓ -1
BWA	Botswana	3,28	48	57	MRT	Mauritania	2,30	128	124	↓ -4
CAF	Central African Republic	2,71	82	102	MUS	Mauritius	2,57	101	96	↓ -5
CAN	Canada	4,12	8	5	MYS	Malaysia	3,65	31	28	↓ -3
CHE	Switzerland	4,20	3	3	NAM	Namibia	2,81	77	70	↓ -7
CHL	Chile	2,98	65	63	NGA	Nigeria	2,48	109	95	↓ -14
CHN	China	3,76	23	21	NIC	Nicaragua	2,47	110	110	⇒ 0
CMR	Cameroon	2,07	136	136	NLD	Netherlands	4,03	12	7	↓ -5
COG	Congo, Rep.	2,67	88	90	NOR	Norway	3,77	22	21	↓ -1
COL	Colombia	3,12	58	63	NZL	New Zealand	3,71	28	28	⇒ 0
CRI	Costa Rica	2,92	69	68	OMN	Oman	3,27	49	46	↓ -3
CUB	Cuba	2,44	114	131	PAN	Panama	3,06	60	59	↓ -1
CYP	Cyprus	3,25	52	56	PER	Peru	2,82	76	65	↓ -11
CZE	Czech Republic	3,42	41	45	PHL	Philippines	3,36	46	48	↑ 2
DEU	Germany	4,16	4	4	PNG	Papua New Guinea	2,72	81	80	↓ -1
DJI	Djibouti	2,87	71	78	POL	Poland	3,71	26	31	↑ 5
DNK	Denmark	4,15	5	6	PRT	Portugal	3,53	36	40	↑ 4
DOM	Dominican Republic	2,71	83	88	PRY	Paraguay	2,66	89	85	↓ -4
DZA	Algeria	2,18	134	112	QAT	Qatar	3,71	27	36	↑ 9
EGY	Egypt	2,93	68	60	ROM	Romania	3,19	57	51	↓ -6
ESP	Spain	3,93	15	14	RUS	Russian Federation	2,63	95	93	↓ -2
EST	Estonia	3,70	29	28	RWA	Rwanda	2,94	67	74	↑ 7
FIN	Finland	4,11	9	2	SAU	Saudi Arabia	3,46	40	40	⇒ 0
FJI	Fiji	2,27	132	130	SDN	Sudan	2,38	122	121	↓ -1
FRA	France	3,86	18	16	SGP	Singapore	4,39	1	1	⇒ 0
GAB	Gabon	2,42	117	119	SLB	Solomon Islands	3,03	64	74	↑ 10
GBR	United Kingdom	3,60	34	24	SLV	El Salvador	2,77	80	85	↑ 5
GEO	Georgia	2,60	97	83	SOM	Somalia	1,97	137	137	⇒ 0
GHA	Ghana	2,57	102	107	SRB	Serbia	2,64	90	78	↓ -12
GIN	Guinea	2,32	126	107	SVK	Slovak Republic	3,40	44	47	↑ 3
GMB	Gambia, The	2,27	131	125	SVN	Slovenia	3,24	53	44	↓ -9
GNB	Guinea-Bissau	2,69	85	89	SWE	Sweden	4,25	2	9	↑ 7
GRC	Greece	3,83	20	20	SYR	Syrian Arab Republic	2,28	130	128	↓ -2
GRD	Grenada	2,41	118	98	TGO	Togo	2,50	107	100	↓ -7
GTM	Guatemala	2,59	98	90	THA	Thailand	3,68	30	34	↑ 4
GUY	Guyana	2,57	99	122	TJK	Tajikistan	2,69	84	105	↑ 21
HKG	Hong Kong SAR, China	4,04	11	9	TTO	Trinidad and Tobago	2,64	93	105	↑ 12
HND	Honduras	2,77	79	72	TUR	Turkey	3,51	38	38	⇒ 0
HRV	Croatia	3,23	54	48	TWN	Taiwan, China	4,05	10	15	↑ 5
HTI	Haiti	2,09	135	135	UKR	Ukraine	2,64	92	87	↓ -5
HUN	Hungary	3,07	59	52	URY	Uruguay	3,04	62	61	↓ -1
IDN	Indonesia	3,05	61	62	USA	United States	3,85	19	17	↓ -2
IND	India	3,26	51	42	UZB	Uzbekistan	2,50	106	93	↓ -13
IRL	Ireland	3,51	37	31	VEN	Venezuela	2,39	121	127	↑ 6
IRN	Iran, Islamic Rep.	2,32	127	122	VNM	Vietnam	3,20	56	50	↓ -6
IRQ	Iraq	2,33	125	117	YEM	Yemen, Rep.	2,41	119	133	↑ 14
ISL	Iceland	3,64	33	33	ZAF	South Africa	3,57	35	25	↓ -10
ISR	Israel	3,72	25	26	ZAR	Congo, Dem. Rep.	2,56	103	110	↑ 7
					ZWE	Zimbabwe	2,46	111	107	↓ -4

