



# UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

## FACULTAD DE ESTUDIOS SUPERIORES ACATLÁN

“Clasificación de países potencia media aplicando análisis de conglomerados”

Tesina

QUE PARA OBTENER EL TÍTULO DE

Actuario

PRESENTA

Jardón Kojakhmetova Acel

Asesor: Mtro. Sánchez Aldana Alberto

Mayo 2011



Universidad Nacional  
Autónoma de México

Dirección General de Bibliotecas de la UNAM

**Biblioteca Central**



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

# Índice general

---

<b>Introducción</b>	<b>III</b>
<b>1. Potencias.</b>	<b>1</b>
1.1. Marco Histórico. . . . .	1
1.1.1. Edad Antigua. . . . .	2
1.1.2. Edad Media. . . . .	3
1.1.3. Edad Moderna. . . . .	4
1.1.4. Edad Contemporánea. . . . .	5
1.2. Noción de potencia. . . . .	6
1.2.1. Indicadores. . . . .	7
1.2.1.1. Indicadores geográficos y demográficos. . . . .	7
1.2.1.2. Indicadores económicos. . . . .	8
1.2.1.3. Indicadores de tipo militar. . . . .	8
1.2.1.4. Indicadores políticos y de recursos humanos. . . . .	11
1.3. Jerarquía entre Estados. . . . .	11
1.3.1. Superpotencias . . . . .	11
1.3.2. Gran potencia. . . . .	12
1.3.3. Potencia medias. . . . .	12
1.3.4. Potencias regionales. . . . .	13
1.3.5. Pequeña potencia. . . . .	14
<b>2. Conglomerados</b>	<b>15</b>
2.1. Antecedentes del análisis de conglomerados. . . . .	15
2.2. Uso de los conglomerados . . . . .	17
2.3. Etapas del análisis de conglomerados . . . . .	18
2.3.1. Elección de las variables. . . . .	18
2.3.2. Estandarización de variables . . . . .	19
2.3.3. Elección de la medida de similitud. . . . .	19
2.4. Técnicas de conglomeración. . . . .	22
2.4.1. Jerárquicos. . . . .	22
2.4.1.1. Aglomerativos . . . . .	23
2.4.1.2. Divisivos. . . . .	23
2.4.1.3. Principales algoritmos de clasificación jerárquica . . . . .	24
2.4.2. No jerárquicos. . . . .	37
2.4.2.1. Métodos de reasignación: . . . . .	37
2.4.2.2. Búsqueda de densidad. . . . .	39
2.4.2.3. Método Directo. . . . .	39
2.4.2.4. Método reductivo. . . . .	40
2.5. Validación del número de conglomerados. . . . .	40
<b>3. Modelo Aplicado.</b>	<b>41</b>
3.1. Elección de variables. . . . .	41
3.1.1. Indicadores . . . . .	41
3.2. Método jerárquico. . . . .	42
3.2.1. Interpretación de resultados jerárquicos . . . . .	45

3.2.2. Validación del método jerárquico. . . . .	50
3.3. Método de k-medias. . . . .	52
<b>Conclusiones</b>	<b>55</b>
<b>Bibliografía</b>	<b>56</b>
<b>A. Indicadores</b>	<b>59</b>
<b>B. Métodos jerárquicos</b>	<b>67</b>
<b>C. Análisis de K-medias</b>	<b>73</b>

# Introducción

---

El hombre recurre a la clasificación como medio para evitar la confusión e instintivamente clasifica el mundo que lo rodea mediante el reconocimiento de similitudes. Por ejemplo, un niño puede clasificar sus juguetes, un estudiante clasifica sus materias en fáciles y difíciles, un adulto puede clasificar opiniones, trabajos, objetos materiales, valores, según ciertas características o prioridades subjetivas.

En el campo científico, podemos encontrar los principios de la clasificación en los trabajos de Aristóteles hasta Carlos Lineo, con la teoría y la práctica de la clasificación de plantas y animales. El estudio de la clasificación correcta en biología es tan importante que se ha desarrollado una disciplina particular, la taxonomía o clasificación numérica. Las ventajas que tiene la clasificación numérica por encima de la clasificación subjetiva, es que permite el análisis de un mayor número de datos, así como la posibilidad de repetir el estudio después de definir ciertos criterios clasificatorios.

Del mismo modo en que se puede clasificar el reino animal en clases, una vez definidas éstas (mamíferos, aves, etc.), se establecen las subclases, las familias, las especies; los actores internacionales pueden jerarquizarse de acuerdo a diversos índices que nos permitan medir la similitud entre ellos.

El objetivo de esta tesina es proponer una metodología de clasificación de países, de acuerdo a índices tangibles, entre ellos: demográficos, económicos y militares, de tal manera que dicha clasificación tenga un sustento estadístico haciendo uso del análisis de conglomerados.

En el primer capítulo se presenta la forma en que se han jerarquizado los países Estado a través de la historia, desde la edad antigua (Imperio Persa), edad media (Imperio Mongol) y la edad moderna y contemporánea; los cambios que se han presentado en la jerarquía de las grandes potencias, principalmente en el Congreso de Viena en 1815, con el reestablecimiento de fronteras en Europa y después de la Segunda Guerra Mundial en 1945. Dentro de los primeros estudios de jerarquización de potencias realizados por Wight (1878), se demostró cómo la clasificación de las primeras potencias se inició con el reconocimiento de que los estados tienen diferentes clases y magnitudes, poniendo un interés particular en la extensión territorial y forma constitucional del estado [11]. Actualmente, se toman en cuenta diversos indicadores para medir y clasificar a las potencias en: superpotencias, grandes potencias, potencias medias, potencias regionales y pequeñas potencias.

Posteriormente, en el capítulo dos se toca el tema del análisis de conglomerados, que es un método estadístico empleado para la clasificación de individuos en grupos lo más distinto posible. La importancia del análisis de conglomerados, se ve reflejado en las diversas aplicaciones que se pueden encontrar de éste, por ejemplo, en biología con la clasificación de plantas y animales; en astronomía, con la clasificación de estrellas o en psicología, con la clasificación de enfermedades mentales.

En análisis de conglomerados existen dos grandes técnicas de clasificación: jerárquicas y no jerárquicas; así como diversos tipos de distancia, empleados según el tipo de variables a analizar. En este capítulo se da una explicación de cada una de ellas con sus respectivos ejemplo, cuando el método así lo permite.

Como capítulo final, se muestra el modelo aplicado para la clasificación de nuestros países, considerados en el amplio sentido de la palabra como potencias (tomando 24 países como objeto de estudio) utilizando índices económicos, demográficos, geográficos y militares. Se muestran los resultados obtenidos aplicando la conglomeración jerárquica con el método de Ward, y la conglomeración aplicando el método de las k- medias, se hacen los comparativos y se presentan las conclusiones. El desarrollo del modelo fue hecho con el paquete estadístico SPSS.

Cabe mencionar que este trabajo no pretende profundizar en el análisis de conglomerados, así como tampoco desarrollar a fondo el tema de países potencia, sino simplemente hacer una propuesta de clasificación que sirva como aportación actuarial en temas de estudio de relaciones internacionales, e introducir al actuario en un área en la cuál, casi no se ha visto involucrado, y así ampliar aún más el campo de aplicación de los conocimientos actuariales.

---

# Capítulo 1

# Potencias.

---

En este capítulo se verá un marco histórico de las potencias. Se definirán los conceptos de super potencia, potencia media y potencia regional, así como, los diversos criterios de clasificación a lo largo de la historia.

En este trabajo, haremos un mayor énfasis en el fenómeno de las potencias medias y potencias regionales.

## 1.1. Marco Histórico.

En un sentido amplio, siempre han existido potencias de primera magnitud relativa en las distintas sociedades internacionales que han ido sucediéndose hasta el advenimiento de la actual sociedad internacional mundial, subrayando el hecho de que las potencias han determinado en última instancia el destino de la sociedad internacional. La historia de la política internacional es fundamentalmente historia de las potencias. El término superpotencia aparece como un neologismo en 1922. Antes del final de la Segunda Guerra Mundial, Francia, Alemania, Italia, Japón, Unión Soviética, Estados Unidos, y Reino Unido poseían el estatus de superpotencia.

Al término de la Segunda Guerra Mundial, las potencias vencedoras - Francia, Reino Unido, la Unión Soviética y los Estados Unidos - se adjudicaron a sí mismos el estatus de superpotencias, adjudicándose además como miembros permanentes del Consejo de Seguridad de la ONU, poseyendo el derecho a veto en las decisiones que tomase el Consejo. Pero debido al declive económico producto de la guerra, los que habían sido grandes imperios como Francia e Inglaterra, y la Guerra Civil que sucedía en China, terminaron por quedar bajo influencia de dos superpotencias: Estados Unidos y la Unión Soviética.

En 1991 con el colapso político de la Unión Soviética, los Estados Unidos se convirtió en la única superpotencia. Sin embargo, este status es contendido por potencias como Rusia, cuando otros especulan que China e India pronto obtendrán estatus de superpotencia, y ser capaces de contender a EE. UU.

Aunque el término superpotencia es contemporáneo, a veces se aplica cuando se habla de las potencias o imperios del pasado, tales como Grecia antigua, el Imperio Romano, el Imperio Bizantino, el Imperio Otomano; y en términos de las edades modernas, el Imperio Español, el Imperio colonial francés y el Imperio Británico.

### 1.1.1. Edad Antigua.

#### Imperio Persa.

El antiguo pueblo ario, procedente de Asia Central se estableció en la meseta de Irán (actualmente Irán y Afganistán), entre el mar Caspio y el Golfo Pérsico. El imperio persa aqueménida fue unificado por Ciro II el Grande, fue el mayor imperio persa que se extendió por los territorios actuales de Irán, Iraq, Turkmenistán, Afganistán, Uzbekistán, Turquía, Chipre, Siria, Líbano, Egipto e Israel. Contaban con un poderoso ejército, destacándose la guardia real, formada por diez mil hombres que realizaban la defensa de su territorio y la conquista de otros que hicieron de él el imperio más grande en extensión hasta entonces. Su extensión concluyó en el año 330 a.c cuando fue derrotado por Alejandro Magno.



Figura 1.1: Imperio Persa



Figura 1.2: Ciro II

#### Imperio Heleno.

Una de las primeras civilizaciones que extendió su influencia y sus territorios lo suficiente para destacarse entre las demás fue la Macedonia de Alejandro Magno, que con su Imperio Heleno abarcó desde Grecia hasta la India ocupando una gran parte del mundo conocido. En su expansión destruyó el Imperio Persa (550 aC–330 aC), el más poderoso de su época.

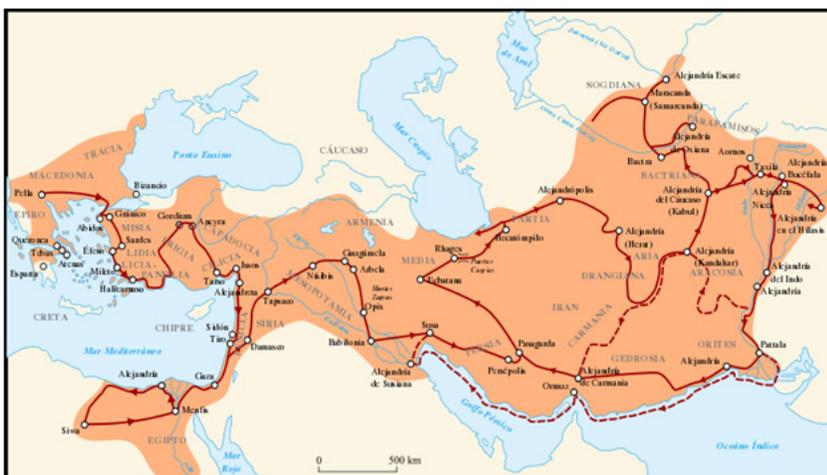


Figura 1.3: Imperio Heleno



Figura 1.4: Alejandro Magno

La civilización romana se convirtió en la primera potencia del mundo conocido, llegando a comerciar con la lejana China. Su tecnología militar y arquitectónica y su fuerte economía superaban con creces a los pueblos de su alrededor.



Figura 1.5: Imperio Romano

En su máxima extensión el Imperio Romano abarcaba desde Inglaterra y la Península Ibérica hasta el Golfo Pérsico, en lo que hoy se conoce como Kuwait. Después de sufrir los ataques durante siglos de los pueblos germánicos, el último emperador romano de occidente fue depuesto en el año 476 d.C. Esto marcó el final de la Edad Antigua.

### 1.1.2. Edad Media.

#### Califato Abasí.

Abbas, descendiente del tío paterno de Mahoma, fundó la dinastía abasí que residía en Humayma, una pequeña aldea de Palestina. El Califato Abasí fue una de las potencias más destacadas de la Edad Media; su cultura y sus conocimientos fueron extendidos por todo su Imperio que abarcó desde la Península Ibérica hasta Indonesia. En el año de 1517, el califato se derrumbó frente a la conquista de los turcos otomanos.



Figura 1.6: Califato Abasí

### Imperio Mongol.

Fue el segundo Imperio más grande de toda la historia. Con Gengis Kan al mando en 1206, este imperio llegó a tener una extensión de unos 33 millones de  $km^2$  y más de 100 millones de habitantes. El imperio acabó decayendo por hambruna mucho antes del principio de la Edad Moderna.

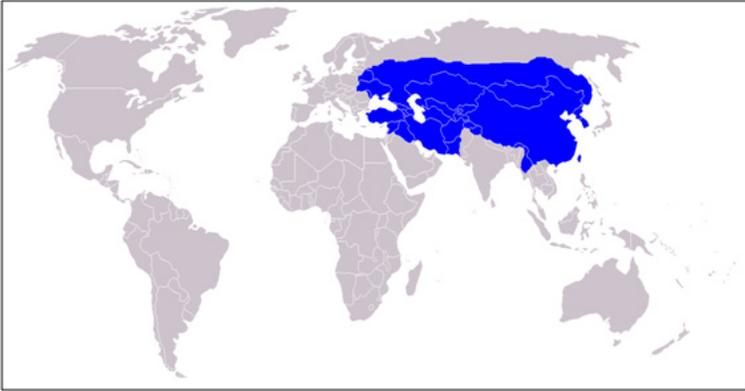


Figura 1.7: Imperio Mongol



Figura 1.8: Gengis Kan

### 1.1.3. Edad Moderna.

A principios de la Edad Moderna los portugueses y los españoles habían iniciado una serie de expediciones para crear rutas económicas hacia la India. Una de estas expediciones fue la de Cristóbal Colón, que al servicio de Castilla descubrió en 1492 América. Este descubrimiento inició el ascenso de Castilla en la escena internacional. Después de la unión dinástica de las Coronas de Castilla y Aragón por parte de los Reyes Católicos, los territorios hispanos se empezaron a expandir vertiginosamente. Durante el siglo XVI y hasta el Tratado de los Pirineos en 1659, el Imperio Español, ya indiscutible primera potencia, conquistó una gran parte de América, las Filipinas y se anexó el Imperio Portugués, además con las uniones dinásticas, consiguió formar un imperio europeo. También la creación de los tercios y una poderosa armada lo llevaron a ser invencible durante casi 150 años. Tras una serie de bancarrotas, malos gobernantes y derrotas militares acabaron con su hegemonía europea, aunque mantuvo un papel importante en la escena mundial gracias a su vasto imperio y su fuerte economía hasta el siglo XIX.

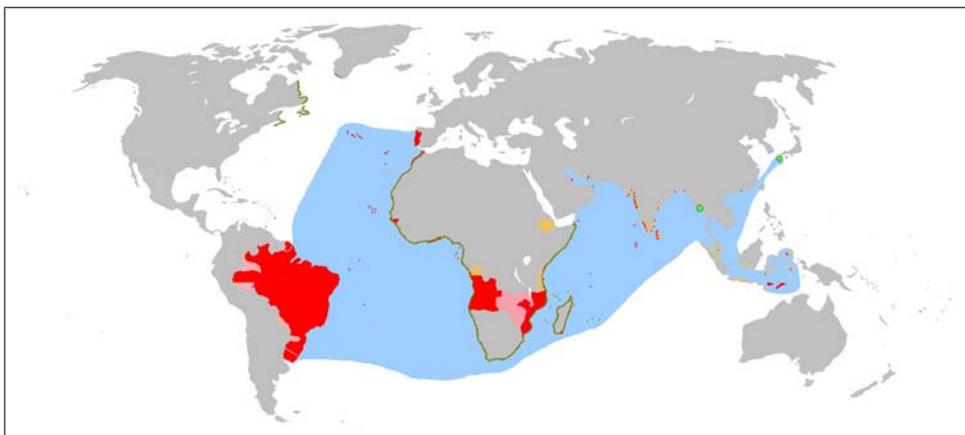


Figura 1.9: Imperio Portugués

Después de la caída oficial española de su título en 1659, Francia y Gran Bretaña (aunque no tan adelante como Francia) se disputaron el dominio mundial por más de un siglo. Gran Bretaña gracias a su marina y la victoria en la Guerra de los 7 años, se logró imponer frente a Francia y las demás potencias; así se empezaba a consolidar más fuertemente el Imperio Británico. También contribuyó a esto la revolución industrial iniciada en el siglo XVIII en Inglaterra que le aseguraría su título; aunque tal y como lo había hecho Gran Bretaña durante

la dominación francesa, estos últimos le seguirían al imperio británico con relativa cercanía hasta principios del siglo XX.



Figura 1.10: Imperio Británico a finales del siglo XIX

#### 1.1.4. Edad Contemporánea.

Con la decadencia del Imperio Británico tras la segunda guerra mundial, quedaron 2 países debatiéndose el puesto de primera potencia mundial: Estados Unidos y la U.R.S.S. Lo que caracterizó y distinguió a estas dos naciones del resto del mundo, fueron las grandes cantidades de arsenal nuclear acumulado al finalizar la segunda guerra mundial. De esas dos naciones, los Estados Unidos era «el primero entre dos iguales» toda vez que en 1950 ejecutaba el 50% del gasto militar mundial, poseía la mitad de las reservas financieras y las dos terceras partes de la producción industrial del mundo. Estos hechos le daban una hegemonía total en la esfera económica del mundo.[18]

Las Naciones Unidas cambiaron bruscamente a partir de 1945, cuando los Estados Unidos obtuvieron el arma nuclear. Fue a partir de la carrera de armamento nuclear entre la Unión Soviética y los Estados Unidos cuando la posesión de amplios arsenales estratégicos otorgó a estos Estados la condición de «superpotencias» y abrió una profunda brecha entre éstas y el resto de potencias con responsabilidades mundiales. Sin embargo, la «estrategia de la disuasión», que dominó las relaciones Este-Oeste durante cuarenta y cinco años (1945-1990), estaba basada en la asunción de que era la posesión, y no el uso del arsenal estratégico, lo que resultaba suficiente para detener el ataque. Por tanto, no hay duda de que el poderío militar (también denominado «poderío beligerante»)<sup>1</sup> de un Estado parece ser la prueba final de su status en la jerarquía de los Estados. Por otra parte, el poderío militar de un Estado es básicamente neutral. El dilema radica en que los propósitos nunca son neutrales; es la actividad interna que desea traslucirse en actividad externa. Si un Estado se propone algo y tiene el poder para hacerlo, intentará conseguirlo, de grado o por la fuerza, por amenaza o recurriendo a la violencia.

De todas formas, en los años setenta se pretendió acuñar la noción de «potencia civil», desarrollada especialmente en la década siguiente, cuando dentro del sistema bipolar de las relaciones internacionales se produjeron una serie de «movimientos anti líder», particularmente en el bloque occidental. Con este término se hacía referencia a aquellos Estados con grandes recursos materiales, industriales, financieros y sociales, pero carentes de una fuerza militar proporcionada a su tamaño. De esta manera se quería desvincular la noción de potencia de la capacidad militar <sup>2</sup>. Para los autores que han analizado la cuestión, la noción de «potencia civil» o «Estado comercial», como también se ha denominado implica:

1. El reconocimiento de la necesidad de una cooperación amplia con otros actores del sistema internacional;
2. La concentración en los asuntos no militares, fundamentalmente los económicos, para establecer una fuerza militar de carácter residual para salvaguardar los intereses esenciales de esa seguridad;

<sup>1</sup>El poder beligerante de un Estado se define como la suma de su poderío marítimo, terrestre y aéreo empleado colectivamente y en cooperación, más su poder civil.

<sup>2</sup>Una gran potencia económica puede optar por permanecer como un «enano» militar, mientras que un Estado con pocos recursos económico-productivos puede organizar una potencia militar

3. La capacidad de cooperación avanzada en las estructuras internacionales donde es posible resolver las cuestiones críticas del sistema internacional; y
4. Asumir la premisa de que la guerra no compensa los sacrificios realizados, porque los intercambios comerciales constituyen el mejor medio para aumentar la riqueza y el poderío de las Naciones.

Finalmente, después de más de 40 años de revoluciones y guerras civiles a lo largo del mundo, en 1991, tras la desintegración de la U.R.S.S. Estados Unidos se convierte oficialmente en la primera potencia mundial hasta la fecha.

## 1.2. Noción de potencia.

El uso del término potencia se desarrolló ampliamente en el concierto diplomático del siglo XIX para referirse a todas las Naciones que participaban en el sistema europeo de Estados <sup>3</sup>.

Actualmente el término potencia hace alusión a un Estado que con mayores recursos tiene más influencia fuera de sus fronteras, mayor seguridad frente a las presiones, las amenazas y el ataque militar y, en definitiva, más prestigio y más amplio campo de elección en la elaboración y ejecución de su política exterior. En cambio, un Estado pequeño es más vulnerable, menos resistente a las presiones del exterior, más limitado en sus opciones políticas y mantiene una vinculación más estrecha entre su política interior y exterior. Así, podemos decir que existen al menos tres tipos de Estados que deben tomarse en consideración en el análisis del concepto de potencia.

El primero de estos tipos es el Estado cuyo interés principal reside en sí mismo y en los países vecinos y que, en palabras de Merle [15], ejerce una «función local». La fuerza que estas Naciones puedan desarrollar se ve limitada a una esfera geográfica muy localizada; ni que decir tiene que la mayoría de los Estados del sistema internacional pertenecen a esta primera categoría.

El segundo tipo lo constituyen aquellos Estados con influencia en un sector particular de las relaciones internacionales o con responsabilidades regionales, que han sido tradicionalmente denominados «potencias medias» o «potencias regionales». Estos Estados no son potencias mundiales porque carecen de los recursos necesarios para ostentar tal rango, pero sus responsabilidades no están limitadas a los países vecinos de su entorno.

El tercer tipo de Estados lo forman aquellos cuyos recursos, intereses y capacidades militares son tales que pueden hacer sentir su influencia en todos los asuntos mundiales y alcanzar sus objetivos de una forma más plena que los otros miembros de la sociedad internacional; se trata de las «grandes potencias». Así pues, cuando nos referimos a las «potencias» hablamos de Estados, pero de Estados con ciertas particularidades de tipo material y subjetivo que hacen que sobresalgan por encima de sus iguales.

