



*Aseorías y Tutorías para la Investigación Científica en la Educación Puig-Salabarría S.C.  
José María Pino Suárez 400-2 esq a Lerdo de Tejada. Toluca, Estado de México. 7223898475*

RFC: ATII20618V12

**Revista Dilemas Contemporáneos: Educación, Política y Valores.**

<http://www.dilemascontemporaneoseducacionpoliticayvalores.com/>

**Año: VIII    Número:3    Artículo no.:25    Período: 1ro de mayo al 31 de agosto del 2021.**

**TÍTULO:** Análisis de clúster de universidades ecuatorianas según su posicionamiento en el ranking mundial.

**AUTORES:**

1. Dr. Vladimir Vega Falcón.
2. Dr. Fernando Castro Sánchez.
3. Dr. Belkis Sánchez Martínez.

**RESUMEN:** El análisis de rankings de universidades y sus variables, ayuda a tomar decisiones a nivel ministerial dentro de cada país, resultando necesario la agrupación de estas instituciones en clústeres que maximicen la homogeneidad intra-grupos y la heterogeneidad inter-grupos, para la toma de decisiones. El objetivo fue conformar los clústeres de las 65 universidades ecuatorianas, según el posicionamiento de las variables que conforman el ranking mundial de estas. Fue un estudio observacional, retrospectivo, transversal y analítico, utilizando las k-medias e hipótesis mediante el ANOVA. Como resultado se conformaron tres clústeres de universidades ecuatorianas, según su posicionamiento en el ranking mundial. Se concluye que mediante el clustering se agruparon satisfactoriamente las universidades de Ecuador para una mejor toma de decisiones.

**PALABRAS CLAVES:** ANOVA, análisis multivariante, clústeres, k-medias, rankings de universidades.

**TITLE:** Cluster analysis of Ecuadorian universities according to their position in the world ranking.

**AUTHORS:**

1. Dr. Vladimir Vega Falcón.
2. Dr. Fernando Castro Sanchez.
3. Dr. Belkis Sanchez Martinez.

**ABSTRACT:** The analysis of university rankings and their variables helps to make decisions at the ministerial level within each country, resulting necessary the grouping of these institutions in clusters that maximize intra-group homogeneity and inter-group heterogeneity, for decision making. The objective was to form the clusters of the 65 Ecuadorian universities, according to the positioning of the variables that make up their world ranking. It was an observational, retrospective, cross-sectional and analytical study, using k-means and hypotheses through ANOVA. As a result, three clusters of Ecuadorian universities were formed, according to their positioning in the world ranking. It is concluded that by means of clustering, Ecuadorian universities were satisfactorily grouped for better decision making.

**KEY WORDS:** ANOVA, multivariate analysis, clusters, k-means, university rankings.

**INTRODUCCIÓN.**

El Ranking Web de Universidades, popularizado como Webometrics, lo crea el Laboratorio de Cibermetría del Consejo Superior de Investigaciones Científicas (CSIC) en el año 2004, a partir de cuya fecha lo actualiza semestralmente. De acuerdo con Aguillo (2021b), su objetivo es promover la presencia académica en la web, lo cual sirve de apoyo a iniciativas de acceso abierto para desarrollar la transferencia del conocimiento cultural y científico creado por la educación superior, a toda la sociedad.

Webometrics, elaborado en España, tiene alto prestigio internacional y evalúa a todas las universidades del mundo, a las cuales les mide el impacto científico de sus actividades, la relevancia económica de la transferencia tecnológica a la industria, el compromiso con la comunidad, la enseñanza, y hasta la influencia política.

Esta valiosa información es útil para la toma de decisiones en los diferentes ministerios de educación superior del mundo, pero sería mejor si pudiese estar clasificada por conglomerados de universidades, según la analogía de los perfiles dentro de cada grupo.

El análisis de conglomerados o análisis clúster (grupos en inglés) es una técnica que permite clasificar a los casos o sujetos en distintos grupos que van a ser relativamente homogéneos entre sí, diferenciándose del análisis discriminante que también sirve para agrupar o clasificar sujetos en grupos, pero en este último ya se conoce a priori los grupos de pertenencia, mientras que en el análisis de conglomerados no. De acuerdo con Levy & Varela (2005), el análisis de conglomerado facilita clasificaciones partiendo de observaciones anteriormente no clasificadas, con aplicaciones en investigaciones que pretendan encontrar conjuntos de datos análogos.

En concordancia con Fontalvo-Herrera y Hoz-Granadillo (2020), para conformar grupos, se manejan medidas de distancias y criterios de eslabonamiento para la clústerización de las observaciones, evaluándose su nivel de disimilaridad o similaridad.

En esencia, el análisis clúster o de conglomerados, según Sánchez, Maldonado & Velasco (2012), representa una técnica estadística multivariante, cuyo propósito es dividir un conjunto de objetos en grupos con vistas a que los perfiles de los objetos dentro de un mismo grupo sean análogos entre sí (cohesión interna) y los de los objetos de clúster diferentes sean desiguales (aislamiento externo).