Tabla 11. Resultados IDL-AHP.

Fuente: elaboración propia a partir de los datos del Banco Mundial.

A continuación, en la Tabla 12 se presenta el “*Top Ten*” de la clasificación IDL-AHP.

Código ISO	Países	IDL		IDL Banco	Comparación de la clasificación
		Score	Clasificación	Mundial Clasificación	
SGP	Singapore	4,39	1	1	0
SWE	Sweden	4,25	2	9	7
CHE	Switzerland	4,20	3	3	0
DEU	Germany	4,16	4	4	0
DNK	Denmark	4,15	5	6	1
ARE	United Arab Emirates	4,15	6	11	5
BEL	Belgium	4,13	7	7	0
CAN	Canada	4,12	8	5	-3
FIN	Finland	4,11	9	2	-7
TWN	Taiwan, China	4,05	10	15	5

Tabla 12. Clasificación de los 10 primeros países del IDL-AHP.

Fuente: elaboración propia a partir de los datos del Banco Mundial.

En esta tabla podemos destacar la subida de Suecia (7 posiciones), pasando del noveno al segundo puesto y el ascenso de 5 posiciones de Emiratos Árabes y de Taiwán. También cabe destacar el descenso de 7 posiciones de Finlandia, pasando del segundo puesto, que ocupaba en la clasificación del Banco Mundial, al noveno puesto que ocupa en la nueva clasificación IDL-AHP.

Por otra parte, en la Tabla 13 se presentan los 10 últimos países de la clasificación IDL-AHP.

Código ISO	Países	IDL		IDL Banco	Comparación de la clasificación
		Score	Clasificación	Mundial Clasificación	
SYR	Syrian Arab Republic	2,28	130	128	-2
GMB	Gambia, The	2,27	131	125	-6
FJI	Fiji	2,27	132	130	-2
MDG	Madagascar	2,18	133	132	-1
DZA	Algeria	2,18	134	112	-22
HTI	Haiti	2,09	135	135	0
CMR	Cameroon	2,07	136	136	0
SOM	Somalia	1,97	137	137	0
LBY	Libya	1,92	138	138	0
AFG	Afghanistan	1,91	139	139	0

Tabla 13. Clasificación de los 10 últimos países del IDL-AHP.

Fuente: elaboración propia a partir de los datos del Banco Mundial.

En las 10 últimas posiciones, predomina el descenso de 22 puestos de Argelia que pasa a ocupar la posición 134 de la clasificación IDL-AHP. Por último, hay que destacar que los últimos 5 países de se mantienen sin cambios en las dos clasificaciones.

5. Conclusiones

El objetivo en este trabajo ha consistido en proponer un método alternativo para calcular un indicador del rendimiento logístico de los países aprovechando el potencial de la combinación de técnicas de aprendizaje automático y métodos de decisión multicriterio.