Algunos autores responden a la pregunta de qué es una potencia atendiendo a sus recursos y a la capacidad de movilización de los mismos. Otros tienen en consideración el efecto que causaría en el sistema la aparición de una potencia, sea mundial o regional, grande o media. Por tanto, podemos decir que potencias son aquellos Estados que establecen las reglas del sistema, ya sea internacional, regional o local, y disponen de los recursos y de capacidades necesarias para movilizarlos de forma apropiada en defensa de tal estructura. Así, los actores que configuran la estructura del sistema son indudablemente las potencias, pues, como ha dicho Barbé [2], «sólo ellas disponen del poder estructural que permite dictar las reglas de juego a nivel global en el sistema internacional»<sup>4</sup> También en los subsistemas regionales y locales las potencias medias, en interacción constante para alterar las reglas del sistema general a su favor. De modo que existe una conexión entre las reglas generales y las reglas regionales. Éstas, por ejemplo la Unión Europea, forman parte de la estructura mundial, en conexión con las reglas generales. Se trata de una dialéctica constante de fuerzas atractivas y repulsivas por mantener la estructura equilibrada y, en consecuencia, las reglas dictadas como generalidad por quien tiene el poder para

<sup>3</sup>El concierto europeo de potencias era un sistema diplomático en el cual concurrían las principales potencias europeas para la resolución de sus controversias. No era un sistema rígido ni basado en pactos que pretendieran colocar a todos los Estados en pie de igualdad. Se reunían cuando lo requerían las circunstancias, y los Estados que lo formaban conservaban intacta su soberanía. Fracasó al no evitar la guerra entre los Estados europeos; sin embargo, evitó la formación de bloques permanentemente hostiles entre grupos de Estados hasta finales del siglo XIX.

<sup>4</sup>La estructura del sistema internacional ha sido definida por Barbé como «la configuración de poder surgida de las relaciones entre actores.»

hacerlo se ven constantemente repelidas por los que se encuentran en disposición de hacerlo; el resultado es el equilibrio sistémico.

### 1.2.1. Indicadores.

Nos interesa ver con qué capacidades efectivas de acción internacional cuenta un Estado, es decir, cuáles son los indicadores que empleamos para evaluar el poder de los Estados y eventualmente establecer una relación jerárquica entre ellos. Los objetivos del Estado tienen que ver, prioritariamente, con la conservación de los elementos primarios que lo caracterizan: territorio (seguridad física), población <sup>5</sup> (seguridad económica) y gobierno (seguridad política).

Esta diversidad de objetivos, por consiguiente, impide la existencia de un único indicador mediante el cual se pueda identificar el poder de un Estado. Para mayor precisión, conviene recurrir a una serie de indicadores distintos, que podemos reunir en los grupos siguientes:

#### 1.2.1.1. Indicadores geográficos y demográficos.

La geografía puede conferir a un Estado ciertas ventajas o desventajas que pueden afectar el comportamiento de la política exterior. Entre las posibles ventajas que puede gozar un país están el control de vías acuáticas estratégicas o la lejanía entre países o potencias en conflicto, por ejemplo, la gran distancia que siempre han tenido los Estados Unidos y Australia respecto de sus enemigos eternos, la mayor parte de su historia [18].

El aspecto demográfico es fundamental para permitir una necesaria mano de obra y demanda interior que garantice ritmos adecuados de producción, acordes con las exigencias tecnológicas y de costes.

Territorio, número de habitantes, distribución de la población, índice de crecimiento demográfico, disponibilidad de recursos naturales, son indicadores tomados en cuenta al momento de jerarquizar un Estado, ya que como Holbraad [11] menciona: “ningún país puede llegar a ser una potencia de primer orden si no se encuentra entre las más pobladas del mundo”. Cabe hacer notar que el volumen de la población determina ciertos límites a la acción exterior de los Estados y fija el potencial económico de los mismos.

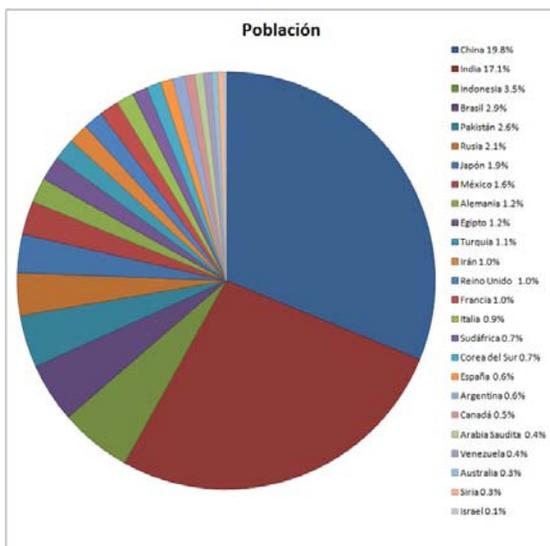


Figura 1.11: Población por país (%mundial)

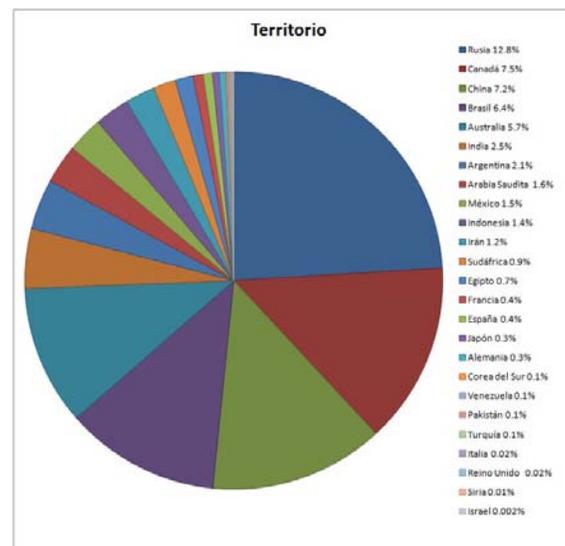


Figura 1.12: Área por país (%mundial)

<sup>5</sup>Holbraad: «ningún país puede llegar a ser una potencia de primer orden si no se encuentra entre las más pobladas del mundo».

### 1.2.1.2. Indicadores económicos.

Los indicadores económicos suelen estar ligados a los indicadores demográficos. Una población adiestrada y avanzada tecnológicamente puede permitir a un país alcanzar altos estándares de vida, gozar ventajas comerciales y proveer asistencia o dominar a otros Estados [18]. Los indicadores económicos han de ser tomados en cuenta al hacer una jerarquización de Estados, debido a que la economía puede ser usada como un “arma” en la política nacional, es decir, su uso para adquirir poder e influencia en el extranjero.

Producto Interno Bruto (PIB) (total y per cápita), distribución de la riqueza, niveles de industrialización y productividad, producción y consumo de energéticos, recursos técnicos (nivel de tecnología); intercambios comerciales con el exterior.

### 1.2.1.3. Indicadores de tipo militar.

El poder militar de un Estado ha sido siempre el primer atributo a tener en cuenta para su valoración. Como afirma Aron [1] «la fuerza militar sigue siendo el fundamento del orden internacional»; para Waltz [22] «un país sin ejército no puede ser una gran potencia»; Barbé [2] dice que «el tamaño militar del Estado constituye un criterio clásico a la hora de establecer jerarquías en el sistema internacional»; Keohane [14] que «la fuerza militar es siempre un componente central del poder nacional»; Morgenthau [16]: «en política internacional, la fuerza armada como amenaza o como amenaza potencial, es el más importante factor que forma el poder político de una nación». Número de hombres movilizados, armamento del que se dispone, calidad del liderazgo y de la organización militar, nivel de gastos militares, industria y comercio de armamento, armas nucleares.

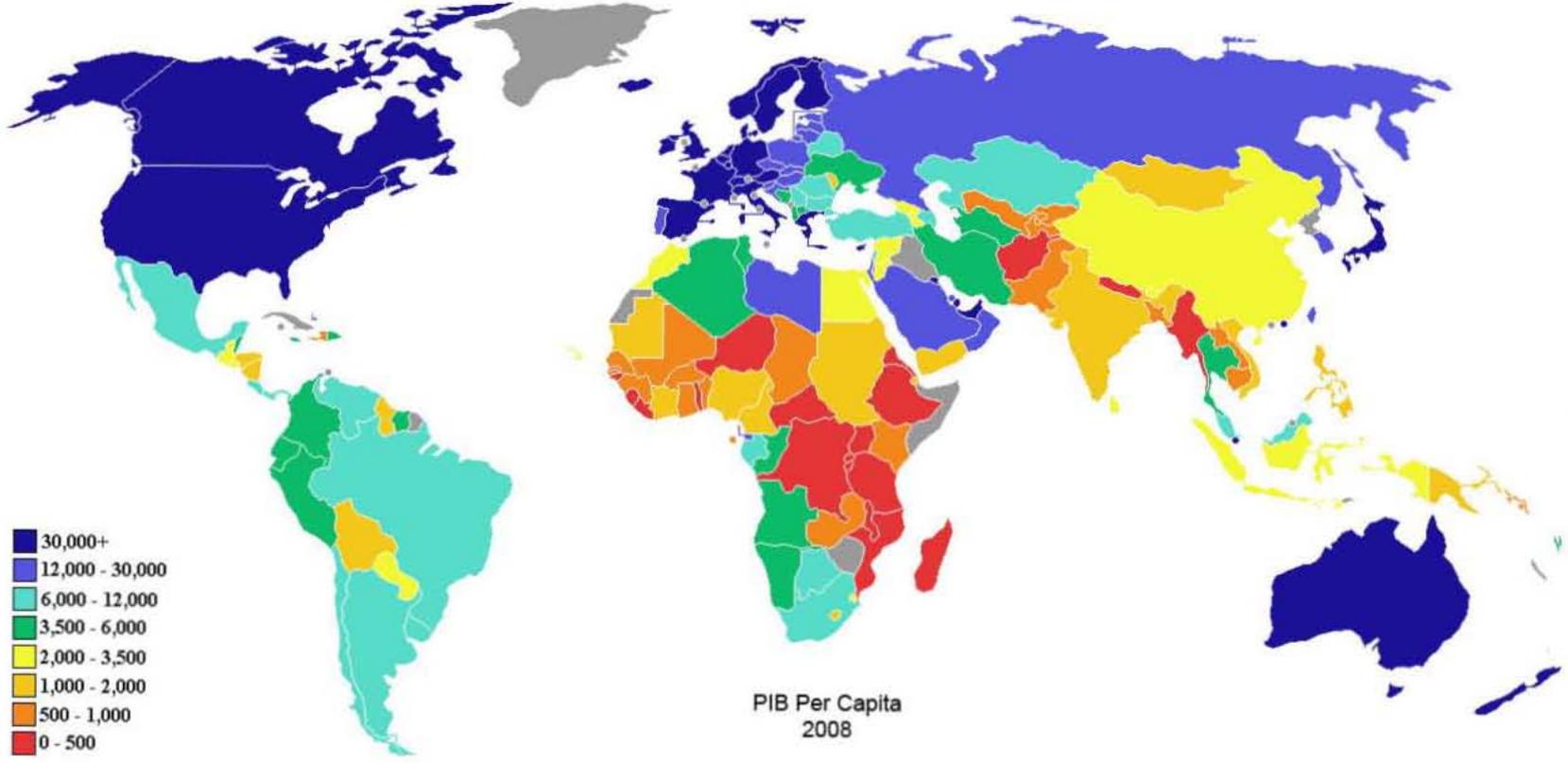


Figura 1.13: PIB

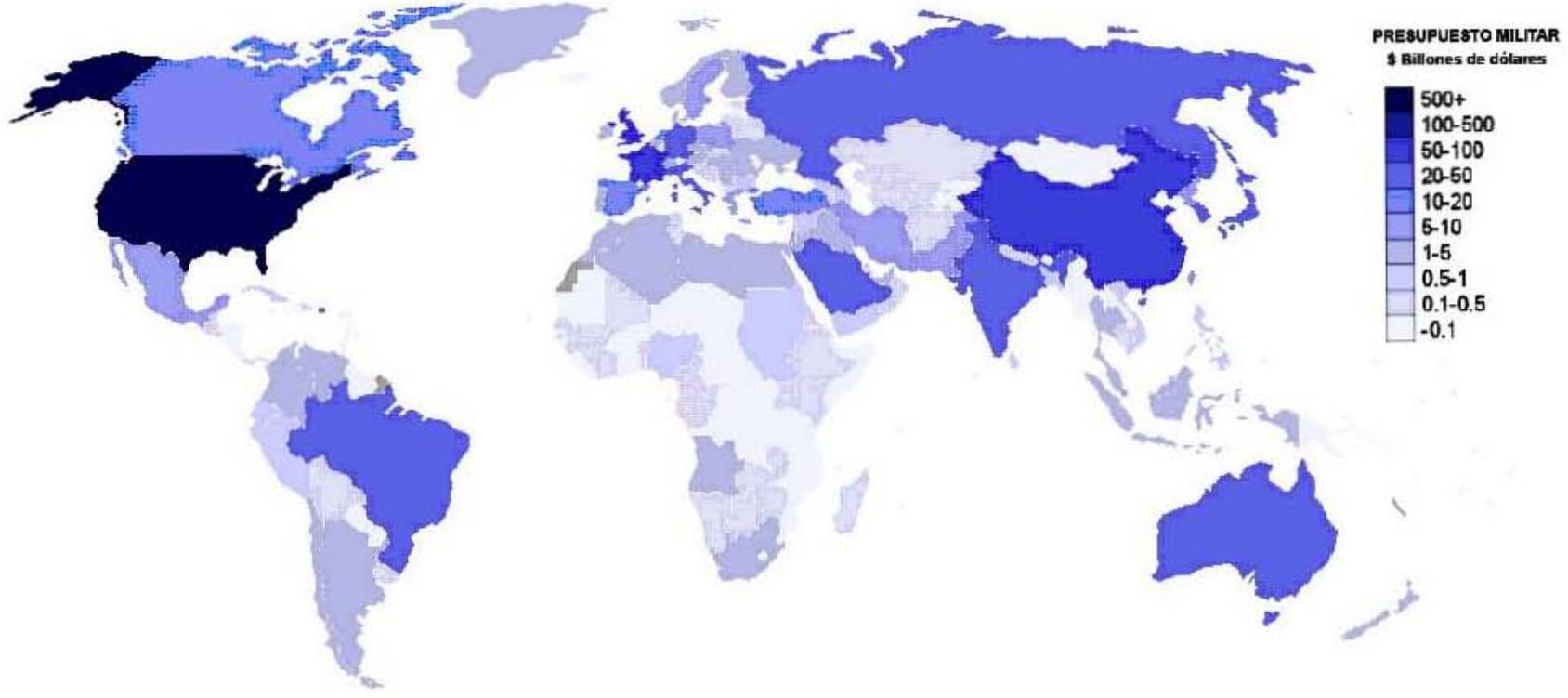


Figura 1.14: Presupuesto Militar (billones de dólares)



Hasta hace poco tiempo, entraban en esta categoría EEUU y la Unión Soviética. Al haberse desintegrado esta última, resulta incierto si podríamos seguir considerando a Rusia como superpotencia en base a su arsenal nuclear heredado de la URSS; por el momento, parece poco probable. Por otra parte, y desde un punto de vista meramente estratégico, la propia existencia de esta categoría podría quedar cuestionada al quedar como integrante un solo país (EEUU): ¿de qué sirve poder responder a un ataque nuclear masivo si ninguna otra potencia puede hacerlo? [12]

Pero para comprender adecuadamente qué es lo que califica a una superpotencia de una gran potencia, y por qué surgen en el período posterior a la Segunda Guerra Mundial y no antes, conviene proceder a una definición del término. Por superpotencia entendemos: aquel Estado con capacidad y voluntad de ejercer una hegemonía, absoluta en el marco de una sociedad mundial mediante la disponibilidad de un poderío militar de naturaleza nuclear susceptible de provocar una guerra de destrucción masiva y simultánea, capaz de poner en peligro la existencia de toda la Humanidad, así como de una potencialidad económica y técnica que permita mantener e incrementar dicho poderío militar nuclear al objeto de garantizar su capacidad disuasoria.

Las superpotencias disponen de una potencialidad militar muy calificado, a saber: un armamento nuclear estratégico cuya utilización provocaría el desencadenamiento de un conflicto bélico, cuya dimensión mundial derivaría no de la participación mayoritaria de los estados como beligerantes, sino de los efectos que toda la Humanidad tendría que soportar, con independencia de su grado de participación en el conflicto, y que ponen en peligro la existencia misma de esa Humanidad. A diferencia de otros períodos históricos en los que las grandes potencias tenían la capacidad de generar conflictos bélicos en todo el ámbito de la sociedad internacional, incluido el ámbito mundial en el siglo actual, dicha capacidad era una consecuencia de la dimensión intercontinental de tales grandes potencias y/o de su capacidad para incorporar a terceros países como partes beligerantes. En la actualidad, las superpotencias por sí solas pueden iniciar y concluir un conflicto, sin la necesidad de intervención de otros estados, cuyos efectos repercutirían definitivamente en la subsistencia del resto de la humanidad.

### 1.3.2. Gran potencia.

Se consideran dentro de esta categoría a los países con gran capacidad de influencia o con intereses de tipo económico, político y/o militar que van más allá de la propia región en la que se encuentran, aunque sus políticas exteriores no lleguen a tener propiamente dimensiones mundiales, como sería el caso de las superpotencias.

Las grandes potencias mundiales son responsables del equilibrio de poder en un determinado sistema (Sistema Internacional). De hecho, este fue el espíritu que animó al Congreso de Viena a establecer formalmente el estatuto de gran potencia, en beneficio de Rusia, Prusia, Gran Bretaña, Francia y Australia.

### 1.3.3. Potencia medias.

Un primer significado se refiere a un atributo tal como la ubicación geográfica estratégica como criterio para clasificar a los países que se consideran de este modo. Este concepto se utilizó durante los siglos XVIII y XIX para referirse a aquellos países que, aunque no contaban con una posición política internacional sobresaliente, disfrutaban de un mayor poder relativo derivado de su posición geográfica.

El término potencia media como tal, surgió después de la Segunda Guerra Mundial y fue aplicado principalmente a Canadá, Australia y Nueva Zelanda, quienes ocuparon una posición intermedia o neutral en la política internacional.

Son diversas las definiciones que encontramos de este concepto, por lo tanto, mencionaremos las que a nuestro parecer son importantes y contienen características similares:

Como señala Guadalupe González [9] “a lo largo de la evolución histórica del sistema internacional, el concepto de potencia media ha sido utilizada en doble sentido para referirse, por una parte, a la posesión de una fuerza económica, política y militar mediana en la escala de poder mundial, y por otra, a la ocupación de una posición intermedia centrista y/o moderadora entre dos países en conflicto”.

Santoro [20] por su parte, define a las potencias medias como “aquél grupo de Estados que ejercen funciones determinadas, debido a su dimensión (territorial, demográfica o económica), a su colocación geopolítica en un área regional o a su capacidad de desestabilizar el sistema global si renunciaran a formar parte de un bloque y pasaran a otro lado”.

Finalmente tenemos la definición de Holbraad [11]<sup>6</sup>

Holbraad [11] presenta una lista de 18 países considerados en base a su territorio, población y PIB; con datos de 1975 su lista está formada por: Alemania, Argentina, Australia, Brasil, Canadá, China, España, Francia, India, Indonesia, Irán, Italia, Japón, México, Nigeria, Polonia, Sudáfrica y el Reino Unido.

Contrastando con esta lista, encontramos la propuesta por Wood [12] quien incluye en su relación a los 30 países que se clasifican inmediatamente tras las grandes potencias según un único clasificador: el PIB. Con datos de 1979, su lista es: las “grandes potencias” serían EEUU, URSS, Japón, República Federal Alemana, Francia y Reino Unido. Las potencias medias serían los siguientes 30 países: Arabia Saudí, Argentina, Australia, Austria, Bélgica, Brasil, Canadá, Corea del Sur, Checoslovaquia, China, Dinamarca, España, Finlandia, Hungría, India, Indonesia, Italia, México, Nigeria, Noruega, Países Bajos, Polonia, RDA, Rumania, Suecia, Suiza, Turquía, Venezuela y Yugoslavia; más Argelia, Irán y Pakistán, añadidos por razones de importancia y equilibrio regional, siendo en total 33 países.

En resumen podemos considerar a las potencias medias como aquellos países de tamaño grande o media, con una política exterior activa en determinadas áreas, regional o sectorial y a los que se les supone una cierta capacidad de influencia, además de un prestigio internacional, e intentan maximizar su influencia sobre países pequeños, minimizar la influencia de las potencias grandes y evitar el surgimiento de otras potencias medias en la región [17]

Dentro de esta categoría podemos encontrar algunos países que tienen un papel importante a nivel internacional pero quizá no sean los más influyentes en la región en la que están ubicados, por ejemplo, el caso de Sudáfrica, que en términos cualitativos sería una potencia media, pero no necesariamente la potencia regional, puesto que su liderazgo es cuestionado por Nigeria, el país más importante en cuanto a población, recursos (petróleo) y potencia económico. Es por esto que vemos importante introducir el concepto de potencia regional.

#### 1.3.4. Potencias regionales.

El concepto de potencia regional ha sido menos desarrollado que el de potencia media, pero podemos entender como potencias regionales a aquellos Estados que carecen de capacidad y/o voluntad de ejercer su hegemonía a nivel mundial, pero intervienen, económica, política y militarmente de modo eficaz en un área geopolítica más restringida.

Estos Estados tienen ciertas características en común:

1. Forman parte de una región definida.
2. Aspiran a un liderazgo regional.
3. Tienen influencia política y geoestratégica en la construcción regional.
4. Disponen de suficientes recursos materiales, organizacionales e ideológicos para crear un proyecto regional.
5. Tienen una notable interdependencia política, económica y cultural regional.
6. Disponen de una destacada influencia en asuntos regionales.
7. Son reconocidos como líderes regionales por los demás países del entorno.
8. Son portavoces en algunos asuntos regionales en foros internacionales.

---

<sup>6</sup>Holbraad: “Las potencias medias son aquellos, debido a sus dimensiones, sus recursos materiales, su voluntad y capacidad de aceptar responsabilidades, su influencia y su estabilidad están en vías de convertirse en grandes potencias”.

Debido a las diferencias que se encuentran en la literatura de relaciones internacionales respecto a la jerarquización de Estados, se puede ver mezclado en algunos casos los conceptos de potencia media y potencia regional con otra definición de potencias emergentes.

### **1.3.5. Pequeña potencia.**

Esta categoría de países podrían considerarse como Estados periféricos emergentes, siendo un grupo numeroso y heterogéneo. En general se caracterizan porque su influencia se limita a su propio territorio exclusivamente, son más vulnerables a la presión internacional, pero cuentan con aspiraciones geopolíticas y geoeconómicas de carácter regional.