Existen diversos métodos de agrupamiento en el análisis de clúster, siendo los más comunes los no jerárquicos, entre los que se encuentra el particional (k-medias) y el probabilístico (algoritmo EM),

así como los jerarquizados, que incluyen los aglomerativos (KNN) y los divisivos (descendentes), entre otros.

En la técnica de clustering no jerárquico se asignan las observaciones a los clústeres que son especificados en el método como punto de partida; es decir, que se necesita fijar, de antemano, el número de conglomerados en que se quieren agrupar los datos. El objetivo consiste en intentar obtener una clasificación por grupos suficientemente aceptable, en el sentido que la dispersión dentro de cada grupo que se forme sea la menor posible.

Cuando se busca la solución clúster, el método más conocido es el algoritmo de k medias (k-means), que se caracteriza porque en cada etapa se comprueba si la configuración clúster que se tiene, hasta ese momento, puede ser mejorada reasignando observaciones de un grupo a otro. K-medias es el algoritmo de clustering por excelencia y también se le denomina algoritmo de Lloyd, pues fue propuesto por primera vez por Stuart Lloyd, en 1957, de acuerdo con Martínez-Pérez (2017).

En particular, la técnica de k medias parte de una interacción 0 que comienza con una posición aleatoria de los centroides (vector de medias para las variables medidas en la investigación) y en la iteración 1 se le asigna cada caso al centroide más próximo. Seguidamente, en una iteración 2, se recoloca el centroide en el centro de los puntos asignados. Posteriormente, en una iteración 3, se continúa iterativamente hasta que la distancia entre centroides sea máxima y, por otra parte, la distancia caso-centroide sea mínima. Los autores consideran que el método de las k-medias es el más manejado en aplicaciones científicas (Alayon, Moncada, Medina, y Rodríguez, 2020).

Cualquiera de los métodos jerárquicos por separado funciona peor que k-medias, el cual es el método más robusto respecto a la presencia de outliers y errores en las medidas de distancia, aunque requiere que el investigador especifique previamente el número de conglomerados y sus centros iniciales. El algoritmo de k-medias, aglomera o pronostica los datos con un aprendizaje no supervisado (Vintimilla, Astudillo-Salinas, Severeyn, Encalada, y Wong, 2017).

En este contexto, el problema de investigación se define de la forma siguiente: ¿Cómo conformar grupos de universidades ecuatorianas según el posicionamiento de las variables que conforman el ranking mundial de éstas?

El estudio incluye una hipótesis del investigador y es que el Impacto, la Apertura y la Excelencia (las tres variables que conforman el ranking) en los grupos que se crean es diferente.

Para abordar este problema, se plantea el objetivo de conformar los clústeres de las 65 universidades ecuatorianas, según el posicionamiento de las variables que conforman el ranking mundial de estas.

## **DESARROLLO.**

### **Taxonomía de la investigación.**

De acuerdo con Supo y Zacarías (2020), se desarrolló un estudio de tipo analítico (el análisis estadístico fue multivariado), observacional (no hubo intervención en la población de estudio) transversal (la variable fue medida en un solo momento, en este caso en enero de 2021), y retrospectivo (los datos ya existían en el momento del estudio).

### **Población de estudio.**

La población estuvo conformada por las 65 universidades de Ecuador. El criterio de inclusión se basó en aparecer en el listado de Webometrics (Aguillo, 2021a), lo cual se cumplió en el 100% de las universidades ecuatorianas, mientras que el criterio de exclusión fue no estar vigente como universidad de Ecuador, en el momento de realizarse el presente estudio, lo cual no ocurrió en ninguno de los casos. Por lo antes mencionado, no fue necesario calcular una muestra.

### **Variables de estudio.**

La variable de agrupación fue la conformación de clústeres de universidades ecuatorianas y las variables clasificadoras fueron: Impacto; Excelencia; y Apertura. Estas tres últimas son las que conforman el Webometrics.

Las variables estudiadas fueron de tipo numérico (de razón) por su propia naturaleza y además se expresan numéricamente según el lugar que ocupe cada universidad dentro de ellas.

De acuerdo con Aguillo (2021b), el ranking mundial de universidades se establece a partir del lugar que ocupa cada universidad, luego de calcularse a partir de la fórmula siguiente:

$$\text{Ranking} = \text{Impacto (0,50)} + \text{Excelencia (0,40)} + \text{Apertura (0,10)}$$

Las variables clasificadoras se conceptualizan a continuación (Aguillo, 2021b):

- **Impacto o Visibilidad (Posición).** Se refiere al impacto del contenido web, o sea, al lugar que ocupa la universidad según el número de redes externas (subredes) afines a las páginas web de la institución (seleccionándose el valor máximo después de la normalización). La fuente es Ahrefs Majestic y está ponderada con el 50% del ranking.
- **Apertura o Transparencia (Posición).** Se refiere al lugar que ocupa la universidad, de acuerdo con sus principales investigadores citados, a partir del número de citas de los 210 mejores autores (exceptuando a los 20 mejores valores atípicos). La fuente son los perfiles académicos de Google. Está ponderada con el 10% del ranking.
- **Excelencia o Académico (Posición).** Se refiere a los principales documentos citados, o sea, al lugar que ocupa la universidad según la cantidad de artículos científicos enmarcados entre el 10% con más citas, en un total de 26 disciplinas académicas, correspondientes al período 2013-2017. La fuente es SCImago Journal Rank (SJR), como factor de medición que define la calidad de las publicaciones científicas según el recuento de citas alcanzadas por cada publicación. Está ponderada con el 40% del ranking.