Debido a la gran cantidad de variables a considerar en el análisis del rendimiento logístico de los diferentes países, alternativas propuestas para mejorar su evaluación son bienvenidas en este campo de investigación, propiciando un marco de reflexión sobre la idoneidad de los indicadores claves del rendimiento de los sistemas logísticos existentes, buscando la mayor objetividad y evitar en lo posible la imparcialidad de los mismos.

Aprovechando la información del IDL referente a nivel mundial, la aplicación de técnicas de clustering permite agrupar los países con patrones similares, no sólo con fines meramente descriptivos en función de variables relevantes como las transacciones comerciales entre países, sino también para ponderar los diferentes componentes del IDL de una manera razonable. Esto, a su vez, permite, a través de una técnica multicriterio, como la aplicada en este trabajo (AHP-difuso), reevaluar el papel que desempeña cada indicador (calidad, puntualidad, infraestructura, trazabilidad, aduanas y envíos internacionales) en el cómputo total del IDL, eliminando la subjetividad implícita e intentando en todo momento que el análisis se sustente en la racionalidad que proporciona esta técnica, ampliamente utilizada en el ámbito en el que se encuadra este estudio.

Por ello, sobre la base del IDL, el indicador propuesto en este trabajo puede constituir una herramienta de evaluación mejorada que sirva como un indicador clave del rendimiento de las cadenas de suministros para todas las partes implicadas, con especial referencia en el desarrollo de estrategias en el ámbito de la logística y transporte de cada país. Asimismo, al igual que el IDL, permite no sólo la comparación entre países, sino la de un mismo país en diferentes momentos del tiempo. Los resultados muestran diferencias significativas en el análisis comparativo de los sistemas logísticos de los diferentes países respecto al IDL publicado por el Banco Mundial.

Finalmente, decir que somos conscientes de las limitaciones del presente trabajo, que constituyen la base para continuar en dicha línea de investigación, valorando el impacto de otras variables en el análisis, así como la adaptación de la metodología empleada.

Referencias

Adıgüzel Mercangöz, B., Yıldırım, B. F., & Kuzu Yıldırım, S. (2020). Time period based COPRAS-G method: Application on the logistics performance index. *Scientific Journal of Logistics*, 16(2), 239-250. <http://doi.org/10.17270/J.LOG.2020.432>

Aghdaie, M. H., & Behzadian, M. (2010). A hybrid fuzzy MCDM approach to thesis subject selection. *J. Mat. Comp. Sci*, 1(4), 355-365.

Alyoubi, B. A. (2021). Clustering analysis of logistics performance in saudi arabia: A roadmap to cloud computing and IoT & blockchain solutions. *International Transaction Journal of Engineering, Management, & Applied Sciences & Technologies*, 12(7), 1-14.

Arvis, J. F., Ojala, L., Shepherd, B., Ulybina, D., & Wiederer, C. (2023). *Connecting to compete 2023: Trade logistics in an uncertain global economy-the logistics performance index and its indicators*. Washington DC.

Asensio, E., Almeida, A., Galiano, A., & Martín-Álvarez, J. M. (2022). Using Customer Knowledge Surveys to Explain Sales of Postgraduate Programs: A Machine Learning Approach. *International Journal of Interactive Multimedia and Artificial Intelligence*, 17(3). <https://doi.org/10.9781/ijimai.2022.01.008>

Banco Mundial. (2023). Logistics performance index (LPI). Recuperado de: <https://lpi.worldbank.org/international/global>

Beysenbaev, R., & Dus, Y. (2020). Proposals for improving the logistics performance index. *The Asian Journal of Shipping and Logistics*, 36(1), 34-42. <https://doi.org/10.1016/j.ajsl.2019.10.001>