# Conglomerados

---

El análisis de conglomerados, conocido también como taxonomía numérica, es un método estadístico multivariante de clasificación, que a partir de un conjunto de datos trata de situarlos en grupos (conglomerados) homogéneos no conocidos de antemano, pero sugeridos por la propia esencia de los datos, de tal manera que los individuos que puedan ser considerados similares sean asignados a un mismo conglomerado, mientras que los individuos diferentes se localicen en conglomerados distintos.

En resumidas palabras, lo que se intenta realizar con este método es una clasificación de individuos en grupos la más distintos posibles.

## 2.1. Antecedentes del análisis de conglomerados.

El proceso de clasificar, el reconocimiento de similitudes y la agrupación de objetos u organismos de acuerdo a estas similitudes comienza con el hombre primitivo, el hombre recurre a la clasificación como medio para evitar la confusión e instintivamente clasifica el mundo que lo rodea.

El origen de la ciencia de clasificación se remonta a los griegos, particularmente a Aristóteles con su sistema *Scala naturae* [8] o escala de la naturaleza. En ésta colocó a todos los organismos en una jerarquía que iba de los más simples a los más complejos de acuerdo a su constitución. De esta manera Aristóteles junto a sus discípulos clasificó algo más de 1000 especies animales y vegetales conocidos en hasta ese momento.

La principal clasificación hecha por Aristóteles fue la clasificación de los animales en dos grandes grupos: *anaíma* o sin sangre, en donde incluyó a los invertebrados, y *enáima*, vertebrados.

Esquema Aristotélico de clasificación de los animales	
<b>Animales sanguíneos (enaima)</b>	<b>Animales no sanguíneos (anima)</b>
<b>Vivíparos</b>	<b>Ovíparos con huevos imperfectos</b>
1. Hombre.	7. Malacodermos (cefalópodos).
2. Cuadrúpedos peludos (mamíferos terrestres).	8. Malacostráceos (crustáceos).
3. Cetáceos (mamíferos marinos).	<b>Vermíparos</b>
<b>Ovíparos</b>	9. Insectos.
4. Aves.	<b>Producidos por limo fértil o generación espontánea.</b>
5. Cuadrúpedos escamosos y ápodos (reptiles y anfibios).	10. Ostracodermos (moluscos, salvo excepciones).
<b>Ovíparos con huevos imperfectos.</b>	11. Zoofitos.
6. Peces.	

Fuente: C. García Dual 1992.

Cuadro 2.1: Esquema Aristotélico de clasificación de los animales.

Aristóteles también contribuyó con la primera clasificación de Economía y Géneros Literarios, llamando “Economía” al estudio de la administración de la “hacienda doméstica”, en su opinión, el Estado debe administrar sus propiedades con la prudencia de un jefe de familia, por lo cual dedujo que hay tantas economías o métodos administrativos especiales como formas de gobierno. La clasificación de Aristóteles es la siguiente:

- Economía regia (reino o monarquía).
- Economía provincial (provincias).
- Economía política (Ciudad-Estado).

En literatura la primera clasificación de géneros literarios según Aristóteles fue la siguiente:

- Épica.
- Lírica.
- Dramática.

Siguiendo las técnicas de Aristóteles, Teofrasto escribió el primer manual fundamental de estructura y clasificación de las plantas basado en sus propiedades médicas.

Posteriormente en el siglo XVII el sueco Carlos Linneo, considerado el Padre de la Taxonomía, publicó el libro *Species Plantarum*, en donde mostró una agrupación de las plantas de acuerdo a similitudes morfológicas entre ellas.

## 2.2. Uso de los conglomerados

La clasificación de objetos siempre ha sido de especial interés en distintas áreas del conocimiento, tales como:

**Biología.** Se ha usado principalmente para hacer clasificaciones de organismos y plantas. Esta clasificación se conoce como taxonomía y es el área en donde más ejemplos aplicativos se pueden encontrar. Un ejemplo claro sería la clasificación del hombre:

ESPECIE	→	SAPIENS
GÉNERO	→	HOMO
FAMILIA	→	HOMÍNIDOS
ORDEN	→	PRIMATES
CLASE	→	MAMÍFEROS
TIPO	→	VERTEBRADOS
REINO	→	ANIMAL

Cuadro 2.2: Clasificación del hombre.

**Astronomía.** Podemos encontrar la clasificación de las estrellas en grandes y pequeñas, usando el diagrama de Hertzsprung-Russell que muestra la relación existente entre la magnitud de una estrella y su temperatura superficial, se puede seguir evolutivamente la vida de las estrellas agrupándolas según unos parámetros concretos.

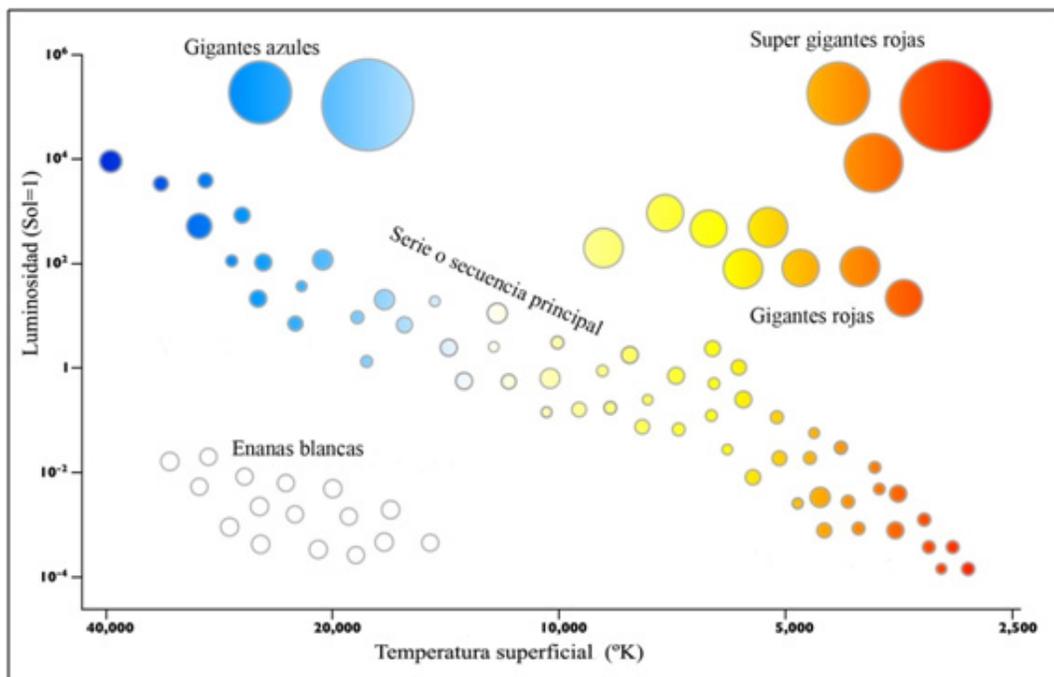


Figura 2.1: Diagrama de Hertzsprung-Russell.

**Psicología y psiquiatría.** Podemos encontrar diversas aplicaciones en la clasificación de trastornos de personalidad y el comportamiento de pacientes esquizofrénicos. Por ejemplo el Cuestionario para el registro de Alteraciones Físicas de origen Psíquico (BEB) [13] de E.Kasielke y K-D Hénsen que determina los trastornos de la vivencia del sujeto de los procesos emocionales, cognitivos y motivacionales que afectan su comportamiento en la vida.

**Medicina.** Se puede clasificar una muestra de aspirantes a la residencia en conglomerados de acuerdo a la posición relativa a sus puntajes estandarizados por encima o por debajo de la media de todo el conjunto [5].

**Química.** La clasificación de los elementos de la tabla periódica desarrollada por Mendeleev. Esta tabla fue la primera representación coherente de las semejanzas de los elementos. Mendeleev se dio cuenta de que clasificando los elementos según sus masas atómicas se veía aparecer una periodicidad en lo que concierne a ciertas propiedades de los elementos.

**Arqueología.** Se ha usado para investigar la estructura social en función de los ajueres que acompañan a los muertos. Los ajueres catalogados en las tumbas representan los conjuntos a agrupar. Luego se pueden examinar los distintos conglomerados para ver qué artefactos desempeñan un papel predominante para llevar a cabo el agrupamiento.

**Investigación de mercados.** Se ha usado para hacer una clasificación de los clientes con diferentes patrones de consumo. Por ejemplo en la venta de automóviles se puede hacer un análisis que agrupe los automóviles de mayor venta de acuerdo a los precios, características físicas y gustos del cliente.

**Lingüística.** Se hace una clasificación de las lenguas teniendo en cuenta sus similitudes gramaticales con el objetivo de reconstruir un árbol de evolución de las lenguas.

## 2.3. Etapas del análisis de conglomerados

1. Elección de las variables.
2. Elección de la medida de similitud.
3. Elección de la técnica de conglomeración.
4. Validación de los resultados.

### 2.3.1. Elección de las variables.

Se debe poner atención en la selección de variables ya que la clasificación efectuada por el análisis de conglomerados depende completamente de las variables seleccionadas, por lo tanto se debe evitar seleccionar aquellas que sean irrelevantes para el análisis ya que éstas podrían dar resultados erróneos, en general la elección de las variables a medir se hace de manera empírica, ya que no hay una base teórica general que ayude a determinarlas. A diferencia del análisis de discriminantes, aquí no hay variables independientes, es por esto, que el análisis tiene un fin meramente exploratorio.

Al seleccionar las variables podemos encontrarnos con algunos problemas:

- Que están medidas en diferentes escalas.
- Que están correlacionadas.
- Que el número de variables sea muy grande.

Para solucionar el problema de la diferencia de escalar, se puede recurrir a la estandarización de las variables de tal manera que todas las variables tengan media 0 y desviación típica 1 y además que no existan unidades entre los valores. La solución a los otros dos problemas se puede resolver con un análisis factorial previo a la realización del análisis de conglomerados, o utilizar la distancia de Mahalanobis (esta medida se verá más adelante).

Dentro del análisis pueden presentarse diversas variables:

**Variables cualitativas.** Su valor es un atributo o categoría:

- a) Binarias. Sólo pueden tomar dos valores posibles.
- b) Generales. Pueden tomar muchos valores.

**Variables cuantitativas.** Su valor se expresa numéricamente:

- a) Discretas. Sólo toman valores enteros.
- b) Continuas. Toman cualquier valor real en un intervalo.

### 2.3.2. Estandarización de variables

La escala a la que están medidas las variables puede afectar el resultado. Si las variables tienen escalas de medida muy distintas, es conveniente estandarizarlas previamente:

$$z_{ic} = \frac{x_{ic} - m_c}{s_c} \quad (2.1)$$

donde  $m_c$  y  $s_c$  son medidas muestrales de centro y dispersión respectivamente, ambas medidas en la misma escala que  $x_{ic}$ .

Las más conocidas son la media y la desviación estándar muestrales:

$$m_c = \frac{1}{n} \sum_{i=1}^n x_{ic} \quad (2.2)$$

y

$$s_c = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_{ic} - m_c)^2} \quad (2.3)$$

### 2.3.3. Elección de la medida de similitud.

Existen diversos tipos de medidas que pueden ser utilizados en el análisis de conglomerados, pero antes tenemos que introducir el concepto de distancia:

**Definición 1** Distancia: Una distancia  $d$  definida en un conjunto  $E$  es una aplicación entre el punto cartesiano  $U \times U$  y los números reales no negativos  $\mathbb{R}^+ \cup \{0\} = [0, \infty)$ , de modo que a cada par de elementos  $(x, y) \in U \times U$  se le asigna un número real no negativo  $r$  que define la distancia entre los puntos  $x$  y  $y$  de  $U$ .

$$\begin{aligned} d : U \times U &\rightarrow \mathbb{R}^+ \cup \{0\} = [0, \infty) \\ (x, y) &\mapsto r \Leftrightarrow d(x, y) = r \end{aligned} \quad (2.4)$$

La distancia  $d$  cumple con las siguientes condiciones:

- a)  $d(x, y) \geq 0$
- b)  $d(x, y) = 0 \Leftrightarrow x = y$
- c)  $d(x, y) = d(y, x)$
- d)  $d(x, z) \leq d(x, y) + d(y, z)$  ,  $\forall z \in U$

**Definición 1.1** Similitud: Una función  $s : U \times U \rightarrow R$  se llama de similitud si  $\forall x, y \in U$ :

- a)  $s(x, y) \leq s_0$
- b)  $s(x, x) = s_0$
- c)  $s(x, y) = s(y, x)$

donde  $s_0$  es un número real finito arbitrario.

**Definición 1.2** Similaridad métrica: Una función  $s$  se llama de similaridad métrica si, además de cumplir la definición anterior:

- a)  $s(x, y) = s_0 \implies x = y$   
 b)  $|s(x, y) + s(y, z)|s(x, z) \geq s(x, y)s(y, z) \quad \forall z \in U$

Las principales clasificaciones de medidas de similitud:

**Medidas de asociación.** Cuando se elige una medida de asociación, los grupos estarán formados por individuos parecidos de tal forma que la distancia entre ellos sea la más pequeña. Se utilizan cuando se está trabajando con objetos cuyas características están medidas en términos no métricos (medidas ordinales o nominales), y son una forma de medir la concordancia o conformidad entre los estados de dos columnas de datos. Suelen usarse para el caso de variables cualitativas y casos con datos binarios.

$$\text{Coeficientes de asociación} \left\{ \begin{array}{l} \text{Jaccard-Sneath} \quad S_J = \frac{a}{(a+u)} = \frac{a}{(a+b+c)} \\ \text{Coeficiente de emparejamiento simple} \quad S_{SM} = \frac{m}{(m+u)} = \frac{m}{n} = \frac{(a+d)}{(a+b+c+d)} \\ \text{Coeficiente de Yule} \quad S_Y = \frac{(ad-dc)}{(ad+bc)} \end{array} \right.$$

Estas medidas parten de la siguiente tabla de frecuencias:

	Variable 1	
Variable 2	Presencia	Ausencia
Presencia	a	b
Ausencia	c	d

**Medidas de correlación.** Las medidas de correlación representan la similitud por la correspondencia de patrones en las características. Elevadas correlaciones indican similitud, y bajas correlaciones indican falta de ella. Una medida de correlación de similitud no ve la magnitud de los valores, en su lugar ve los patrones de los valores. Estas medidas rara vez son utilizadas en el análisis de conglomerados, ya que en el análisis se hace énfasis en la magnitud de los objetos y no en su patrón de valores.

**Medidas de distancia.** Las medidas de distancia son las más utilizadas en el análisis de conglomerados; se trata de las distintas medidas entre los puntos del espacio definido por los individuos.

$$\text{Distancias} \left\{ \begin{array}{l} \text{Distancia euclidiana} \quad d(i, j) = \sqrt{\sum_k (x_{ik} - x_{jk})^2} \\ \text{Distancia de Minkowski} \quad d_q(i, j) = \left( \sum_k |x_{ik} - x_{jk}|^q \right)^{\frac{1}{q}} \\ \text{Distancia de Manhattan} \quad d_1(i, j) = \sum_k |x_{ik} - x_{jk}| \\ \text{Distancia de Tchebichez} \quad d_\infty(i, j) = \text{Max}_k (|x_{ik} - x_{jk}|) \end{array} \right.$$

**Distancia euclidiana.** La distancia euclidiana es sensible a las unidades de medida de las variables, las diferencias entre los valores de variables medidas con valores altos contribuirán en mayor medida que las diferencias entre los valores de las variables con valores bajos, por lo tanto, la variación en la escala determinará también cambios en la distancia entre los individuos. Es una de las distancias más utilizadas.

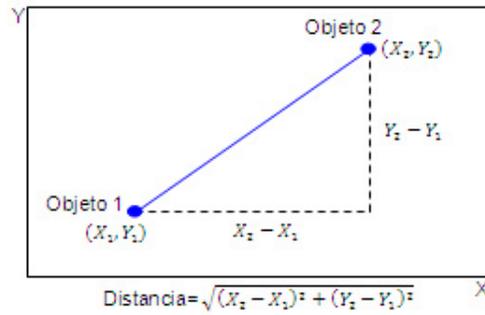


Figura 2.2: Distancia euclidiana

**Distancia de Mahalanobis.** Esta distancia es útil si se desea solucionar los problemas de escalas distintas y correlación de las variables, ya que es invariante bajo transformaciones lineales no singulares de las variables.

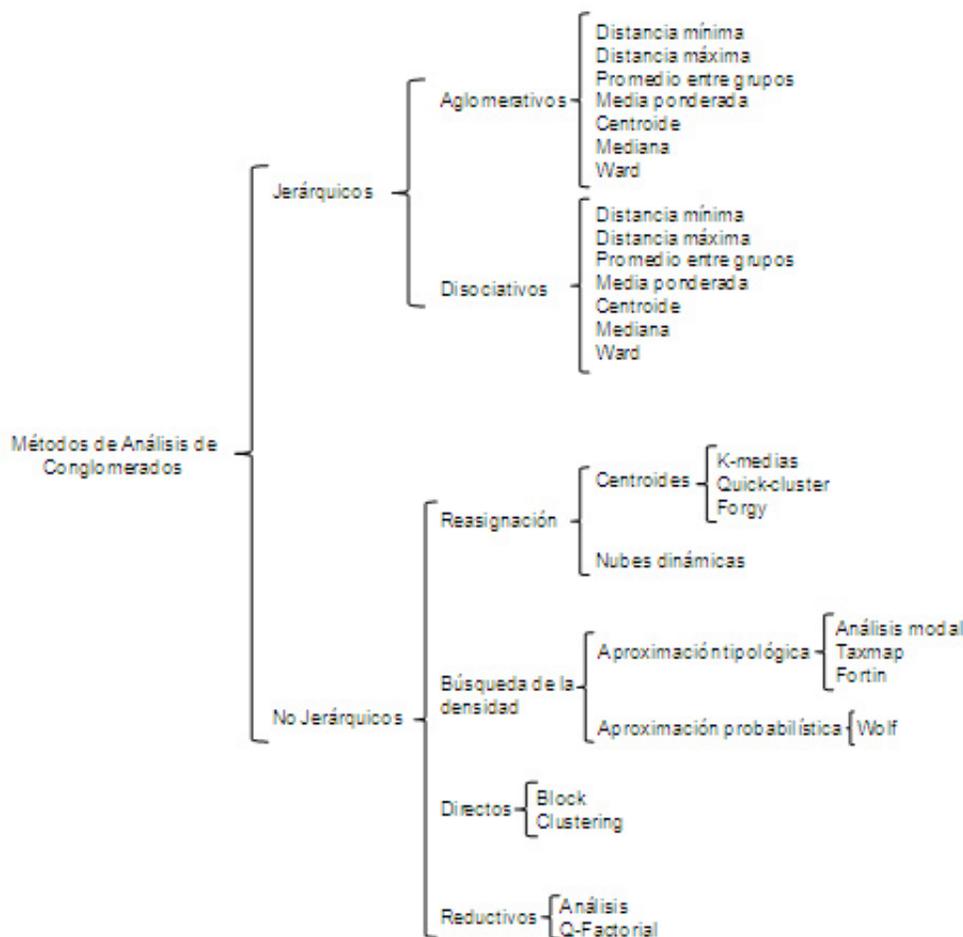
**Distancia Manhattan o ciudad.** Es un caso particular de la distancia de Minkowski cuando  $q=1$ . El valor de esta distancia es cero cuando la similitud es perfecta y aumenta a medida que los objetos son más disimilares.

**Distancia de Tchebichev.** Es el caso de la distancia de Minkowski cuando  $q$  tiende a infinito.

**Distancia de Minkowski.** Es una distancia genérica que da lugar a otras distancias en casos particulares y se define como la raíz  $q$ -ésima de la suma de las potencias  $q$ -ésimas de las diferencias, en valor absoluto, de las coordenadas de dos puntos considerados.

## 2.4. Técnicas de conglomeración.

Existen dos grandes grupos de técnicas de clasificación del análisis de conglomerados: Jerárquicos y No jerárquicos.



Fuente: Blsquerra A.R. Introducción conceptual al análisis multivariante.

Cuadro 2.3: Métodos de Análisis de conglomerados

### 2.4.1. Jerárquicos.

Los métodos jerárquicos se dividen en dos tipos de algoritmos: aglomerativos y divisivos. El algoritmo aglomerativo va agrupando unidades de análisis de tal forma que en cada paso se va disminuyendo el número de conglomerados hasta llegar a un solo conglomerado que contenga la totalidad de las unidades.

#### Propiedades matemáticas de los métodos jerárquicos.

Varias propiedades matemáticas pueden ser definidas en el análisis de conglomerados jerárquicos:

- Se utiliza una matriz de distancias o similitudes entre los elementos.
- Propiedad ultra métrica, dicha propiedad establece que:

$$h_{ij} \leq \max(h_{ik}, h_{jk}) \forall i, j, k. \text{ Donde } h_{ij} \text{ es la distancia entre los conglomerados } i \text{ y } j.$$

**2.4.1.1. Aglomerativos**

En estos algoritmos se produce una sucesión de conglomerados de tal manera que en cada paso el número de conglomerados va disminuyendo.

**Etapas de procedimiento del método aglomerativo:**

1. En esta primera etapa hay tantos conglomerados como elementos a agrupar. Se trata al conjunto de  $N$  elementos como una primera partición en conglomerados de máxima homogeneidad.
2. Se agrupan los conglomerados de la primera partición que sean más parecidos según la medida de similitud o disimilitud que se haya elegido. La segunda partición contiene un conglomerado menos que la primera. Posteriormente se recalculan las distancias entre la nueva clase y el resto, y nos encontramos en la misma situación que en el paso 1 pero con  $N - 1$  conglomerados.

El proceso se repite hasta que en la última partición se obtiene un solo grupo que contiene a todos los elementos. El resultado total de este proceso, que es una jerarquía indexada, podemos observarlo gráficamente en un dendrograma.

**2.4.1.2. Divisivos.**

Al igual que los algoritmos aglomerativos, estos algoritmos también producen unas sucesiones de conglomerados, pero en este caso, el número de ellos va disminuyendo en cada paso.

**Etapas de procedimiento del método divisivo:**

1. Se calcula el diámetro de cada conglomerado<sup>1</sup>.
2. Se selecciona el conglomerado con diámetro más grande. (Se divide en A y B).
3. Se selecciona el objeto  $i$  del conglomerado A que maximice el promedio de las disimilitudes a todos los demás objetos de A.
4. Si la diferencia entre la distancia promedio de  $i$  a los demás elementos de A y la distancia promedio de  $i$  a los elementos de B es mayor que cero, se mueve  $i$  de A a B.
5. Si no, se repite el proceso desde el paso 1) hasta que hayan tantos conglomerados como elementos.

**Dendrograma.**

El dendrograma o árbol jerárquico es la representación gráfica del proceso de agrupación en forma de árbol. Se forma de la siguiente manera:

1. En la parte inferior del gráfico se disponen los  $n$  elementos iniciales.
2. Las uniones entre elementos se indican por tres líneas rectas. Dos dirigidas a los elementos que se unen, y que son perpendiculares al eje de los elementos, y una paralela a este eje, que se sitúa al nivel en que se unen.
3. El proceso se repite hasta que todos los elementos están conectados por líneas rectas.

Si el dendrograma es cortado a un nivel de distancia dado, se obtiene una clasificación del número de conglomerados existentes a ese nivel y los elementos que los forman. El dendrograma es útil cuando los puntos tienen una estructura jerárquica, pero pueden ser engañosos si se interpretan mecánicamente, ya que dos puntos pueden parecer alejados cuando están próximos, y pueden parecer próximos cuando están alejados.

---

<sup>1</sup>El diámetro de un conglomerado es la distancia más grande entre cualquiera de dos de sus objetos.

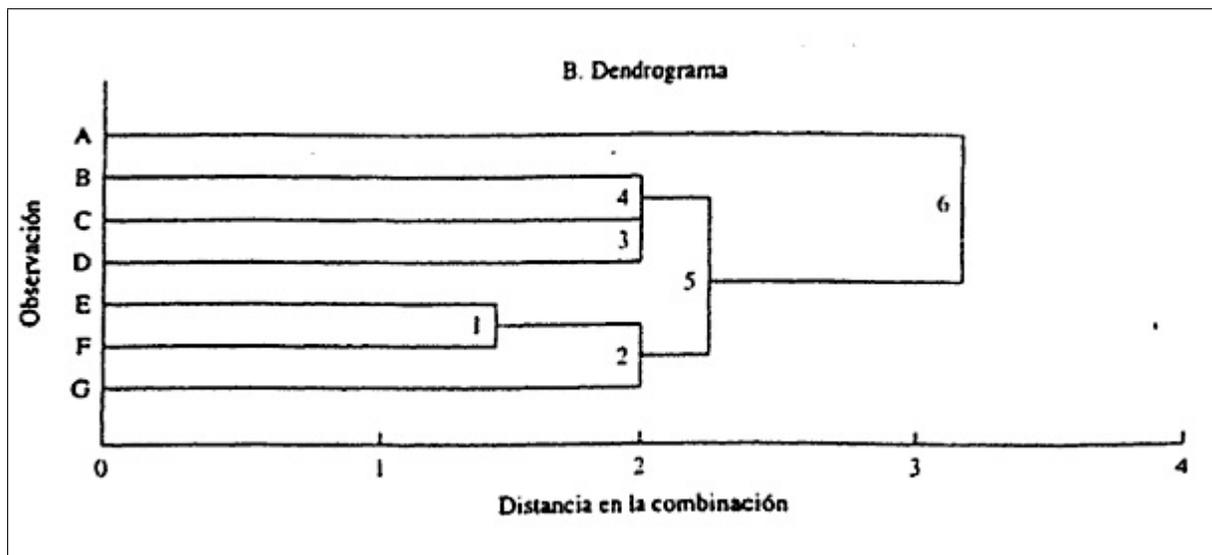


Figura 2.3: Ejemplo de dendrograma

**Gráfico de carámbanos.**

Se le denomina carámbanos porque se parece a las estalactitas de hielo que a veces se forma en la cornisa de los tejados cuando ha nevado.

En este gráfico aparece estructurado en columnas y filas en las que se sitúan los objetos de análisis o casos y el número de conglomerados correspondientes para cada paso del análisis respectivamente. La lectura del gráfico se hace de abajo a arriba. La ventaja de este gráfico es que en cada paso podemos ver fácilmente el número de conglomerados correspondientes.

		Diagrama de témpanos vertical										
		Caso										
Número de conglomerados		6:Libano	7:Libia	10:Tunez	8:Marrueco	5:Jordania	4:Iraq	3:Irán	11:Turquia	9:Siria	2:Egipto	1:Argelia
1		X	X	X	X	X	X	X	X	X	X	X
2		X	X	X	X	X	X	X	X	X	X	X
3		X	X	X	X	X	X	X	X	X	X	X
4		X	X	X	X	X	X	X	X	X	X	X
5		X	X	X	X	X	X	X	X	X	X	X
6		X	X	X	X	X	X	X	X	X	X	X
7		X	X	X	X	X	X	X	X	X	X	X
8		X	X	X	X	X	X	X	X	X	X	X
9		X	X	X	X	X	X	X	X	X	X	X
10		X	X	X	X	X	X	X	X	X	X	X

Figura 2.4: Gráfico de carámbanos (témpanos).

**2.4.1.3. Principales algoritmos de clasificación jerárquica**

**Distancia mínima.**

En este método se relaciona un elemento con un grupo si tiene la mayor similitud con cualquiera de los elementos individuales de ese grupo. Los dos primeros casos que se combinan son aquellos cuya distancia es la menor o cuya similitud es la máxima. La distancia entre el nuevo conglomerado y un caso individual se calcula como la mínima distancia entre el caso individual y un caso del conglomerado. Este método utiliza la distancia:

$$\bar{d}(h_k, h_i \cup h_j) = \min(d(h_i, h_k), d(h_j, h_k))$$

**Ejemplo 2.1:**

Supongamos que tenemos la siguiente matriz de datos originales con 4 individuos y 4 variables:

Matriz 1:

Individuo				
A	1	3	2.5	4
B	2	1	5	7
C	3	9	2	6.5
D	3.5	8	4.5	6

Ahora, calculamos la matriz de distancias entre individuos, utilizando la distancia euclidiana:

$$\begin{aligned}
 d(A, B) &= \sqrt{(1-2)^2 + (3-1)^2 + (2.5-5)^2 + (4-7)^2} = 4.5 \\
 d(A, C) &= \sqrt{(1-3)^2 + (3-9)^2 + (2.5-2)^2 + (4-6.5)^2} = 6.82 \\
 d(A, D) &= \sqrt{(1-3.5)^2 + (3-8)^2 + (2.5-4.5)^2 + (4-6)^2} = 6.26 \\
 d(B, C) &= \sqrt{(2-3)^2 + (1-9)^2 + (5-2)^2 + (7-6.5)^2} = 8.62 \\
 d(B, D) &= \sqrt{(2-3.5)^2 + (1-8)^2 + (5-4.5)^2 + (7-6)^2} = 7.25 \\
 d(C, D) &= \sqrt{(3-3.5)^2 + (9-8)^2 + (2-4.5)^2 + (6.5-6)^2} = 2.78
 \end{aligned}$$

entonces, tenemos la matriz de distancias:

Matriz 2:

	A	B	C	D
A	0	4.5	6.82	6.26
B	4.5	0	8.62	7.25
C	6.82	8.62	0	2.78
D	6.26	7.25	2.78	0

Ahora, iniciamos el proceso de conglomeración con el algoritmo de la distancia mínima:

*Etapa 0:*

Se inicia con la partición:

$$C_0 = (A), (B), (C), (D)$$

*Etapa 1:*

Se agrupan los individuos más cercanos. En este caso, de acuerdo a la matriz de distancias, son los elementos C y D, a una distancia de 2.78.

$$C_1 = (A), (B), (CD)$$

Ahora, se recalculan las distancias de un elemento al grupo (CD):

$$\begin{aligned}
 d(A, CD) &= \min(6.82, 6.26) = 6.26 \\
 d(B, CD) &= \min(8.62, 7.25) = 7.25
 \end{aligned}$$

y la nueva matriz de distancias es:

	A	B	CD
A	0	4.5	6.26
B		0	7.25
CD			0

Etapa 2:

Se fusionan los elementos más próximos a una distancia de 4.5, correspondiente a los elementos A y B:

$$C_2 = (AB), (CD)$$

Recalculando las distancias, tenemos:

$$d((AB), (CD)) = \min(6.26, 7.25) = 6.26$$

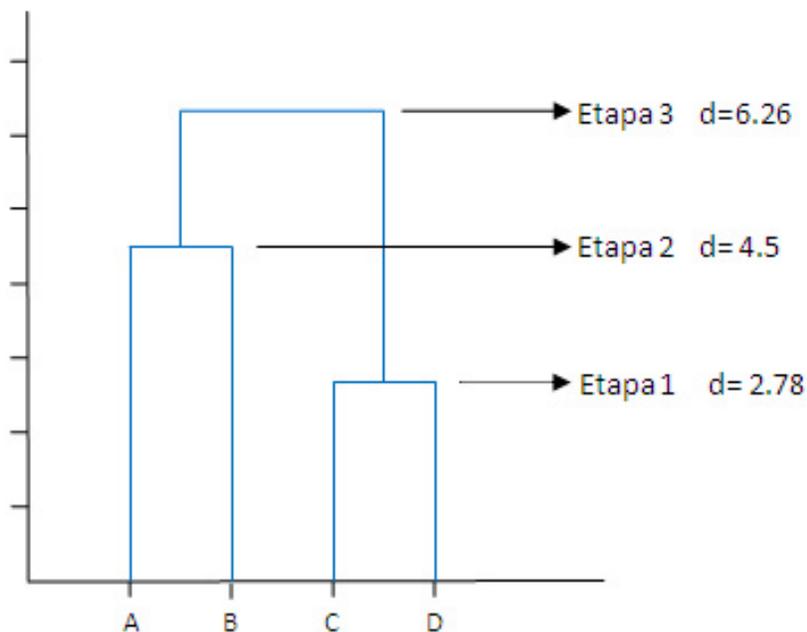
La nueva matriz de distancias es:

$$\begin{array}{c} \begin{array}{cc} & AB & CD \\ AB & \left( \begin{array}{cc} 0 & 6.26 \end{array} \right) \\ CD & \left( \begin{array}{cc} & 0 \end{array} \right) \end{array} \end{array}$$

Etapa 3:

En este último paso, se fusionan ambos conglomerados en uno solo, a una distancia de 6.26.

Gráficamente podemos observar el procedimiento mediante un dendrograma:



### Distancia máxima.

En este método la similitud de un elemento con un grupo se calcula como la similitud de dicho elemento con el individuo más alejado de ese grupo. La distancia entre dos conglomerados se calcula como la distancia entre sus dos puntos más alejados. La distancia utilizada es:

$$\bar{d}(h_k, h_i \cup h_j) = \text{Max}(d(h_i, h_k), d(h_j, h_k))$$

Este método tiende a producir grupos de igual diámetro y se ve distorsionado por la presencia de datos atípicos moderados, por lo tanto, su uso se encuentra en decadencia, ya que presenta el inconveniente de alargar mucho el proceso y dar como resultado agrupaciones encadenadas.

**Ejemplo 2.2:**

Utilizamos la misma matriz de distancias (matriz 2).

*Etapa 0:*

Se inicia con la partición:

$$C_0 = (A), (B), (C), (D)$$

*Etapa 1:*

Se fusionan, igual que en el método anterior, los elementos más próximos: C y D a una distancia de 2.78.

$$C_1 = (A), (B), (CD)$$

Se recalculan las distancias de cada elemento al grupo formado (CD) como la máxima distancia de un elemento a C y a D:

$$d(A, CD) = \max(6.82, 6.26) = 6.82$$

$$d(B, CD) = \max(8.62, 7.25) = 8.62$$

Se forma una nueva matriz de distancias:

$$\begin{array}{c} A \quad B \quad CD \\ A \quad \left( \begin{array}{ccc} 0 & 4.5 & 6.82 \\ B & & 0 & 8.62 \\ CD & & & 0 \end{array} \right) \end{array}$$

*Etapa 2.*

Se fusionan los elementos más próximos a una distancia de 4.5, formando el grupo (AB):

$$C_2 = (AB), (CD)$$

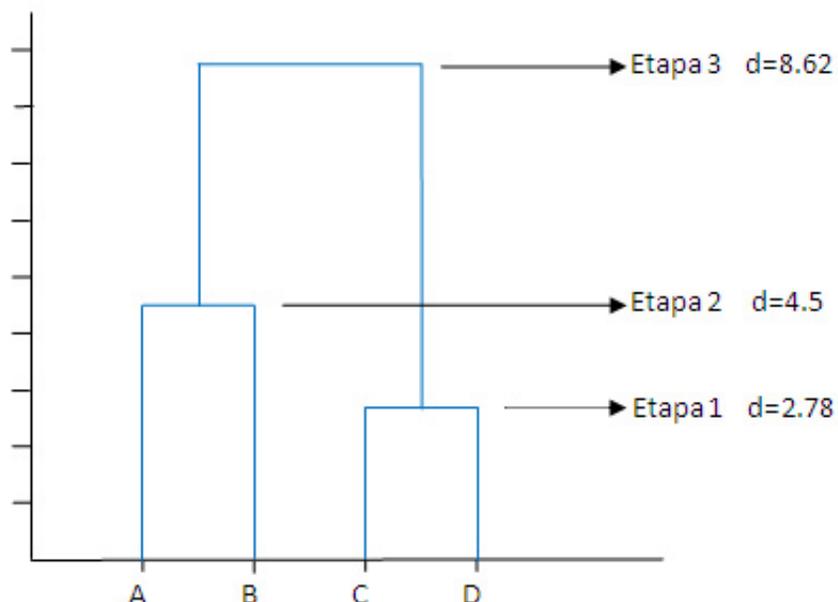
Recalculando las distancias tenemos:

$$d((AB), (CD)) = \max(6.82, 8.62)$$

La nueva matriz de distancias es:

$$\begin{array}{c} AB \quad CD \\ AB \quad \left( \begin{array}{cc} 0 & 8.62 \\ CD & & 0 \end{array} \right) \end{array}$$

*Etapa 3.* Se fusionan ambos grupos (AB) y (CD) a una distancia de 8.62.



### Promedio entre grupos.

En este método se ponderan los nuevos miembros admitidos en un conglomerado con el mismo peso que los existentes hasta entonces. El método combina conglomerados de modo que la distancia media entre todos los casos en el conglomerado resultante sea la menor posible. La distancia entre dos conglomerados se toma como la media de las distancias entre todos los posibles pares de casos en el conglomerado resultante. La distancia utilizada es:

$$\bar{d}(h_k, h_i \cup h_j) = \left(\frac{1}{2}\right) d(h_i, h_k) + \left(\frac{1}{2}\right) d(h_j, h_k)$$

Este método tiene tendencia a producir conglomerados compactos.

El gráfico ilustra los algoritmos: distancia mínima, distancia máxima y promedio.

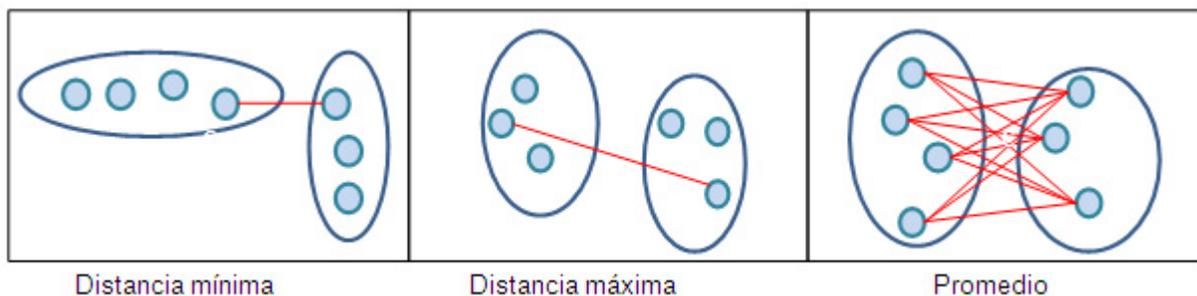


Figura 2.5: Algoritmos de clasificación

### Ejemplo 2.3.

*Etapa 0:*

Se inicia con la partición:

$$C_0 = (A), (B), (C), (D)$$

*Etapa 1:* Se fusionan los elementos más cercanos, C y D a una distancia de 2.78.

$$C_1 = (A), (B), (CD)$$

Se recalculan las distancias:

$$d(A, CD) = \frac{1}{2} (6.82) + \frac{1}{2} (7.25) = 6.54$$

$$d(B, CD) = \frac{1}{2} (8.62) + \frac{1}{2} (7.25) = 7.935$$

Se forma una nueva matriz de distancias:

$$\begin{array}{c} A \quad B \quad CD \\ A \left( \begin{array}{ccc} 0 & 4.5 & 6.54 \\ B & & 0 & 7.935 \\ CD & & & 0 \end{array} \right)$$

*Etapa 2:*

Se fusionan los elementos más cercanos A y B, a una distancia de 4.5

$$C_2 = (AB), (CD)$$

Se recalculan las distancias:

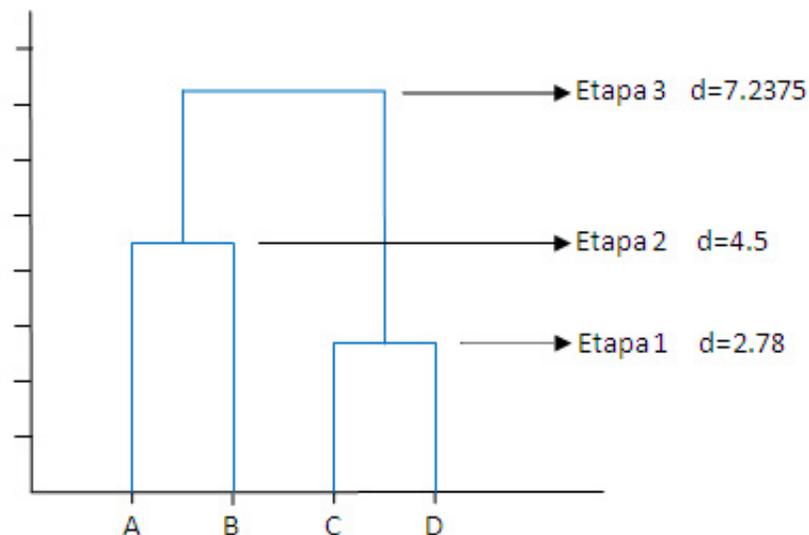
$$d((AB), (CD)) = \frac{1}{2} (6.54) + \frac{1}{2} (7.935) = 7.2375$$

Se forma una nueva matriz de distancias:

$$\begin{array}{c} AB \quad CD \\ AB \left( \begin{array}{cc} 0 & 7.2375 \\ CD & & 0 \end{array} \right)$$

*Etapa 3:*

Se fusionan ambos conglomerados (AB) y (CD) en uno solo, a la distancia de 7.2375.



### Media ponderada.

En este método, la distancia entre dos conglomerados se define como la media de las distancia entre todos los pares de casos en los que un miembro del par es de cada uno de los conglomerados. La distancia se define ponderando respecto a  $n_i$  y  $n_j$ ; es decir, ponderando con respecto al número de individuos de  $h_i$  y de  $h_j$  de la siguiente forma:

$$\bar{d}(h_k, h_i \cup h_j) = \left( \frac{n_i}{(n_i + n_j)} \right) d(h_i, h_k) + \left( \frac{n_j}{(n_i + n_j)} \right) d(h_j, h_k)$$

### Centroide.

Este método calcula la distancia entre dos conglomerados como la distancia entre sus medias para todas las variables. Este método hace intervenir el número de individuos de  $h_i$  y de  $h_j$  que son  $n_i$  y  $n_j$  respectivamente. La distancia utilizada es:

$$\bar{d}(h_k, h_i \cup h_j) = \left( \frac{n_i}{(n_i + n_j)} \right) d(h_i, h_k) + \left( \frac{n_j}{(n_i + n_j)} \right) d(h_j, h_k) - \left( \frac{n_i n_j}{(n_i + n_j)^2} \right) d(h_i, h_j)$$

### Ejemplo 2.4:

Para los métodos de Centroides, Mediana y Ward utilizaremos la distancia euclidiana al cuadrado, con la misma matriz de datos original (matriz 1).

Calculando la distancia euclidiana al cuadrado:

$$\begin{aligned} d(A, B) &= \left( \sqrt{(1-2)^2 + (3-1)^2 + (2.5-5)^2 + (4-7)^2} \right)^2 = 20.25 \\ d(A, C) &= \left( \sqrt{(1-3)^2 + (3-9)^2 + (2.5-2)^2 + (4-6.5)^2} \right)^2 = 46.5 \\ d(A, D) &= \left( \sqrt{(1-3.5)^2 + (3-8)^2 + (2.5-4.5)^2 + (4-6)^2} \right)^2 = 39.25 \\ d(B, C) &= \left( \sqrt{(2-3)^2 + (1-9)^2 + (5-2)^2 + (7-6.5)^2} \right)^2 = 74.25 \\ d(B, D) &= \left( \sqrt{(2-3.5)^2 + (1-8)^2 + (5-4.5)^2 + (7-6)^2} \right)^2 = 52.25 \\ d(C, D) &= \left( \sqrt{(3-3.5)^2 + (9-8)^2 + (2-4.5)^2 + (6.5-6)^2} \right)^2 = 7.75 \end{aligned}$$

La matriz de distancias original que utilizaremos para los siguientes métodos es:

Matriz 3:

$$\begin{array}{c} \\ A \\ B \\ C \\ D \end{array} \begin{array}{cccc} A & B & C & D \\ \left( \begin{array}{cccc} 0 & 20.25 & 46.5 & 39.25 \\ 20.25 & 0 & 74.25 & 52.5 \\ 46.5 & 74.25 & 0 & 7.75 \\ 39.25 & 52.5 & 7.75 & 0 \end{array} \right) \end{array}$$

Ahora iniciamos con el análisis de conglomerados con el algoritmo del centroide:

*Etapa 0.*

Se inicia con la partición:

$$C_0 = (A), (B), (C), (D)$$

*Etapa 1.*

Se fusionan los elementos más próximos, que corresponde a C y D, a una distancia de 7.75. A este conglomerado para fines prácticos, lo llamaremos  $C_{CD}$ .

Ahora, se calcula el vector de medias (centroide) del grupo formado (CD), utilizando la matriz de datos originales:

$$\left( \frac{(3+3.5)}{2}, \frac{(9+8)}{2}, \frac{(2+4.5)}{2}, \frac{(6.5+6)}{2} \right) = (3.25 + 8.5 + 3.25 + 6.25)$$

Y se forma una nueva matriz de datos:

$$\begin{array}{l} A \\ B \\ C_{CD} \end{array} \begin{pmatrix} 1 & 3 & 2.5 & 4 \\ 2 & 1 & 5 & 7 \\ 3.25 & 8.5 & 3.25 & 6.25 \end{pmatrix}$$

Se recalculan las distancias de A y B al conglomerado  $C_{CD}$ :

$$d(A, C_{CD}) = \left( \sqrt{(1 - 3.25)^2 + (3 - 8.5)^2 + (2.5 - 3.25)^2 + (4 - 6.25)^2} \right)^2 = 41.37$$

$$d(B, C_{CD}) = \left( \sqrt{(2 - 3.25)^2 + (1 - 8.5)^2 + (5.5 - 3.25)^2 + (7 - 6.25)^2} \right)^2 = 61.23$$

Y la nueva matriz de distancias es:

$$\begin{array}{l} A \\ B \\ C_{CD} \end{array} \begin{array}{c} A \quad B \quad C_{CD} \\ \begin{pmatrix} 0 & 20.25 & 41.37 \\ & 0 & 61.43 \\ & & 0 \end{pmatrix} \end{array}$$

*Etapa 2:*

Se fusionan los elementos más próximos, A y B con una distancia de 20.25. A este conglomerado le llamaremos  $C_{AB}$ .