Çakır, S. (2017). Measuring logistics performance of OECD countries via fuzzy linear regression. *Journal of Multi-Criteria Decision Analysis*, 24(3-4), 177-186. <https://doi.org/10.1002/mcda.1601>

Çalık, A., Erdebilli, B., & Özdemir, Y. S. (2023). Novel integrated hybrid multi-criteria decision-making approach for logistics performance index. *Transportation Research Record*, 2677(2), 1392-1400. <https://doi.org/10.1177/03611981221113314>

Çemberci, M., Civelek, M. E., & Canbolat, N. (2014). The moderator effect of global competitiveness index on dimensions of logistics performance index. *Procedia-Social and Behavioral Sciences*, 195, 1514-1524. <https://doi.org/10.1016/j.sbspro.2015.06.453>

Chang, D. Y. (1992). Extent analysis and synthetic decision. *Optimization Techniques and Applications*, 1(1), 352-355.

d'Aleo, V. (2015). The mediator role of logistic performance index: A comparative study. *Journal of International Trade, Logistics and Law*, 1(1), 1-7.

Faria, R. N. D., Souza, C. S. D., & Vieira, J. G. V. Evaluation of logistic performance indexes of Brazil in the international trade. *RAM (Revista De Administração Mackenzie)*, 16, 213-235.

<https://doi.org/10.1590/1678-69712015/administracao.v16n1p213-235>

Gocer, F., & Sener, N. (2022). Spherical fuzzy extension of AHP-ARAS methods integrated with modified k-means clustering for logistics hub location problem. *Expert Systems*, 39(2). <https://doi.org/10.1111/exsy.12886>

Işik, Ö., Aydin, Y., & Koşarolu, S. (2020). The assessment of the logistics performance index of CEE countries with the new combination of SV and MABAC methods. *LogForum*, 16(4), 549-559. <http://doi.org/10.17270/J.LOG.2020.504>

Kabir, G., & Hasin, M. A. A. (2011). Comparative analysis of AHP and fuzzy AHP models for multicriteria inventory classification. *International Journal of Fuzzy Logic Systems*, 1(1), 1-16.

Kálmán, B., & Tóth, A. (2021). Links between the economy competitiveness and logistics performance in the Visegrád group countries: Empirical evidence for the years 2007-2018. *Entrepreneurial Business and Economics Review*, 9(3), 169-190. <https://doi.org/10.15678/EBER.2021.090311>

Kawa, A., & Anholcer, M. (2019). Intangible assets as a source of competitive advantage for logistics service providers. *Transport Economics and Logistics*, 78, 29-41. <https://doi.org/10.26881/etil.2018.78.03>

Kira, K., & Rendell, L. A. (1992). The feature selection problem: Traditional methods and a new algorithm. In *Proceedings of the Tenth National Conference on Artificial Intelligence*, 129-134.

MacKay, D. J. (2003). Information theory, inference and learning algorithms. *Cambridge University Press*.

Maimon, O., & Rokach, L. (2005). Data mining and knowledge discovery handbook. *Springer*, 2 (2005).

Martí, L., Martín, J. C., & Puertas, R. (2017). A DEA-logistics performance index. *Journal of Applied Economics*, 20(1), 169-192. [https://doi.org/10.1016/S1514-0326\(17\)30008-9](https://doi.org/10.1016/S1514-0326(17)30008-9)

Martí, L., Puertas, R., & García, L. (2014). The importance of the logistics performance index in international trade. *Applied Economics*, 46(24), 2982-2992. <https://doi.org/10.1080/00036846.2014.916394>

Mešić, A., Miškić, S., Stević, Ž., & Mastilo, Z. (2022). Hybrid MCDM solutions for evaluation of the logistics performance index of the western balkan countries. *Economics*, 10(1), 13-34. <https://doi.org/10.2478/eoik-2022-0004>