Se calcula el vector de medias de  $C_{AB}$ :

$$\left( \frac{(1+2)}{2}, \frac{(3+1)}{2}, \frac{(2.5+5)}{2}, \frac{(4+7)}{2} \right)$$

Se forma una nueva matriz de datos:

$$\begin{array}{l} C_{AB} \\ C_{CD} \end{array} \begin{pmatrix} 1.5 & 2 & 3.75 & 5.5 \\ 3.25 & 8.5 & 3.25 & 6.25 \end{pmatrix}$$

Se recalculan las distancias entre  $C_{AB}$  y  $C_{CD}$ :

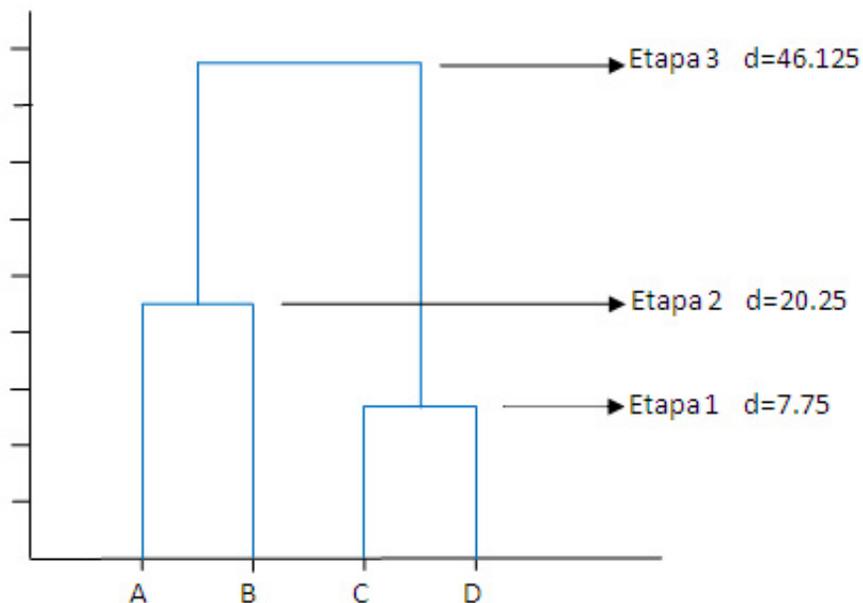
$$d(C_{AB}, C_{CD}) = \left( \sqrt{(1.5 - 3.25)^2 + (2 - 8.5)^2 + (3.75 - 3.25)^2 + (5.5 - 6.25)^2} \right)^2 = 46.125$$

Se forma una nueva matriz de distancias:

$$\begin{array}{l} C_{AB} \\ C_{CD} \end{array} \begin{array}{c} C_{AB} \quad C_{CD} \\ \begin{pmatrix} 0 & 46.125 \\ & 0 \end{pmatrix} \end{array}$$

*Etapa 3:*

Se termina el análisis, fusionándose ambos conglomerados en uno solo, con distancia de 46.125.



### Mediana.

En este método los dos conglomerados que están siendo combinados pesan lo mismo en el cálculo del centroide y es indiferente el número de casos de cada uno. Solamente se utiliza la distancia euclidiana al cuadrado:

$$\bar{d}(h_k, h_i \cup h_j) = \left(\frac{1}{2}\right) d(h_i, h_k) + \left(\frac{1}{2}\right) d(h_j, h_k) - \left(\frac{1}{4}\right) d(h_i, h_j)$$

### Ejemplo 2.5

*Etapa 0.*

Se inicia con la partición:

$$C_0 = (A), (B), (C), (D)$$

*Etapa 1.*

Se fusionan los elementos más próximos C y D, con una distancia de 7.75.

Se recalculan las distancias:

$$d(A, CD) = \left(\frac{1}{2}\right) d(A, C) + \left(\frac{1}{2}\right) d(A, D) - \left(\frac{1}{4}\right) d(C, D) = \left(\frac{1}{2}\right) (46.5) + \left(\frac{1}{2}\right) (39.25) - \left(\frac{1}{4}\right) (7.75) = 40.93$$

$$d(B, CD) = \left(\frac{1}{2}\right) d(B, C) + \left(\frac{1}{2}\right) d(B, D) - \left(\frac{1}{4}\right) d(C, D) = \left(\frac{1}{2}\right) (74.25) + \left(\frac{1}{2}\right) (52.5) - \left(\frac{1}{4}\right) (7.75) = 61.4375$$

Se forma una nueva matriz de distancias:

$$\begin{array}{c} A \quad B \quad (C, D) \\ A \quad \begin{pmatrix} 0 & 20.25 & 40.93 \\ B & & 0 & 61.4375 \\ (C, D) & & & 0 \end{pmatrix} \end{array}$$

*Etapa 2.*

Se fusionan los elementos A y B, con una distancia de 20.25

Se recalculan las distancias entre conglomerados:

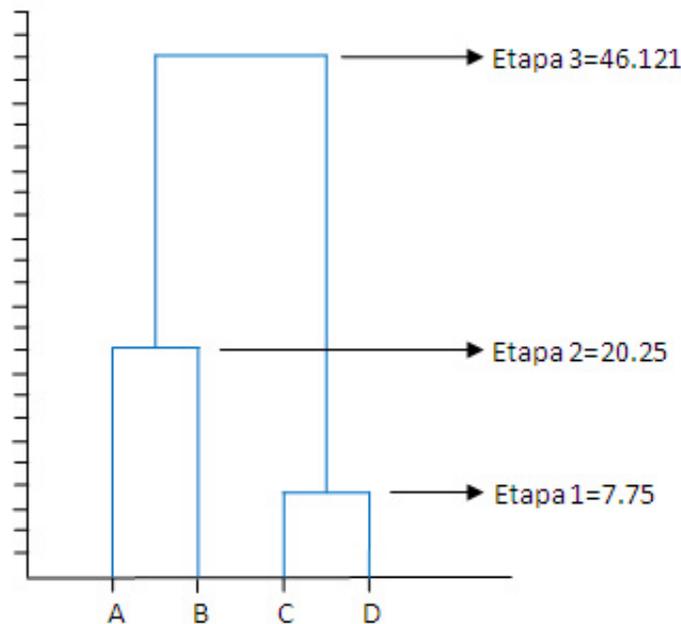
$$d((CD), (AB)) = \frac{1}{2}d((CD), A) + \frac{1}{2}d((CD), B) - \frac{1}{4}d(A, B) = \frac{1}{2}(40.93) + \frac{1}{2}(61.4375) - \frac{1}{4}(20.25) = 46.121$$

Se forma una nueva matriz de distancias:

$$\begin{array}{c} (A, B) \quad (C, D) \\ (A, B) \left( \begin{array}{cc} 0 & 20.25 \\ 20.25 & 0 \end{array} \right) \\ (C, D) \left( \begin{array}{cc} 40.93 & 61.4375 \\ 61.4375 & 0 \end{array} \right) \end{array}$$

*Etapa 3.*

Se fusionan ambos conglomerados a una distancia de 46.121.



**Ward.**

En este método se calculan las medias para todas las variables en cada uno de los conglomerados. Posteriormente, para cada objeto se calcula la distancia euclidiana cuadrada a las medias de los grupos; estas distancias se suman a todos los objetos, es decir, su desviación cuadrática. En cada etapa del análisis se van fusionando los dos conglomerados con el menor incremento en la suma total de los cuadrados de sus distancias,

$$W = \sum_g \sum_{i \in g} (X_{ig} - \bar{X}_g)' (X_{ig} - \bar{X}_g)$$

donde  $g$  es un vector de conglomerados,  $\bar{X}_g$  es el vector que contiene la media de cada conglomerado (centroide) y  $X_{ig}$ ,  $i \in g$ , es un vector que contiene los elementos del conglomerado  $i$ .

**Ejemplo 2.6.**

*Etapa 0.*

Se inicia con la partición:

$$C_0 = (A), (B), (C), (D)$$

*Etapa 1.*

En esta etapa se formarán 3 conglomerados: uno con dos elementos, y dos con un elemento cada uno. Para saber cuáles se unirán, hacemos los siguientes cálculos:

- El vector de medias (centroide) del conglomerado formado. Para calcular el centroide de un conglomerado, supongamos que queremos medir la distancia entre los conglomerados  $C_j$  (compuesto por  $n_j$  elementos) y  $C_i$  (formado a su vez por dos conglomerados  $C_{i1}$  y  $C_{i2}$  con  $n_{i1}$  y  $n_{i2}$  elementos respectivamente). Sean  $m^i$ ,  $m^{i1}$  y los centroides ( $n$  dimensionales) de los conglomerados mencionados anteriormente. Así, el centroide del conglomerado  $C_i$  vendrá dado en notación vectorial por:

$$m^i = \frac{n_{i1}m^{i1} + n_{i2}m^{i2}}{n_{i1} + n_{i2}}$$

cuyos componente serán:

$$m_l^i = \frac{n_{i1}m_l^{i1} + n_{i2}m_l^{i2}}{n_{i1} + n_{i2}} \quad l = 1, \dots, n$$

- La suma de cuadrados de los errores del conglomerado  $k$ , es decir, la distancia euclidiana al cuadrado entre cada elemento del conglomerado  $k$  a su centroide:

$$E_k = \sum_{i=1}^{n_k} \sum_{j=1}^n (x_{ij}^k - m_j^k)^2 = \sum_{i=1}^{n_k} \sum_{j=1}^n (x_{ij}^k)^2 - n_k \sum_{j=1}^n (m_j^k)^2$$

- La suma de cuadrados de los errores para todos los conglomerados.
- El incremento de  $E$ :  
Supongamos ahora que los conglomerados  $C_p$  y  $C_q$  se unen resultando un nuevo conglomerado  $C_i$ . Entonces el incremento de  $E$  será:

$$\Delta E_{pq}$$

Así, el menor incremento de los errores cuadráticos es proporcional a la distancia euclidiana al cuadrado de los centroides de los conglomerados unidos.

Entonces, tenemos como posibles fusiones una combinación  $\binom{4}{2} = 6$

Partición	Centroide	$E_k$	$E$	$\Delta E$
(A,B),C,D	$C_{(A,B)} = (1.5, 2, 3.75, 5.5)$	$E_{AB} = 10.125$ $E_C = E_D = 0$	10.125	10.125
(A,C),B,D	$C_{(A,C)} = (2, 6, 2.25, 5.25)$	$E_{AC} = 23.25$ $E_B = E_D = 0$	23.25	23.25
(A,D),B,C	$C_{(A,D)} = (2.25, 5.5, 3.5, 5)$	$E_{AD} = 19.6235$ $E_B = E_C = 0$	19.625	19.625
(B,C),A,D	$C_{(B,C)} = (2.5, 5, 3.5, 6.75)$	$E_{BC} = 37.125$ $E_A = E_D = 0$	37.125	37.125
(B,D),A,C	$C_{(B,D)} = (2.75, 4.5, 4.75, 6.5)$	$E_{BD} = 26.25$ $E_A = E_C = 0$	26.25	26.25
(C,D),A,B	$C_{(C,D)} = (3.25, 8.5, 3.25, 6.25)$	$E_{CD} = 3.875$ $E_A = E_B = 0$	3.875	3.875

Se deduce que se fusionan los elementos C y D.

*Etapa 2.*

La configuración queda: (C,D), A, B, y combinaciones posibles  $\binom{3}{2} = 3$ .

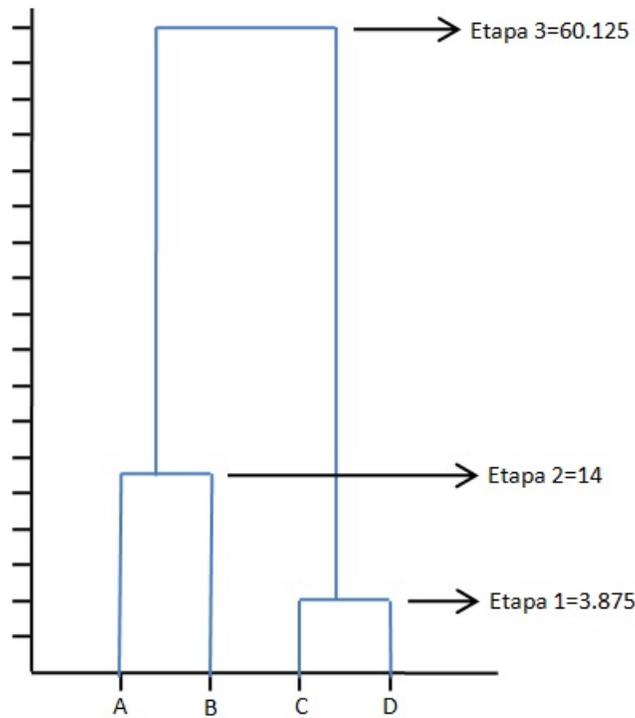
Partición	Centroide	$E_k$	$E$	$\Delta E$
(A,C,D),B	$C_{(A,C,D)} = (2.5, 6.7, 3, 5.5)$	$E_{ACD} = 31.17$ $E_B = 0$	31.17	27.29
(B,C,D),A	$C_{(B,C,D)} = (2.8, 6, 3.8, 6.5)$	$E_{BCD} = 44.83$ $E_A = 0$	44.83	40.96
(A,B),(C,D)	$C_{(A,B)} = (1.5, 2, 3.75, 5.5)$ $C_{(CD)} = (3.25, 8.5, 3.25, 6.25)$	$E_{AB} = 10.125$ $E_{CD} = 3.875$	14	10.125

En esta etapa se fusionan los elementos A y B.

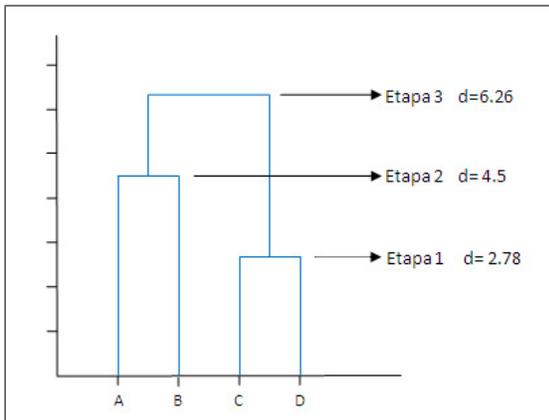
*Etapa 3.*

En esta etapa se fusionarán los dos conglomerados existentes (A,B) y (C,D). Los valores del centroide de los incrementos de distancias son:

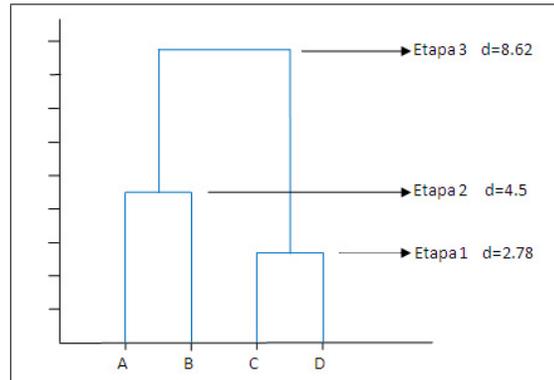
Partición	Centroide	$E_k$	$E$	$\Delta E$
(A,B,C,D)	$C_{(A,B,C,D)} = (2.375, 5.25, 3.5, 5.875)$	$E_{ABCD} = 60.125$	60.125	46.125



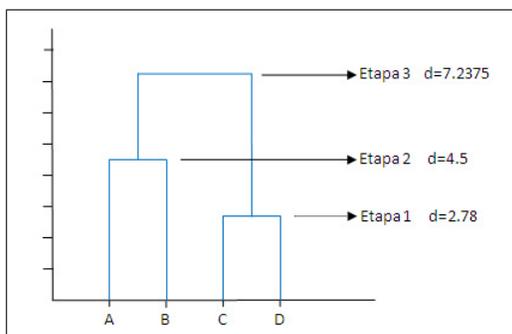
MÉTODOS JERÁRQUICOS



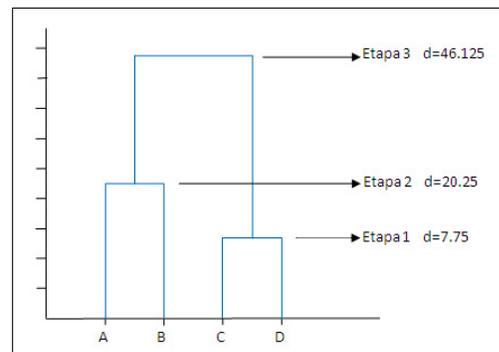
(a) Distancia mínima



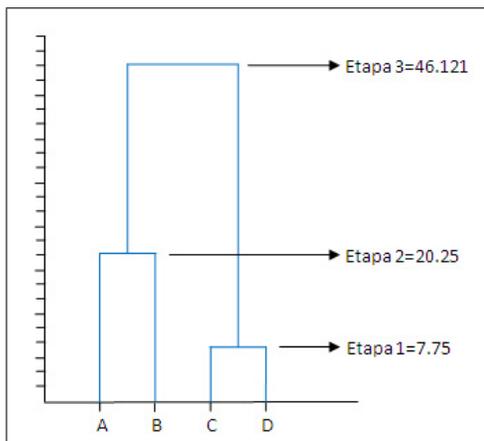
(b) Distancia máxima



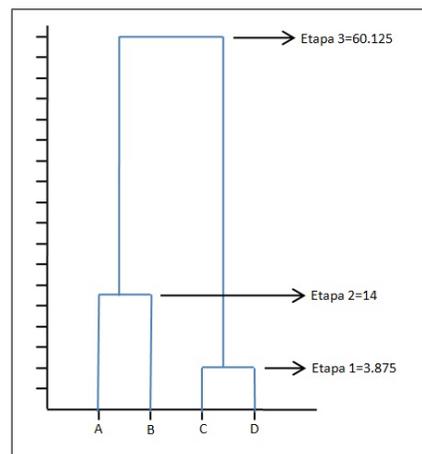
(c) Promedio entre grupos



(d) Centroide



(e) Mediana



(f) Ward

### 2.4.2. No jerárquicos.

Los métodos de conglomeración no jerárquicos también son denominados como *métodos partitivos* o de optimización, dado que éstos, tienen por objetivo realizar una sola partición de los individuos en  $K$  grupos. La denominación *no jerárquica* alude a la no existencia de una estructura vertical de dependencia entre los grupos formados, y por consiguiente, éstos no se presentan en distintos niveles de jerarquía. El análisis precisa que el investigador fije de antemano el número de conglomerados en que se quieren agrupar los datos. Esto es, probablemente, la mayor diferencia con los métodos de conglomeración jerárquicos; otra diferencia es, que los métodos no jerárquicos trabajan con la matriz de datos original y no requieren la conversión en una matriz de similitud. Cabe mencionar que no existe una metodología que indique cuál es el número óptimo de conglomerados, por lo tanto, esto debe ser resuelto repitiendo la prueba de conglomeración con diferentes números a fin de tantear la clasificación que mejor se ajuste al objetivo del problema, o la de más clara interpretación.

#### 2.4.2.1. Métodos de reasignación:

En los métodos de reasignación un individuo asignado a un grupo en un determinado paso del procedimiento puede ser reasignado a otro grupo en un paso posterior del procedimiento si esto mejora el criterio de selección. El proceso termina cuando ya no quedan individuos para reasignar y mejorar la clasificación. Dentro de los algoritmos conocidos está:

##### 1. Centroides.

Estos métodos se denominan de Centroides porque hacen la clasificación calculando la distancia respecto a los centroides de los conglomerados.

- a) K-medias. En este trabajo se hace uso del método de k- medias cuyo objetivo es el de minimizar las distancias de los elementos dentro de cada conglomerado, al mismo tiempo que busca maximizar la distancia de los elementos pertenecientes a otro conglomerado.

El algoritmo usado es:

Paso 1. Se seleccionan los centroides iniciales de los  $k$  conglomerados:  $c_1, c_2, \dots, c_k$ .

Los centroides iniciales pueden ser considerados de las siguientes maneras:

- 1) Usando las primeras  $k$  observaciones.
- 2) Eligiendo aleatoriamente  $k$  observaciones.
- 3) Tomando cualquier partición al azar en  $k$  conglomerados y calculando sus centroides iniciales.

Paso 2. Se asigna cada observación  $x_i$  del conjunto de datos al conglomerado  $C(i)$  cuyo centroide  $c_i$  está más cerca a  $x_i$ :

$$C_i = \operatorname{argmin}_{1 \leq k \leq K} \|x_i - c_k\|$$

Paso 3. Para cada uno de los conglomerados obtenidos se vuelve a calcular su centroide con base en los elementos que los conforman y se minimiza la suma de cuadrados dentro del conglomerado.

Paso 4. Se repite el paso 2 hasta que la nueva asignación de los casos a grupos no cambie respecto a la iteración anterior.

- b) Quick-cluster.

Este método es utilizado cuando se conoce el número de conglomerados que se desea obtener. El algoritmo utilizado es el siguiente:

Utilizando la siguiente notación:

$NC$ - Número de conglomerados a obtener.

$M_i$ - Media del  $i$ -ésimo conglomerado.

$X_k$ - Vector de la  $k$ -ésima observación.

$d(x_i, x_j)$ - Distancia Euclidiana entre el vector  $x_i$  y  $x_j$ .

$d_{mn}$ - La distancia  $\min_{i,j} d(M_i, M_j)$

$\varepsilon$ - Criterio de convergencia.

Algoritmo:

Paso 1. Selección de los centros iniciales de conglomerados.

a) Si la  $\min_i d(x_k, M_i) > d_{mn}$  y  $d(x_k, M_m) > d(x_k, M_n)$ , entonces,  $x_k$  reemplaza a  $M_n$ .

Si la  $\min_i d(x_k, M_i) > d_{mn}$  y  $d(x_k, M_m) < d(x_k, M_n)$ , entonces,  $x_k$  reemplaza a  $M_n$ .

b) Si no son especificados los centros de conglomerados, entonces el proceso inicia con cada media del conglomerado como su centro inicial.

Paso 2. Se asigna cada observación (individuo) al conglomerado más cercano, donde la distancia euclidiana es mínima entre la observación y el centro del conglomerado.

El algoritmo se detiene cuando se ha alcanzado el máximo número de iteraciones, en donde el mayor cambio entre los conglomerados es  $\varepsilon$  veces la menor distancia entre los centros iniciales de los conglomerados [10].

c) Forgy.

Este método inicia con cualquier configuración inicial, entendiéndose por la misma, un conjunto de  $k$  centroides escogidos (ya sea al azar, o con una estrategia en la cual dichos centroides se encuentre lo suficientemente separados unos de otros, o por medio de una partición del conjunto de objetos en  $k$  conglomerados). En el caso de que se haya comenzado por un conjunto de  $k$  centroides, se obtendrá la partición correspondiente sin más que asignar cada objeto que se pretende clasificar al centroide más cercano. Una vez que se haya llevado a cabo la asignación de todos los objetos a clasificar, se calculan los  $k$  nuevos centroides, los cuales, se utilizarán como nuevos "atractores", es decir, se asignará cada uno de los  $n$  objetos a clasificar al centroide más cercano. Estos dos pasos anteriores de recálculo de centroides y reasignación de objetos se mantendrá hasta que se alcance algún criterio de parada. Ejemplo de criterios de parada pueden ser: un número máximo de iteraciones, el que no se produzcan cambios en los conglomerados en dos iteraciones consecutivas, entre otras.

El algoritmo utilizado es:

Paso 1. Comenzar con cualquier configuración inicial.

Ir al paso 2 si se comienza por un conjunto de  $k$  centroides.

Ir al paso 3 si se comienza por una partición del conjunto de objetos en  $k$  grupos.

Paso 2. Asignar cada objeto a clasificar al centroide más próximo. Los centroides permanecen fijos en este paso.

Paso 3. Calcular los nuevos  $k$  centroides como los baricentros de los  $k$  conglomerados obtenidos.

Paso 4. Se alternan los pasos 2 y 3 hasta que se alcance un determinado criterio de convergencia.

## 2. Nubes dinámicas.

En este método los conglomerados no están caracterizados por su centro de gravedad, sino por un cierto número de individuos a clasificar que constituyen los nódulos de los conglomerados. El recentrado de los

conglomerados en torno a sus núcleos se hace cuando todos los individuos han sido asignados a un conglomerado.

#### 2.4.2.2. Búsqueda de densidad.

En estos métodos los grupos se forman por localización de regiones con alta densidad o concentración de datos. Presenta dos tipos de aproximación:

##### 1. Aproximación tipológica.

Los grupos se forman buscando las zonas en las cuales se da una mayor concentración de individuos. Los algoritmos utilizados en este método son:

###### a) Análisis modal

El análisis modal de Wishart [7] es una derivación del algoritmo del vecino más cercano, en donde se buscan subgrupos naturales de los datos mediante la búsqueda de superficies de densidad discontinuas en la distribución de los datos. La búsqueda se realiza considerando una esfera de algún radio,  $R$ , alrededor de cada punto y contando el número de puntos que caen dentro de ese radio. Los objetos (individuos) son etiquetados posteriormente como *densos* o *no densos* dependiendo si hay esferas que contienen más o menos puntos que el parámetro de vinculación,  $K$ , el cual es ajustado dependiendo del número de objetos en el conjunto de datos.

###### b) Taxmap.

Una de las principales características de este método es que presenta el problema de no presentar conglomerados claramente aislados, sino que forman uno continuo. Se caracteriza porque se debe introducir un valor corte, lo que da cierta subjetividad a los resultados y a su vez, permite adaptar la técnica al tipo de datos y análisis que se desea. Esta técnica utiliza el concepto natural de agrupación, de este modo, tiende a controlar el desplazamiento del centroide, por la adición de un nuevo miembro al centro del conglomerado.

##### 2. Aproximación probabilística. Donde se parte del postulado de que las variables siguen una ley de probabilidad según la cual los parámetros varían de un grupo a otro. Se trata de encontrar los individuos que pertenezcan a la misma distribución. El algoritmo más conocido es el método de Wolf.

#### 2.4.2.3. Método Directo.

Aquí se pueden clasificar simultáneamente individuos y variables. Las entidades agrupadas ya no son los individuos o las variables, sino que son las observaciones.

##### 1. Block-Clustering.

En este método, cada bloque es definido por un conglomerado de casos y un conglomerado de variables como si cada variable en el bloque es constante para los casos en el bloque, excepto aquellas que pertenecen a otro bloque [4].

#### 2.4.2.4. Método reductivo.

##### 1. Análisis Q-Factorial.

Este método consiste en buscar factores en el espacio de los individuos, correspondiendo cada factor a un grupo.

## 2.5. Validación del número de conglomerados.

No hay un criterio óptimo para la elección del número de conglomerados, por lo tanto esta elección se deja a manos del investigador, y en este caso se solucionará con la inspección de dendrogramas y posteriormente se realizará un análisis discriminante con los resultados obtenidos.

De igual manera, no existe una metodología exacta para la validación de la selección de los conglomerados, por lo tanto, se proponen las siguientes:

##### 1. Validación interna.

Se verifica:

Homogeneidad de los conglomerados respecto a las variables utilizadas para construir la partición, medida por la varianza intra-clase.

- a) Separación de los conglomerados: variación inter-clases.
- b) Estabilidad frente a pequeñas perturbaciones.

##### 2. Validación externa (empírica).

Se verifica que:

- a) Caracterización de los conglomerados mediante la información conocida sobre los individuos.

# Modelo Aplicado.

---

Con lo ya visto en los dos capítulos anteriores, en este capítulo se procederá clasificar ciertos países en conglomerados, de tal forma que, de acuerdo al análisis de los indicadores militares y económicos de cada uno, los países pertenecientes a un conglomerado sean semejantes entre sí y diferentes a los pertenecientes a otros conglomerados. Para la realización del análisis se utilizó el paquete estadístico SPSS.

## 3.1. Elección de variables.

Se eligieron 25 países, los cuales figuran en la mayoría de las listas de países considerados potencias medias o potencias regionales:

Alemania	España	Pakistán
Arabia Saudita	Francia	Reino Unido
Argentina	India	Rusia
Australia	Indonesia	Siria
Brasil	Irán	Sudáfrica
Canadá	Israel	Turquía
China	Italia	Venezuela
Corea del Sur	Japón	
Egipto	México	

### 3.1.1. Indicadores

Ya que en la política internacional, los criterios fundamentales de jerarquización, tradicionalmente se han basado en dos: el poder y el estatus, se tomaron en cuenta indicadores militares y económicos para realizar la jerarquización de los países<sup>1</sup>.

---

<sup>1</sup>Ver Apéndice A

Personal	Sistemas	Logística	Petróleo	Finanzas	Geografía
Población total	Armas aéreas	Puertos	Producción	Presupuesto de defensa	Área(= $km^2$ )
Disposición de mano de obra militar	Helicópteros	Aeropuertos	Consumo	PIB	
Personal activo militar	Armas terrestres	Marina mercante	Reserva	Reservas de div y oro	
Activos de reserva militares	Artillería				
Disponibles para obra militar	Armas nucleares				
Alcanzan edad para milicia	Marina				
Paramilitares					
Fuerza laboral					

### 3.2. Método jerárquico.

Después de realizar varias pruebas con distintas técnicas de conglomeración <sup>2</sup>, se llegó a la conclusión de que el método que satisfacía de mejor manera el interés de este trabajo fue el método de Ward con la distancia euclidiana al cuadrado.

El procedimiento realizado fue:

#### 1. Estandarización de variables.

Debido a que las variables de cada uno de los países son presentadas en escalas distintas, se realizó una estandarización de las variables con análisis descriptivo de puntuaciones Z. Dicha estandarización consistió en restar a cada una de las variables su media y dividir la diferencia por su desviación típica .

#### 2. Conglomeración con algoritmo de Ward.

Posteriormente se realizó el análisis jerárquico por casos, con técnica de conglomeración Ward (vinculación) y distancia euclidiana al cuadrado:

La representación gráfica del análisis se presenta con el dendrograma , en donde se muestra para cada paso los conglomerados que se combinan y los valores de los coeficientes de distancia.

Observando el dendrograma 3.1, podemos ver que una posible solución sería la de 7 conglomerados. (La asignación de los colores para distinguir los conglomerados fue hecha a mano, el programa SPSS no muestra el gráfico así).

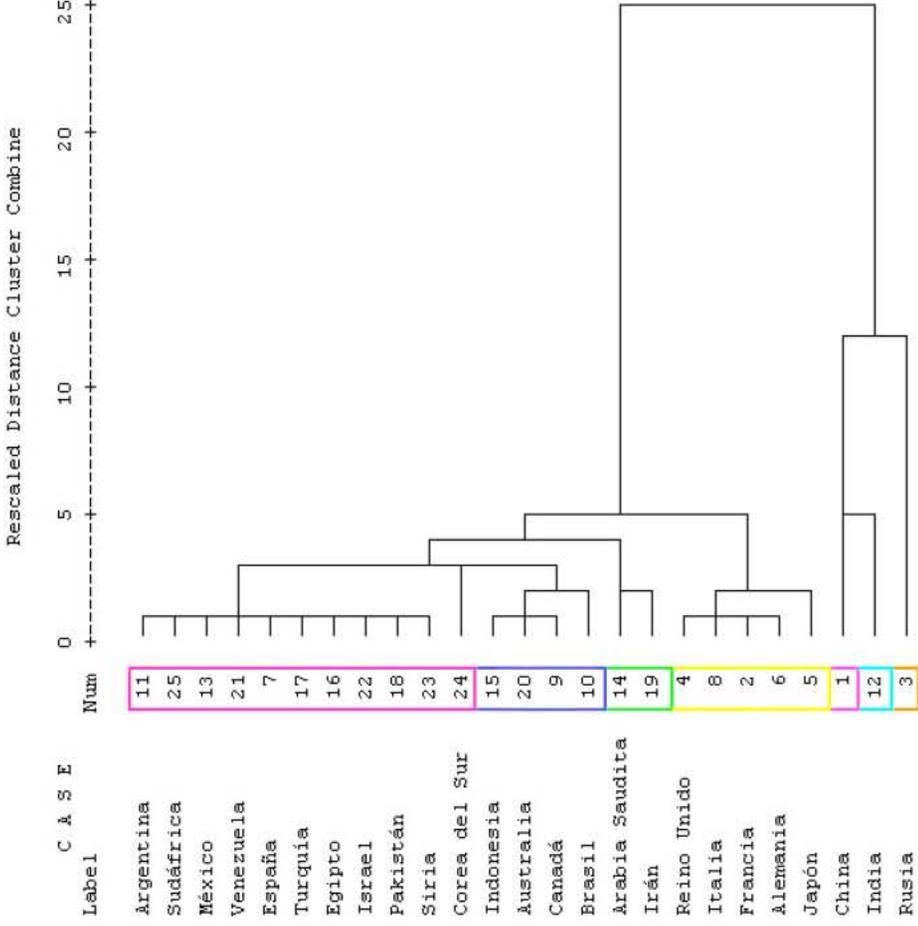
---

<sup>2</sup>Ver apéndice B

**Historial de conglomeración**

Etapa	Conglomerado que se combina		Coeficientes	Etapa en la que el conglomerado aparece por primera vez		Próxima etapa
	Conglomerado 1	Conglomerado 2		Conglomerado 1	Conglomerado 2	
1	11	25	615	0	0	8
2	4	8	1.539	0	0	9
3	2	6	2.732	0	0	9
4	16	22	3.997	0	0	7
5	18	23	5.431	0	0	7
6	7	17	7.068	0	0	12
7	16	18	9.429	4	5	12
8	11	13	12.019	1	0	10
9	2	4	14.870	3	2	15
10	11	21	18.247	8	0	14
11	15	20	23.226	0	0	13
12	7	16	28.246	6	7	14
13	9	15	36.111	0	11	16
14	7	11	44.280	12	10	18
15	2	5	54.513	9	0	21
16	9	10	70.560	13	0	19
17	14	19	88.800	0	0	20
18	7	24	112.067	14	0	19
19	7	9	136.673	18	16	20
20	7	14	172.328	19	17	21
21	2	7	213.433	15	20	24
22	1	12	254.855	0	0	23
23	1	3	355.765	22	0	24
24	1	2	576.000	23	21	0

\* \* \* \* \* H I E R A R C H I C A L C L U S T E R A N A L Y S I S \* \* \* \* \*  
Dendrogram using Ward Method



Cuadro 3.1: Dendrograma. Método de Ward

Para saber cuáles países se han unido, una herramienta muy útil es el gráfico de carámbanos 3.2. Observando la opción de 7 conglomerados, tenemos:

**Diagrama de témpanos horizontal**

Caso	Número de conglomerados																							
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
19:Irán	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
14:Arabia Saudita	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
10:Brasil	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
20:Australia	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
15:Indonesia	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
9:Canadá	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
24:Corea del Sur	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
21:Venezuela	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
13:México	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
25:Sudáfrica	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
11:Argentina	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
23:Siria	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
18:Pakistán	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
22:Israel	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
16:Egipto	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
17:Turquía	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
7:España	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
5:Japón	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
8:Italia	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
4:Reino Unido	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
6:Alemania	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
2:Francia	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
3:Rusia	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
12:India	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
1:China	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X

Cuadro 3.2: Témpanos, método de Ward

### 3.2.1. Interpretación de resultados jerárquicos

Si analizamos conjuntamente el historial de conglomeración y el dendrograma, tenemos que al inicio del proceso se considera que cada país es un conglomerado, y cada uno de ellos adopta la denominación del número del caso correspondiente en el archivo de datos:

$$C_1=\{1\}, C_2=\{2\}, \dots, C_{25}=\{25\}$$

En la primera etapa se combinan los dos países 11 y 25 (“Conglomerado que se combina”: Conglomerado 1=11, Conglomerado 2=25) o, lo que es equivalente, los conglomerados  $C_{11}$  y  $C_{25}$ , y la distancia entre ellos (“Coeficiente”) es 0.615. Si observamos el dendrograma, las líneas correspondientes a los países 11 y 25 son las dos primeras líneas que se cierran. En consecuencia, a partir de la altura del cierre, únicamente quedarán 24 líneas, correspondientes a los 24 conglomerados resultantes después de combinar los conglomerados iniciales  $C_{11}$  y  $C_{25}$  en un único conglomerado, que adoptará el nombre del mínimo número de caso al que contenga, en este caso  $C_{11}$ . Entonces, después de la primera etapa la solución obtenida es:

$$C_1=\{1\}, \dots, C_{11}=\{11,25\}, \dots, C_{24}=\{24\}$$

La próxima vez que el conglomerado  $C_{11}$  se combina con algún otro (“Próxima etapa”) es en la etapa 8, combinándose con el conglomerado  $C_{13}$ .

Si seguimos el proceso desde la etapa 18, observando de dónde proviene cada uno de los conglomerados que se combinan, el número de conglomerados es  $25-18=7$  y la solución obtenida es:

$$C_1=\{1\}, C_2=\{2,4,5,6,8\}, C_3=\{3\}, C_7=\{7,11,13,16,17,18,21,22,23,24,25\}, C_9=\{9,10,15,20\}, C_{12}=\{12\}, C_{14}=\{14,19\}$$

**Conglomerado 1:** Formado únicamente por China.



Figura 3.1: China

País considerado super potencia y en vías posibles de igualar e incluso superar la hegemonía de Estados Unidos. China es país miembro de lo que se conoce en economía internacional como BRIC (siglas para referirse a Brasil, Rusia, India y China); estos países tienen en común una gran población, gran territorio, vastos recursos naturales y crecimientos significativos en el PIB. Es miembro permanente del Consejo de Seguridad de Naciones Unidas con derecho de veto, lo cual le da el privilegio para actuar en asuntos importantes que conciernen al sistema de seguridad mundial. En todos los demás foros de las Naciones Unidas tiene un lugar privilegiado, por ejemplo en el Banco Mundial y Fondo Monetario Internacional, debido al tamaño de las acciones que tiene,

puede nombrar, sin necesidad de coalición internacional, a un director ejecutivo adjunto y un vicepresidente en ambas instituciones. Es miembro del Club de Potencias Nucleares. Debido a su desarrollo nuclear para fines de guerra, cuando ocupó su lugar como miembro permanente del consejo de seguridad de la ONU (1971), obtuvo la legitimidad de poseer esas armas. Es la tercera economía del mundo y la primera en Asia, destina el 4.3% del PIB a gasto militar, y es el 4 país con mayor crecimiento del PIB, estimado a 2008 en 9%.

**Conglomerado 2:** Formado por Francia, Reino Unido, Japón, Alemania e Italia.



Figura 3.2: Francia, Reino Unido, Japón, Alemania e Italia

Países miembros todos del G-7. Son países cuyo peso económico, político y militar son relevantes a escala mundial. La pertenencia a este grupo no se basa en ningún criterio es particular, pero podemos ver con el conglomerado obtenido, que estos países se encuentran dentro de las 10 principales economía mundiales: Japón es la 2, Alemania es la 4, Francia es la 5, Reino Unido es la 6 e Italia es la 7 economía mundial.

**Conglomerado 3:** Formado únicamente por Rusia.



Figura 3.3: Rusia

País potencia mundial y tiene un papel hegemónico en Europa. Rusia es el país más extenso del mundo, con un territorio de 17 millones de km<sup>2</sup> repartidos entre Asia (77%) y Europa (23%). Es el noveno país más poblado del mundo y cuenta con abundantes recursos naturales, sobre todo petróleo siendo el primer país en producción y el segundo en exportación. Militarmente tienen un papel de potencia internacional debido a su ejército y por ser el segundo país con mayor cantidad de armas nucleares después de Estados Unidos.

**Conglomerado 4:** Formado por España, Argentina, México, Egipto, Turquía, Pakistán, Venezuela, Siria, Corea del Sur, Sudáfrica e Israel.

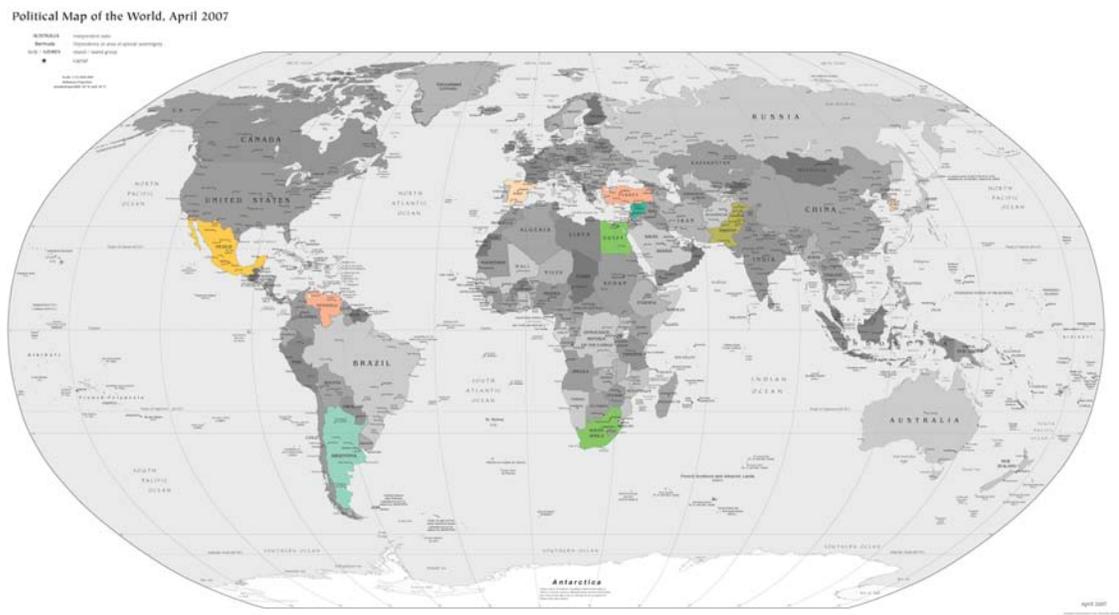


Figura 3.4: España, Argentina, México, Egipto, Turquía, Pakistán, Venezuela, Siria, Corea del Sur, Sudáfrica e Israel

España tiene una importancia diplomática, relativa importancia económica y militar y tiene una amplia influencia en América Latina y el Mediterráneo.

Argentina tiene una importancia regional siendo el 8 país del mundo y el segundo en América Latina.

México tiene una importancia económica por el petróleo, siendo el séptimo productor a nivel mundial, una importancia diplomática y cierta influencia a nivel regional.

Egipto tiene una importancia diplomática, influencia regional en el mundo Árabe. En los últimos años Egipto ha recuperado su papel de líder mundial, además de ser el país más poblado de la región, conserva su liderazgo cultural e intelectual y mantiene contactos con todos los actores implicados en los conflictos árabe-israelíes.

Turquía tiene una importancia militar e influencia regional formando un importante puente económico y geoestratégico con el Mediterráneo y Asia.

Pakistán tiene una importancia militar e influencia regional, tiene buenas relaciones con Irán y Arabia Saudita; coopera con el Consejo de Cooperación para los Estados Árabes del Golfo, coopera militarmente a cambio de ayuda económica. En los últimos años su ejército se ha expandido y modernizado principalmente por la ayuda de Estados Unidos.

Venezuela es un país tiene una limitada influencia regional, tiene importantes relaciones comerciales con China, Rusia e Irán.

Siria tiene un papel de potencia regional en Próximo Oriente.

Corea del Sur tiene una importancia militar e influencia regional, es la 13 economía del mundo. Militarmente destina el 2.7% del PIB a gastos militares y tiene ocupa el segundo lugar de soldados activos en la milicia, después de Corea del Norte.

Sudáfrica es una potencia emergente; tiene influencia en la región, siendo la principal fuerza mediadora en los conflictos.

Israel tiene una importancia política y militar; es el único país del conglomerado que cuenta (no confirmado) con armas nucleares.

España, México y Turquía son países miembros de la Organización para la Cooperación y el Desarrollo Económico (OCDE)<sup>3</sup>

Este conglomerado de países podría considerarse como potencias emergentes regionales.

**Conglomerado 5:** Formado por: Canadá, Brasil, Indonesia y Australia.

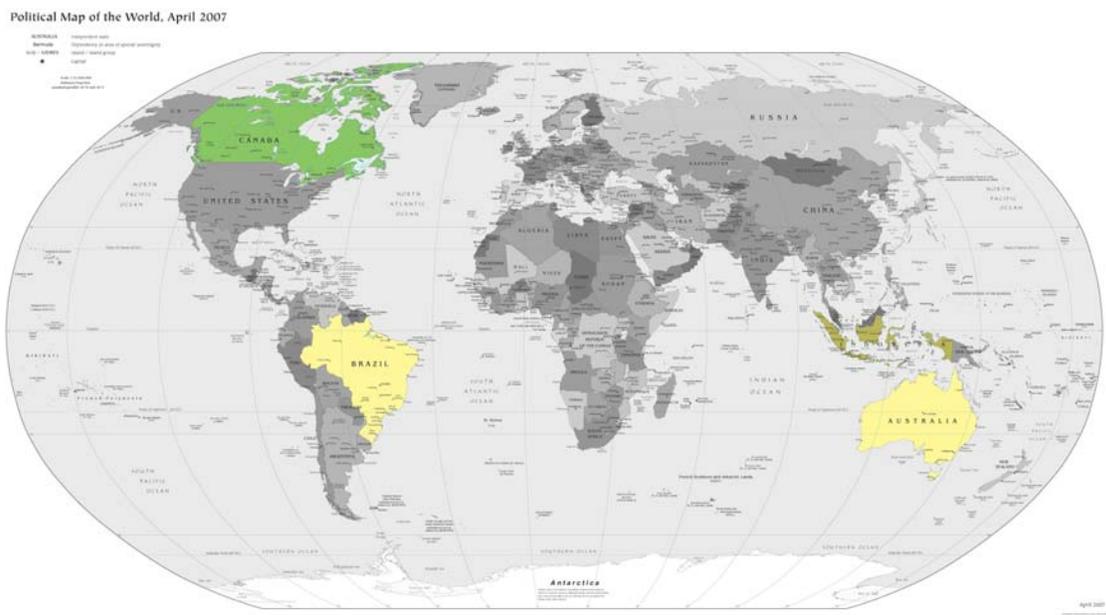


Figura 3.5: Canadá, Brasil, Indonesia y Australia

Canadá ha sido considerada potencia media debido principalmente a las contribuciones pacíficas que ha tenido. Es miembro del G-7 y es la 15 economía mundial. Es el cuarto país en producción de petróleo y el segundo

<sup>3</sup>OCDE, organización con 34 países miembros.

con mayor reserva de éste en su territorio.