Miškić, S., Stević, Ž., Tadić, S., Alkhayyat, A. y Krstić, M. (2023). Assessment of the LPI of the EU countries using MCDM model with an emphasis on the importance of criteria. *World Review of Intermodal Transportation Research*, 11(3), 258-279. <https://doi.org/10.1504/WRITR.2023.132501>

Nedyalkova, M., Madurga, S., & Simeonov, V. (2021). Combinatorial k-means clustering as a machine learning tool applied to diabetes mellitus type 2. *International Journal of Environmental Research and Public Health*, 18(4), 1919. <https://doi.org/10.3390/ijerph18041919>

Petrović, M., Jeremić, V., & Bojković, N. (2017). Exploring logistics performance index using I-distance statistical approach.

Polat, M., Kara, K., & Yalcin, G. C. (2022). Clustering countries on logistics performance and carbon dioxide (CO2) emission efficiency: An empirical analysis. *Business and Economics Research Journal*, 13(2), 221-238. <http://dx.doi.org/10.20409/berj.2022.370>

Rashidi, K., & Cullinane, K. (2019). Evaluating the sustainability of national logistics performance using data envelopment analysis. *Transport Policy*, 74, 35-46. <https://doi.org/10.1016/j.tranpol.2018.11.014>

Rezaei, J., van Roekel, W. S., & Tavasszy, L. (2018). Measuring the relative importance of the logistics performance index indicators using best worst method. *Transport Policy*, 68, 158-169. <https://doi.org/10.1016/j.tranpol.2018.05.007>

Rojas-Torres, I. L., Ahmad, M., Martín Álvarez, J. M., Golpe, A. A., & Gil Herrera, R. D. J. (2022). Mental health, suicide attempt, and family function for adolescents' primary health care during the COVID-19 pandemic. *F1000Research*, 11, 529. <https://doi.org/10.12688/f1000research.109603.2>

Rokach L., & Maimon, O. (2005). "Clustering methods" in *Data mining and knowledge discovery handbook* (pp. 321-352). Springer, Boston, MA.

Saaty, R. W. (1987). The analytic hierarchy process—what it is and how it is used. *Mathematical Modelling*, 9(3-5), 161-176.

Sun, Z. L., Choi, T. M., Au, K. F., & Yu, Y. (2008). Sales forecasting using extreme learning machine with applications in fashion retailing. *Decision Support Systems*, 46(1), 411-419. <https://doi.org/10.1016/j.dss.2008.07.009>.

Ulkhaq, M. M. (2023). Clustering countries according to the logistics performance index. *JATISI (Jurnal Teknik Informatika Dan Sistem Informasi)*, 10(1). <https://doi.org/10.35957/jatisi.v10i1.4755>

Ulutaş, A., & Karaköy, Ç. (2019). An analysis of the logistics performance index of EU countries with an integrated MCDM model. *Economics and Business Review*, 5(4), 49-69. <https://doi.org/10.18559/ebr.2019.4.3>

Urbanowicz, R. J., Meeker, M., La Cava, W., Olson, R. S., & Moore, J. H. (2018). Relief-based feature selection: Introduction and review. *Journal of Biomedical Informatics*, 85, 189-203. <https://doi.org/10.1016/j.jbi.2018.07.014>

Van Steenberg, R. M., & Mes, M. R. (2020). Forecasting demand profiles of new products. *Decision support systems*, 139, 113401. <https://doi.org/10.1016/j.dss.2020.113401>.

Yildirim, B. F., & Adiguzel Mercangoz, B. (2020). Evaluating the logistics performance of OECD countries by using fuzzy AHP and ARAS-G. *Eurasian Economic Review*, 10(1), 24-45. <https://doi.org/10.1007/s40822-019-00131-3>

Yu, M. M., & Rakshit, I. (2023). An alternative assessment approach to global logistics performance evaluation: Common weight H-DEA approach. *International Transactions in Operational Research*. <https://doi.org/10.1111/itor.13360>