Brasil, únicamente por su territorio podría considerarse potencia global. En cuanto a su población y PIB es el más grande en América del Sur y uno de los más importantes a nivel mundial.

Indonesia es una potencia demográfica y debido a su posición entre los estrechos que separan a los océanos Índico y Pacífico es una potencia regional; ocupa el cuarto lugar en densidad demográfica. Miembro del G-20, Indonesia tiene planes de seguir con su desarrollo económico, principalmente en agricultura y convertirse en la segunda potencia de Asia del Este después de China y antes de Japón.

Australia posee una importancia económica diplomática y militar, con una influencia regional en Asia pacífico, teniendo estrechas relaciones con Estados Unidos e Inglaterra. Es miembro fundador del Sistema de Naciones Unidas y ocupa una posición de liderazgo en temas de seguridad, no proliferación, desarme y lucha contra terrorismo.

Una característica de este conglomerado es que todos los países son considerados potencias medias y además, potencias regionales, es decir, dominan en ámbitos políticos, económicos y militares en su región.

**Conglomerado 6:** Formado únicamente por India.



Figura 3.6: India

India es un país con un peso demográfico, económico y militar importantes. Es el séptimo país más grande del mundo y el segundo en densidad demográfica y la quinta economía mundial.

Es miembro del llamado BRIC, un importante socio económico con la UE y un posible candidato a ocupar un lugar en el Consejo de Seguridad de la ONU.

Militarmente tiene el segundo ejército más grande y es potencia nuclear en la región sur de Asia, contando con aproximadamente 100 armas y con la tecnología necesaria para producir más.

**Conglomerado 7:** Formado por Arabia Saudita e Irán.



Figura 3.7: Arabia Saudita e Irán

Arabia Saudita e Irán son países que en los últimos años han ganado terreno en la política internacional y han emergido como potencias regionales, teniendo como características su alta producción de petróleo: Arabia Saudita es el primer productor mundial e Irán el cuarto.

Irán se perfila a ser potencia regional con su ideología de crecimiento territorial en la región de Palestina y Líbano, su desarrollo nuclear pacífico y la influencia que tiene sobre los precios del petróleo y gas. Ejerce una influencia religiosa y política sobre determinados sectores del mundo musulmán.

Por otro lado, Arabia Saudita es la segunda potencia petrolera más grande del mundo, y el país con más reservas de petróleo en su territorio, dándole así un papel importante a nivel internacional y regional en Oriente Medio.

### 3.2.2. Validación del método jerárquico.

Para la validación de la elección del número de conglomerados se realizó un análisis discriminante para analizar si existen diferencias entre los grupos en cuanto a su comportamiento con respecto a las variables consideradas y averiguar en qué sentido se dan dichas diferencias.

El planteamiento seguido en el análisis fue: tenemos  $n=25$  países agrupados en  $q=7$  conglomerados con tamaño  $n_1=1, n_2=5, n_3=1, n_4=11, n_5=4, n_6=1, n_7=2$  y  $Y=24$  variables numéricas observadas sobre dichos países.

El cálculo de las funciones discriminantes fue con Lambda de Wilks que es un estadístico que mide el poder discriminante de un conjunto de variables, viene dado por:

$$L = \frac{|W|}{|W + B|} = \frac{1}{\prod_{i=1}^{\min(q-1)} (1 + \lambda_i)}$$

y toma los valores entre 0 y 1, cuanto más cerca esté del cero, mayor es el poder discriminante de las variables consideradas y cuanto más cercano esté a 1, menor será el poder discriminante de dichas variables.

El número máximo de funciones discriminantes es igual a  $\min\{7 - 1, 24\} = 6$

En las tablas siguientes observamos que las 6 funciones discriminantes tienen un Lambda Wilks pequeña por lo tanto significativas y tienen un poder discriminante elevado dado su valor de correlación canónica.

ANÁLISIS DISCRIMINANTE<sup>4</sup>

## Autovalores

Función	Autovalor	% de varianza	% acumulado	Correlación canónica
1	13758.496 <sup>a</sup>	94.5	94.5	1.000
2	674.421 <sup>a</sup>	4.6	99.1	.999
3	101.404 <sup>a</sup>	.7	99.8	.995
4	17.459 <sup>a</sup>	.1	99.9	.973
5	8.790 <sup>a</sup>	.1	100.0	.948
6	4.891 <sup>a</sup>	.0	100.0	.911

a. Se han empleado las 6 primeras funciones discriminantes canónicas en el análisis.

## Lambda de Wilks

Contraste de las funciones	Lambda de Wilks	Chi-cuadrado	gl	Sig.
1 a la 6	.000	400.839	72	.000
2 a la 6	.000	262.662	55	.000
3 a la 6	.000	168.189	40	.000
4 a la 6	.001	101.070	27	.000
5 a la 6	.017	58.795	16	.000
6	.170	25.715	7	.001

Finalmente tenemos la tabla resumen que nos indica que el 100 % de nuestros casos fueron asignados correctamente:

Resultados de la clasificación<sup>b,c</sup>

	Ward Method	Grupo de pertenencia pronosticado							Total	
		1	2	3	4	5	6	7		
Original	Recuento	1	1	0	0	0	0	0	0	1
		2	0	5	0	0	0	0	0	5
		3	0	0	1	0	0	0	0	1
		4	0	0	0	11	0	0	0	11
		5	0	0	0	0	4	0	0	4
		6	0	0	0	0	0	1	0	1
		7	0	0	0	0	0	0	2	2
	%	1	100.0	.0	.0	.0	.0	.0	.0	100.0
		2	.0	100.0	.0	.0	.0	.0	.0	100.0
		3	.0	.0	100.0	.0	.0	.0	.0	100.0
		4	.0	.0	.0	100.0	.0	.0	.0	100.0
		5	.0	.0	.0	.0	100.0	.0	.0	100.0
		6	.0	.0	.0	.0	.0	100.0	.0	100.0
		7	.0	.0	.0	.0	.0	.0	100.0	100.0
Validación cruzada <sup>a</sup>	Recuento	1	0	0	0	0	0	1	0	1
		2	0	5	0	0	0	0	0	5
		3	0	0	0	0	0	1	0	1
		4	0	0	0	11	0	0	0	11
		5	0	0	0	0	4	0	0	4
		6	0	0	1	0	0	0	0	1
		7	0	1	0	0	0	1	0	2
	%	1	.0	.0	.0	.0	.0	100.0	.0	100.0
		2	.0	100.0	.0	.0	.0	.0	.0	100.0
		3	.0	.0	.0	.0	100.0	.0	.0	100.0
		4	.0	.0	.0	100.0	.0	.0	.0	100.0
		5	.0	.0	.0	.0	100.0	.0	.0	100.0
		6	.0	.0	100.0	.0	.0	.0	.0	100.0
		7	.0	50.0	.0	.0	50.0	.0	.0	100.0

a. La validación cruzada sólo se aplica a los casos del análisis. En la validación cruzada, cada caso se clasifica mediante las funciones derivadas a partir del resto de los casos.

b. Clasificados correctamente el 100.0% de los casos agrupados originales.

c. Clasificados correctamente el 80.0% de los casos agrupados validados mediante validación cruzada.

<sup>4</sup>Análisis realizado en SPSS

### 3.3. Método de k-medias.

De manera complementaria, se realizó el análisis de k-medias <sup>5</sup> con una selección a priori de 7 conglomerados para comparar los resultados obtenidos en el análisis jerárquico. Cabe señalar que las diferencias encontradas se deben principalmente a la medida de distancia utilizada, ya que el método de k-medias realiza la clasificación con la distancia euclidiana, a diferencia del método de Ward que utiliza la distancia euclidiana al cuadrado, por otro lado, los resultados obtenidos por el método de k-medias varían según el orden en que son seleccionados los centros de los conglomerados, y por lo tanto también depende del orden en que son introducidas las variables; en este caso no hicimos ningún cambio en el orden de éstas.

#### **Estandarización de variables.**

Iniciamos el proceso estandarizando las variables con puntuaciones Z, evitando así que alguna variable, debido a las diferentes escalas en que son medidas, tenga mayor peso en el análisis, en consecuencia, las nuevas variables tendrán media 0 y desviación típica 1.

#### **Selección de los centros iniciales de los conglomerados.**

En este caso, los centros de los conglomerados son desconocidos, por lo tanto, serán estimados iterativamente y se clasificarán los países con arreglo a los centros estimados. Se hizo la estimación de los centros según los resultados del método jerárquico.

#### **Proceso de clasificación.**

Una vez seleccionados los centros de los conglomerados, cada país es asignado al conglomerado cuyo centro se encuentra más próximo y comienza un proceso de ubicación iterativa de los centros. En la primera iteración se reasignan los casos por su distancia al nuevo centro y, tras la reasignación se vuelve a actualizar el valor del centro. En nuestro caso en la tabla Historial de conglomeración tenemos 2 iteraciones, es decir, en la cuarta iteración ya no hay un desplazamiento de los centros de conglomerados.

En la tabla *Centros de conglomerados iniciales* podemos ver que el primer centro corresponde a un país con un valor grande de territorio (1.56050), análogamente con los centros siguientes.

En la tabla *Centros de conglomerados finales* se muestran los centros de los conglomerados tras el proceso de actualización iterativa, podemos observar que no hubo cambio en los centros de los conglomerados 1, 3 y 4, lo cual nos podría indicar que son conglomerados unitarios.

Por último tenemos la tabla de Pertenencia a los conglomerados, en donde tenemos:

---

<sup>5</sup>Ver Apéndice C

## Pertenenca a los conglomerados

Número de caso	País	Conglomerado	Distancia
1	China	1	.000
2	Francia	5	1.356
3	Rusia	3	.000
4	Reino Unido	5	1.261
5	Japón	5	2.861
6	Alemania	5	1.414
7	España	6	1.306
8	Italia	5	1.259
9	Canadá	2	2.474
10	Brasil	2	3.424
11	Argentina	6	1.914
12	India	4	.000
13	México	2	1.526
14	Arabia Saudita	7	3.003
15	Indonesia	2	2.581
16	Egipto	6	1.168
17	Turquía	6	1.809
18	Pakistán	6	1.747
19	Irán	7	3.003
20	Australia	2	2.229
21	Venezuela	6	1.892
22	Israel	6	1.346
23	Siria	6	1.568
24	Corea del Sur	6	4.568
25	Sudáfrica	6	1.424

Como se mencionó anteriormente, encontramos diferencias en la agrupación de países, en este caso México. Esta variación en el resultado se debe debido a la diferencia en la media de distancia utilizada, recordando que en la clasificación jerárquica se hace uso del método de Ward (con distancia euclidiana cuadrada); a pesar de esto, podemos decir que los resultados son constantes.

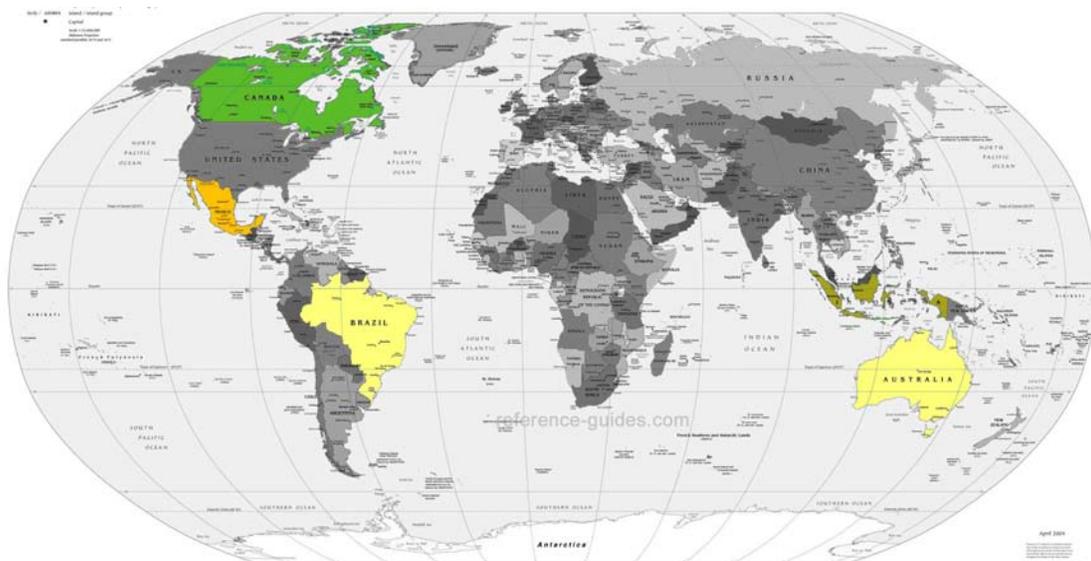


Figura 3.8: Conglomerado. Ubicación de México

**El caso de México.**

En este caso México se ubicó en el mismo conglomerado que Brasil, siendo estos dos países, dentro del conjunto de Estados Latinoamericanos, los únicos que cuentan con las capacidades para ser reconocidos como potencias regionales-mediadoras.

México puede desplegar su:

- **Geoeconomía.** Es este aspecto debe remarcarse especialmente el petróleo, cuya exportación ha permitido mantener unos niveles aceptables de ingresos y ha aumentado considerablemente la importancia económica y estratégica del país a nivel regional, e incluso internacional.[12]
- **Geopolítica** Por una parte se encuentra al lado de la superpotencia mundial (EEUU), con la que comparte una frontera de más de 3,000km; por la otra, es vecino de los pequeños países de Centroamérica y el Caribe, quienes tradicionalmente han advertido el peso económico, demográfico y cultural de México, aunque no siempre haya existido una clara voluntad de proyectar una regional.[12]
- **Geoestratégica** México se ha caracterizado por un bajo desarrollo militar acorde a sus dimensiones geográficas y demográficas, en cierta medida debido a la vecindad con los EEUU. Así México ha optado por la utilización de instrumentos de carácter político y diplomático que implican costos bajos, y se ha negado a jugar un papel más grande en los esquemas militaristas de seguridad internacional y regional [12].

# Conclusiones

---

Se analizaron 25 países considerados potencias medias, y se clasificaron dados unos índices demográficos, geográficos, económicos y militares. En primer lugar se realizó una clasificación jerárquica con el método de Ward y de manera adicional se realizó una clasificación no jerárquica mediante la técnica de k-medias, obteniendo resultados consistentes, en vista de que se presentó la variación de un solo país (México) en la asignación de conglomerados. La variación se debe a la diferencia en criterios al realizar la clasificación por método jerárquico y no jerárquico.

Se eligió la solución de 7 conglomerados, lo cual satisfizo el interés de la tesina, al obtener como conglomerados unitarios a países considerados superpotencias (India, Rusia y China). En un segundo grupo, grandes potencias (Francia, Reino Unido, Japón, Alemania e Italia). En un tercer grupo a potencias medias, entre estos países se distinguen claras potencias regionales, como es el caso de Canadá, Brasil, Arabia Saudita e Irán.

En el caso de México, en la clasificación jerárquica, se ubicó con países considerados potencias emergentes regionales y consideramos que actualmente México guarda una mayor similitud con los países formados en este conglomerado 4 fig.3.4 y que a futuro se puede proyectar en una ubicación junto con las potencias medias obtenidas en el método no jerárquico y efectivamente junto con Brasil, sean los dos países con hegemonía regional y subhegemonía mundial. 3.3.

Se puede concluir que la aplicación del análisis de conglomerados resulta útil para fines exploratorios de clasificación de países, dejando a criterio del investigador la elección de las variables, así como el método y número de conglomerados óptimos a considerar. En un trabajo futuro se pueden añadir más indicadores, por ejemplo, el índice de desarrollo humano y comparar los resultados. Otra posible mejora sería la clasificación de países según el PIB per capita y comparar los resultados de organismos internacionales con los obtenidos en la investigación.



# Bibliografía

---

- [1] Aron. *Paz y guerra entre las naciones*. Ed. Alianza. Madrid, España, 1985.
- [2] Barbé, Esther. *Relaciones Internacionales*. Ed. Tecnos. Madrid, España, 1995
- [3] Bisquerra
- [4] *BMDP Statistical Software*. W.J.Dixon. Universidad de California, 1983.
- [5] Borrachi, Raúl. *Aplicación de análisis de conglomerados y redes neuronales artificiales para la clasificación y selección de candidatos a residencias médicas*. Universidad de Buenos Aires, Argentina, 2005.
- [6] Dillon, R.William. *Multivariate Analysis*. Wiley series in probability and mathematical statistics. Estados Unidos, 1984.
- [7] Evettitt. *Cluster analysis*. Ed. Wiley. Londres, 2001.
- [8] González, Bueno Antonio. *Historia de la Ciencia y de la Técnica*. Ed. Akal. Madrid, España,1998.
- [9] González, Guadalupe. *Incertidumbre de una potencia media regional: las nuevas dimensiones de la política exterior mexicana*. CIDE. México
- [10] Hartigan, J.A. *Clustering algorithms*. New York: John Wiley & Sons, Inc. 1975
- [11] Holbraad, Carsten. *Las potencias medias en la política internacional*. Fondo de Cultura Económica. México, 1989.
- [12] Jordi Palou. *El concepto de potencia media. Los casos de España y México*. Fundación CIDOB
- [13] Kasielke, E. *Cuestionario para el Registro de Alteraciones Físicas de Origen Psíquico (BEB)*. Universidad de Humboldt. Berlín, 1982.
- [14] Keohane. *Después de la hegemonía. Cooperación y discordia en la política económica mundial*. Grupo Editorial Buenos Aires. Argentina, 1988.
- [15] Merle. *Sociología de las relaciones internacionales*. Ed. Alianza. Madrid,1991.
- [16] Morgenthau.*Escritos sobre política internacional*. Tecnos, Madrid, España, 1990.
- [17] Paulo Gilberto F.*Brasil como potencia media*. Sao-Paulo, 1998
- [18] Pearson, Frederics. *Relaciones Internacionales. Situación global en el siglo XX*. Ed.McGraw Hill. Bogotá, Colombia.2003.
- [19] Rocha Valencia, Alberto. *Sistema político internacional y potencias regionales-mediadoras: los casos de Brasil y México*. UDG. México.
- [20] Santoro, C.M.¿A dónde va la política exterior italiana? *Cinco hipótesis de una potencia media*. Italia: Relazioni Internazionali,1988.
- [21] Truyo y Serra. *La sociedad internacional*. Ed. Alianza. Madrid, España, 1974.
- [22] Waltz. *Teoría de la política internacional*. Grupo Editorial Latinoamericano. Buenos Aires, Argentina, 1988.

- [23] <https://www.cia.gov/library/publications/the-world-factbook/>
- [24] <http://www.globalfirepower.com>
- [25] <http://www.indexmundi.com/es>
- [26] <http://www.militarypower.com.br/index3.htm>
- [27] <http://www.nationmaster.com>
- [28] <http://revistas.um.es/analesps>

---

## Apéndice A

# Indicadores

---

### Personal

**Población.** Basado en estadísticas de los censos de población, estadísticas de sistemas de registros o encuestas por muestreo relativas al pasado reciente.

**Disposición de mano de obra militar.** Número de hombre y mujeres que caen dentro del rango de edad apto para el servicio militar de un país (definido de 16 a 49 años) y se supone que cada individuo está en condiciones de servir. No se toma en cuenta cualquier tipo de entrenamiento militar, paramilitar o de seguridad de ningún tipo, únicamente se refleja la cantidad total de mano de obra de una nación podría participar en caso de guerra si es necesario.

**Personal militar activo.** Número de hombres y mujeres activos en el servicio militar.

**Activos de reserva.** Número de hombre y mujeres que operan en una base de fuerza militar principal.

**Aptos para el servicio militar.** Número de hombres y mujeres que caen dentro del rango de edad militar de un país (16 a 49 años) y que no son descalificados por cuestiones de salud; con este indicador se puede explicar la situación de salud de un país.

**Alcanzan edad para milicia.** Número de hombres y mujeres que están alcanzando la edad militar cada año, afectan directamente a la mano de obra militar posible.

**Paramilitares.** Se consideran aquellos hombres y mujeres cuya función es similar a la de las fuerzas militares, pero no forman parte de manera oficial de los servicios de milicia.

**Fuerza laboral.** Se muestra la fuerza laboral total; como se vio en la Segunda Guerra Mundial, la capacidad de fabricación de una nación es tan importante como el poder general para hacer la guerra. Con una mano de obra de gran potencial, el material de guerra se puede producir en grandes cantidades y menor tiempo.

### Sistemas

**Armas aéreas.** Número (aproximado) de armas aéreas, se incluyen las de sistema fijo como aéreas.

**Helicópteros.** Número total de helicópteros, se incluyen los de transporte, scout y ataque.

**Armas terrestres.** Número total de armas terrestres.

**Artillería.** Número total de armas de artillería, entendiendo como éstas, aquellas máquinas de guerra que operan con la proyección de municiones, más allá de la gama de armas personales.

**Armas nucleares.** Total de armas nucleares (cifras aproximadas).

**Marina.** Total de armas de la marina.

## Logística

**Puertos.** Se enumeran los principales puertos y terminales, principalmente con base en la cantidad de tonelaje de mercancías transportadas a través de las instalaciones anualmente.

**Aeropuertos.** Número total de aeropuertos y aeródromos reconocibles desde el aire.

**Marina mercante.** Se puede definir como todos los buques dedicados al transporte de mercancías, o todos los buques comerciales.

## Petróleo.

**Producción.** Número total de barriles producidos por día.

**Consumo.** Número total de barriles consumidos por día.

**Reserva.** Stock de reservas probadas de petróleo crudo en barriles.

## Finanzas

**Presupuesto de defensa.** Cantidad de gasto en programas de defensa.

**Producto Interno Bruto (PIB).** Muestra el valor de todos los bienes y servicios finales producidos por una nación en un año determinado. (Valuado a precios vigentes en EEUU).

**Reservas de divisas y oro.** Muestra el valor en dólares de la población de todos los activos financieros disponibles para la autoridad monetaria central para su uso en una reunión de balance de pagos que el país necesita a partir de la fecha final del período especificado. Esta categoría incluye no sólo moneda extranjera y oro, sino también las explotaciones de un país de derechos especiales de giro en el Fondo Monetario Internacional, y su posición de reserva en el Fondo.

## Geografía

**Área.** Territorio de un país en km<sup>2</sup>.

## PERSONAL

País	Población (millones)	Disposición de mano de obra militar (millones)	Personal activo militar (miles)	Activos de reserva (miles)	Aptos para servicio militar (millones)	Alcanzan la edad para la milicia (miles)	Paramilitares (miles)	Fuerza Laboral (millones)
China	1330.04	729.32	2255	800	609.27	20470.41	3969.00	807.30
Francia	64.06	29.03	225	419	23.96	783.79	101.40	27.97
Rusia	140.70	73.24	1245	2400	50.25	1602.67	359.10	75.70
Reino Unido	65.10	28.86	195	234	23.74	784.52	0.00	31.23
Japón	127.29	54.68	239	58	45.10	1212.32	12.25	66.50
Alemania	82.37	38.14	250	359	30.96	863.77	40.00	43.60
España	40.49	19.80	177	329	16.22	395.00	72.60	22.85
Italia	58.15	27.04	240	65	21.97	564.31	328.37	25.11
Canadá	33.21	15.89	62	25	13.06	442.99	9.35	18.22
Brasil	196.34	104.83	287	1115	83.37	3275.15	385.60	93.65
Argentina	40.48	19.92	71	0	16.72	681.25	47.24	16.27
India	1148.00	584.14	1325	1155	467.80	22229.37	1293.30	523.50
México	109.96	57.15	192	300	47.07	2183.77	25.00	45.32
Arabia Saudita	28.15	14.93	124	20	12.92	533.70	15.00	6.74
Indonesia	237.51	125.53	316	400	104.50	4291.70	207.00	112.00
Egipto	81.71	41.65	450	254	35.56	1611.89	405.00	24.60
Turquía	71.89	39.65	514	380	33.44	1298.98	148.70	24.06
Pakistán	172.80	82.75	650	528	63.82	3998.98	302.00	50.58
Irán	65.88	39.85	545	350	34.34	1494.32	11390.00	24.35
Australia	21.01	9.87	54	20	8.16	282.45	0.00	11.25
Venezuela	26.41	13.45	100	100	11.05	549.43	600.00	12.59
Israel	7.11	3.35	187	408	2.84	118.13	8.05	2.96
Siria	19.75	10.22	296	133	8.46	418.84	108.00	5.59
Corea del Sur	48.38	26.72	687	4500	21.97	696.52	22.00	24.22
Sudáfrica	49.05	23.12	60	0	11.51	1051.88	0.00	17.79

*Datos a 2008.*

*Fuente: <https://www.cia.gov/library/publications/the-world-factbook/>  
<http://www.globalfirepower.com>*

## SISTEMAS

País	Armas aéreas	Helicópteros	Armas terrestres	Artillería	Armas nucleares	Total marina
China	1,900	491	31,300	14,000	410	1,299
Francia	1,023	892	8,536	787	350	235
Rusia	3,888	2,625	79,985	13,585	11,000	735
Reino Unido	1,891	779	5,121	455	200	217
Japón	1,957	745	2,040	5,760	0	260
Alemania	350	748	5,699	1,364	0	207
España	691	311	2,869	1,330	0	139
Italia	1,594	716	3,355	643	0	164
Canadá	399	168	2,194	362	0	65
Brasil	1,272	372	1,676	655	0	151
Argentina	632	156	0	0	0	64
India	1,007	240	10,340	4,175	100	241
México	365	273	1,171	130	0	160
Arabia Saudita	453	140	5,695	200	0	114
Indonesia	313	194	2,122	293	0	164
Egipto	1,230	243	9,357	2,393	0	157
Turquía	1,199	336	6,672	685	0	271
Pakistán	710	198	3,919	3,952	60	64
Irán	84	84	5,449	2,010	0	216
Australia	466	309	705	270	0	93
Venezuela	515	92	649	92	0	113
Israel	1,230	386	14,200	2,783	500	34
Siria	607	235	12,350	630	0	31
Corea del Sur	538	502	8,325	4,000	0	210
Sudáfrica	425	132	0	0	0	39

*Datos a 2008.*

*Fuente:*

<https://www.cia.gov/library/publications/the-world-factbook/>

<http://www.globalfirepower.com>

## LOGÍSTICA

<b>País</b>	<b>Puertos</b>	<b>Aeropuertos</b>	<b>Marina mercante</b>
China	8	467	1,822
Francia	9	476	138
Rusia	8	1,260	1,074
Reino Unido	10	483	518
Japón	10	176	683
Alemania	8	550	391
España	7	154	158
Italia	8	132	609
Canadá	9	1,343	175
Brasil	7	4,263	136
Argentina	7	1,272	46
India	9	346	501
México	7	1,834	58
Arabia Saudita	4	213	62
Indonesia	10	652	971
Egipto	6	88	67
Turquía	6	117	602
Pakistán	2	146	14
Irán	3	331	92
Australia	11	461	52
Venezuela	4	390	62
Israel	4	53	11
Siria	2	90	77
Corea del Sur	4	116	812
Sudáfrica	5	728	3

*Datos a 2008.*

*Fuente: <https://www.cia.gov/library/publications/the-world-factbook/>*

## PETRÓLEO (millones, barriles/día)

<b>País</b>	<b>Producción</b>	<b>Consumo</b>	<b>Reserva</b>
China	3.97300	7.8500	16,000
Francia	0.07080	1.9860	103
Rusia	9.79000	2.9000	60,000
Reino Unido	1.58400	1.7100	3,410
Japón	0.13310	4.7850	44
Alemania	0.15080	2.5690	276
España	0.02813	1.5620	150
Italia	0.16220	1.6390	407
Canadá	3.35000	2.2600	178,100
Brasil	2.42200	2.5200	12,620
Argentina	0.79230	0.6100	2,616
India	0.88350	2.9400	5,625
México	3.18600	2.1280	10,500
Arabia Saudita	10.78000	2.3800	266,700
Indonesia	1.05100	1.5640	3,990
Egipto	0.63060	0.6970	3,700
Turquía	0.04612	0.6755	300
Pakistán	0.06187	0.3830	339
Irán	4.17400	1.7550	136,200
Australia	0.58640	0.9537	1,500
Venezuela	2.64300	0.7600	99,380
Israel	0.00525	0.2350	2
Siria	0.42610	0.2560	2,500
Corea del Sur	0.01705	2.1300	0
Sudáfrica	0.19500	0.5830	15

*Datos a 2008.*

*Fuente: <https://www.cia.gov/library/publications/the-world-factbook/>*

## FINANZAS (millones de dólares)

País	Presupuesto de defensa	PIB (paridad poder adquisitivo)	Reservas de div y oro
China	59,000	7,992,000	1,955,000
Francia	61,571	2,133,000	102,900
Rusia	43,200	2,271,000	427,100
Reino Unido	53,149	2,236,000	52,980
Japón	44,300	4,340,000	1,011,000
Alemania	45,930	2,925,000	138,000
España	17,700	1,402,000	20,250
Italia	40,050	1,827,000	105,300
Canadá	19,500	1,303,000	43,870
Brasil	24,400	1,998,000	193,800
Argentina	3,000	575,200	46,370
India	32,350	3,304,000	254,000
México	6,070	1,567,000	95,300
Arabia Saudita	31,255	577,900	30,590
Indonesia	4,740	916,700	51,640
Egipto	2,500	444,800	33,850
Turquía	30,936	903,900	73,660
Pakistán	4,260	431,200	8,903
Irán	6,300	843,700	96,560
Australia	22,000	802,900	32,920
Venezuela	7,700	356,300	42,300
Israel	18,700	203,400	42,510
Siria	935	99,060	6,765
Corea del Sur	25,500	1,206,000	262,200
Sudáfrica	3,700	492,200	34,070

*Datos a 2008.*

*Fuente: <https://www.cia.gov/library/publications/the-world-factbook/>*

## GEOGRAFÍA

<b>País</b>	<b>Territorio (<math>km^2</math>)</b>
China	9,596.96
Francia	551.50
Rusia	17,075.20
Reino Unido	24.48
Japón	377.84
Alemania	357.02
España	504.78
Italia	30.12
Canadá	9,984.67
Brasil	8,511.97
Argentina	2,766.89
India	3,287.59
México	1,972.55
Arabia Saudita	2,149.69
Indonesia	1,919.44
Egipto	1,001.45
Turquía	78.06
Pakistán	80.39
Irán	1,648.00
Australia	7,686.85
Venezuela	91.21
Israel	2.08
Siria	18.52
Corea del Sur	99.72
Sudáfrica	1,219.09

*Datos a 2008.*

*Fuente: <https://www.cia.gov/library/publications/the-world-factbook/>*

# Apéndice B

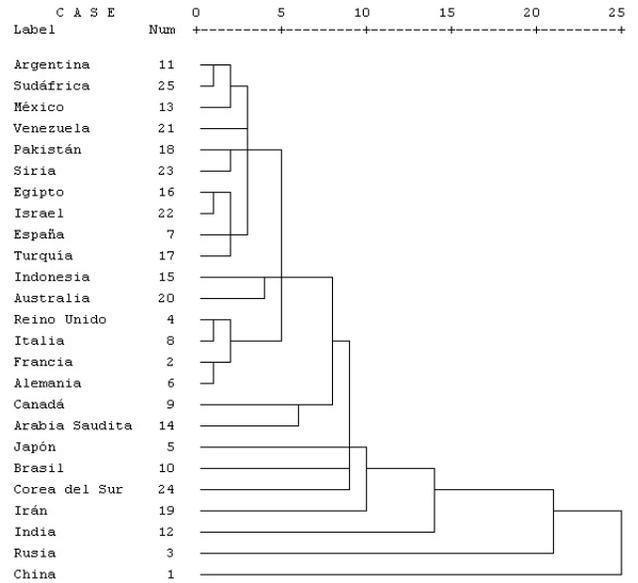
# Métodos jerárquicos

Promedio entre grupos

\*\*\*\*\*HIERARCHICAL CLUSTER ANALYSIS\*\*\*\*\*

Dendrogram using Average Linkage (Between Groups)

Rescaled Distance Cluster Combine



Historial de conglomeración

Etapa	Conglomerado que se combina		Coeficientes	Etapa en la que el conglomerado aparece por primera vez		Próxima etapa
	Conglomerado 1	Conglomerado 2		Conglomerado 1	Conglomerado 2	
1	11	25	1.109	0	0	8
2	4	8	1.359	0	0	7
3	2	6	1.545	0	0	7
4	16	22	1.591	0	0	6
5	18	23	1.693	0	0	11
6	7	16	1.805	0	4	9
7	2	4	1.964	3	2	15
8	11	13	2.029	1	0	10
9	7	17	2.044	6	0	11
10	11	21	2.330	8	0	12
11	7	18	2.346	9	5	12
12	7	11	2.589	11	10	14
13	15	20	3.156	0	0	14
14	7	15	3.406	12	13	15
15	2	7	3.651	7	14	17
16	9	14	4.212	0	0	17
17	2	9	5.098	15	16	18
18	2	5	5.343	17	0	19
19	2	10	5.499	18	0	20
20	2	24	5.705	19	0	21
21	2	19	6.279	20	0	22
22	2	12	8.378	21	0	23
23	2	3	11.929	22	0	24
24	1	2	14.184	0	23	0

Diagrama de témpanos vertical

Número de conglomerados	Caso																								
	3 Rusia	12 India	19 Italia	Sur	10 Brasil	5 Japón	Saudia	9 Canadá	20 Australia	15 Indonesia	21 Venezuela	13 México	25 Sudáfrica	11 Argentina	23 Snp	18 Pakistán	17 Turquía	22 Israel	16 Egipto	7 España	8 Italia	4 Reino Unido	6 Alemania	2 Francia	1 China
1	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
2	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
3	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
4	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
5	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
6	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
7	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
8	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
9	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
10	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
11	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
12	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
13	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
14	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
15	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
16	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
17	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
18	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
19	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
20	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
21	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
22	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
23	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
24	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X

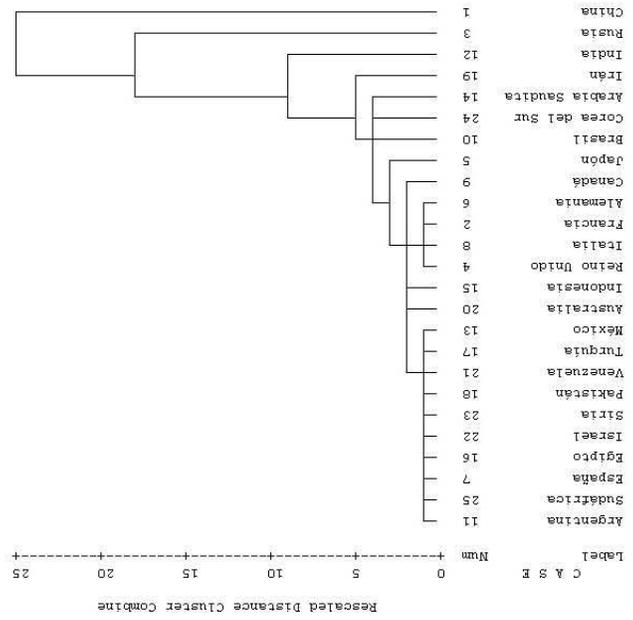






\*\*\*\*\* HIERARQUICAL CLUSTER ANALYSIS \*\*\*\*\*

Dendrogram using Centroid Method



Etapa	Conglomerado que se combina		Etapa en la que el conglomerado aparece por primera vez	Conglomerado	Proxima etapa
	Coefficientes	Conglomerado			
1	11	25	1.230	0	8
2	4	8	1.848	0	7
3	6	2	2.386	0	7
4	16	22	2.530	0	5
5	16	23	2.656	0	6
6	18	4	2.699	0	9
7	18	3	2.851	2	15
8	11	11	3.058	0	9
9	16	16	3.204	8	10
10	7	16	3.204	6	20
11	7	9	3.591	0	11
12	7	17	4.802	0	12
13	13	13	6.501	0	13
14	20	12	8.883	0	14
15	7	7	9.026	13	15
16	7	7	9.024	14	16
17	2	5	14.445	0	17
18	2	9	14.445	15	18
19	2	5	22.861	0	18
20	2	10	23.488	0	19
21	2	17	24.808	0	20
22	2	18	24.808	0	20
23	2	19	29.166	0	21
24	2	20	31.211	0	22
25	2	21	60.885	0	23
26	2	22	130.824	0	24
27	1	2	186.413	0	23

Diagrama de tiempos vertical

Caso	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	
3 Rusia	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
12 India	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
19 Iran	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
14 Arabia Saudita	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
24 Corea del Sur	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
10 Brasil	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
5 Japon	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
9 Canada	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
15 Indonesia	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
20 Australia	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
13 Mexico	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
17 Turquía	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
21 Venezuela	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
18 Pakistan	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
23 Siria	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
22 Israel	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
16 Egipto	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
25 Sudánica	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
11 Argentina	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
7 España	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
8 Italia	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
4 Reino Unido	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
6 Alemania	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
21 Francia	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
1 China	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X



# Análisis de K-medias

## Centro iniciales.

Los centros iniciales de los conglomerados fueron elegidos de manera aleatoria. La forma en que esta tabla se lee es, por ejemplo, en el conglomerado 1, observando la variable “Puntua (Territorio)” (puntua hace referencia a que es una variable estandarizada), tenemos un valor de 1.56050 será el valor del centro de conglomerado con el que se calculará la distancia euclidiana para agrupar a los países que sean similares a éste. La interpretación es análoga para los otros conglomerados y todas las variables.

Centros iniciales de los conglomerados

	Conglomerado						
	1	2	3	4	5	6	7
Puntua(Territorio)	1.56050	.96660	3.28790	.10310	-.59441	-.49180	-.21770
Puntua(Población)	3.53750	-.14830	-.09130	2.98200	-.27838	-.33530	-.37720
Puntua: DISPOSICION DE MANO DE OBRA MILITAR	3.66400	-.14010	-.08740	2.83380	-.30291	-.33060	-.34960
Puntua: Personal activo militar	3.61750	-.49570	1.61570	1.77420	-.39648	-.24220	-.18900
Puntua: Activos de reserva	.23330	-.18990	1.88500	.59980	-.35834	.05790	-.40160
Puntua: Disponibles para obra militar	3.73280	-.13650	-.15060	2.75000	-.29717	-.33000	-.33540
Puntua: Alcanzan la edad para milicia	3.10110	-.14110	-.22400	3.41110	-.35805	-.29800	-.32770
Puntua(Paramilitares)	1.35130	-.27390	-.18510	.21250	-.29688	-.27060	2.08910
Puntua: Fuerza Laboral	3.97960	-.14190	-.04880	2.41700	-.25151	-.34200	-.38000
Puntua: Armas aéreas (Bases)	1.12350	-.46530	3.57660	.02150	.46082	-.30770	-.88980
Puntua(Helicópteros)	.07110	-.37950	4.24760	-.42020	.62886	-.38030	-.67070
Puntua: Armas Terrestres	1.38120	-.44960	4.38980	.08600	-.24712	-.21870	-.20870
Puntua(Artillería)	3.07540	-.53850	2.96520	.46560	-.16479	-.25710	-.34990
Puntua: armas nucleares	-.04330	-.23040	4.78990	-.18480	-.18019	-.20720	-.23040
Puntua: Total marina	4.09390	-.37660	1.95850	.08810	-.00424	-.38310	-.19960
Puntua(Puertos)	.48810	.96480	.48810	.86950	.86945	-.69060	-1.22790
Puntua(Aeropuertos)	-.20160	1.16720	.69340	-.33820	-.31856	-.21690	-.42170
Puntua: Marina mercante	3.27830	-.07170	1.59480	.30530	.23055	-.43150	-.64900
Puntua: Petróleo producción	.72920	-.01160	2.76140	-.35010	-.51197	-.40370	1.95330
Puntua: Petróleo consumo	3.64190	-.05450	.60530	.62990	.38313	-.61490	.09460
Puntua: Petróleo reserva	-.23940	.24970	.41170	-.39290	-.46363	-.31540	2.50480
Puntua: Presupuesto de Defensa	1.81200	-.34980	.98580	.41840	1.28909	-.69810	-.29140
Puntua: PIB (Paridad poder adquisitivo)	3.75530	-.23130	.36980	.98110	.61907	-.56090	-.55340
Puntua: Reservas de div y oro	4.17280	-.30050	.52650	.11340	.18033	-.34820	-.34100

### Historial de iteraciones

Las iteraciones necesarias para que ya no existiera movimientos en los centros de conglomerados que optimicen la clasificación, fueron 2:

Historial de iteraciones(a)

Iteración	Cambio en los centros de los conglomerados						
	1	2	3	4	5	6	7
1	.000	.382	.000	.000	6.67E-006	.236	.000
2	.000	.000	.000	.000	.000	.000	.000

a Se ha logrado la convergencia debido a que los centros de los conglomerados no presentan ningún cambio o éste es pequeño. El cambio máximo de coordenadas absolutas para cualquier centro es de .000. La iteración actual es 2. La distancia mínima entre los centros iniciales es de 2.896.

### Pertenencia a los conglomerados.

Aquí podemos ver la pertenencia de cada país a cada conglomerado, es decir, China pertenece al conglomerado 1; Canadá, Brasil, México, Indonesia y Australia, pertenecen al conglomerado 2. Análogamente para los siguientes casos.

Pertenencia a los conglomerados

Número de caso	País	Conglomerado	Distancia
1	China	1	.000
2	Francia	5	1.356
3	Rusia	3	.000
4	Reino Unido	5	1.261
5	Japón	5	2.861
6	Alemania	5	1.414
7	España	6	1.306
8	Italia	5	1.259
9	Canadá	2	2.474
10	Brasil	2	3.424
11	Argentina	6	1.914
12	India	4	.000
13	México	2	1.526
14	Arabia Saudita	7	3.003
15	Indonesia	2	2.581
16	Egipto	6	1.168
17	Turquía	6	1.809
18	Pakistán	6	1.747
19	Irán	7	3.003
20	Australia	2	2.229
21	Venezuela	6	1.892
22	Israel	6	1.346
23	Siria	6	1.568
24	Corea del Sur	6	4.568
25	Sudáfrica	6	1.424

### Centro finales de conglomerados

Aquí se presentan los centros finales, por ejemplo el conglomerado 1 no tuvo un cambio entre su centro inicial y centro final, lo cual nos indica que este conglomerado está formado por un solo país, que revisando la pertenencia de los conglomerados, es China.

## Centros de los conglomerados finales

	Conglomerado						
	1	2	3	4	5	6	7
Puntua(Territorio)	1.56050	.73310	3.28794	.10306	-59441	-.52095	-21774
Puntua(Población)	3.63746	-.15569	-.09132	2.98201	-.27839	-.35034	-37718
Puntua: DISPOSICION DE MANO DE OBRA MILITAR	3.66397	-.14793	-.08740	2.83364	-.30291	-.34571	-.34956
Puntua: Personal activo militar	3.61747	-.49082	1.61565	1.77421	-.39646	-.21929	-.18896
Puntua: Activos de reserva	.23327	-.20851	1.88504	.69976	-.36834	.09194	-.40162
Puntua: Disponibles para obra militar	3.73279	-.14372	-.15055	2.74999	-.29717	-.34569	-.33544
Puntua: Alcanzan la edad para milicia	3.10108	-.13715	-.22395	3.41106	-.36804	-.31569	-.32769
Puntua(Paramilitares)	1.35130	-.28454	-.18508	.21252	-.29688	-.26498	2.08908
Puntua: Fuerza Laboral	3.97964	-.15676	-.04877	2.41695	-.25151	-.35464	-.38001
Puntua: Armas aéreas (Bases)	1.12346	-.52537	3.57662	.02152	.46082	-.25143	-.88977
Puntua(Helicópteros)	.07108	-.37475	4.24755	-.42015	.62886	-.38277	-.67066
Puntua: Armas Terrestres	1.38121	-.46578	4.38977	.08595	-.24712	-.19250	-.20870
Puntua(Artillería)	3.07543	-.55255	2.96519	.46561	-.16479	-.22198	-.34988
Puntua: armas nucleares	-.04327	-.23039	4.78989	-.18475	-.18018	-.20483	-.23039
Puntua: Total marina	4.09387	-.34489	1.95849	.06814	-.00424	-.39951	-.19960
Puntua(Puertos)	.48811	.79318	.48811	.86945	.86945	-.77030	-1.22791
Puntua(Aeropuertos)	-.20163	1.20202	.69343	-.33820	-.31857	-.37274	-.42173
Puntua: Marina mercante	3.27827	-.19571	1.59484	.30527	.23055	-.40546	-.84897
Puntua: Petróleo producción	.72920	.08153	2.76135	-.35011	-.51197	-.48949	1.95331
Puntua: Petróleo consumo	3.64168	-.01724	.60532	.62966	.38313	-.68957	-.09463
Puntua: Petróleo reserva	-.23941	.13569	-.41169	-.39294	-.46363	-.31488	2.50483
Puntua: Presupuesto de Defensa	1.81204	-.47107	.98578	.41837	1.28909	-.67235	-.29141
Puntua: PIB (Paridad poder adquisitivo)	3.75530	-.19441	.36982	.98111	.61907	-.61226	-.55345
Puntua: Reservas de div y oro	4.17275	-.29345	.52651	-.11342	.18033	-.35650	-.34102