### **Herramientas estadísticas.**

El análisis multivariante se efectuó a través del método de las k-medias por resultar beneficioso cuando los datos a clasificar son cuantiosos (65 universidades en el presente estudio), en concordancia

con Gallego & Aranque (2019), partiendo del supuesto de que el número de grupos fue conocido a priori, siendo en este caso tres, por coincidirse con (Kakushadze y Yu, 2016b; Kakushadze y Yu, 2017a; y con Kakushadze y Yu, 2017b), respecto a que el número de factores estadísticos (Impacto; Apertura; y Excelencia) es válido para identificar el número objetivo de clústeres.

Además, dicha cifra se consideró una cantidad prudente para la toma de decisiones que pueda hacer en un futuro la Secretaría de Educación Superior, Ciencia, Tecnología e Innovación (SENESCYT) y el Consejo de Educación Superior (CES) de Ecuador, a partir de la conformación de estos grupos de universidades en función a su posicionamiento en el ranking mundial y las variables que conforman el mismo.

Para utilizar el método de las k-medias, se siguió el diagrama de flujo de Saxena et al. (2017), el cual se muestra en la Figura 1.

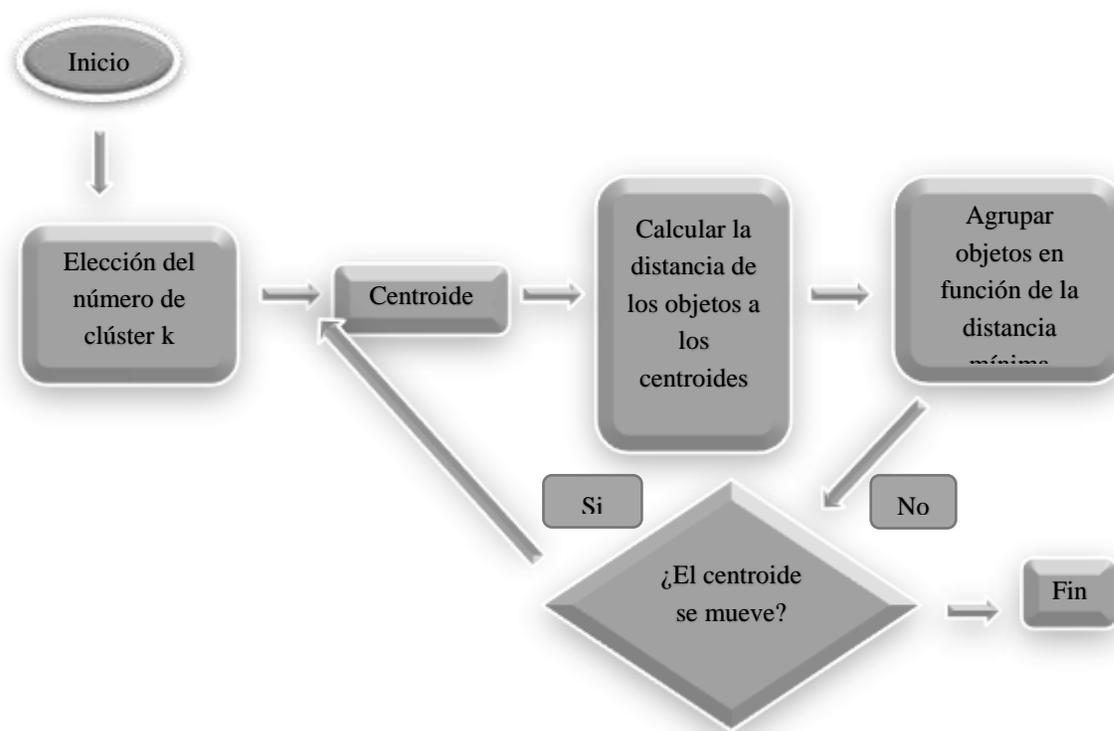


Figura 1. Diagrama de flujo del algoritmo K -means. Fuente: elaboración propia a partir de (Saxena et al., 2017).

Se utilizó el software IBM SPSS Statistics (v. 26.0, Edición de 64 bits), creándose previamente una base de datos que contuviera toda la información estadística del estudio en Microsoft Excel 2019.

### **Hipótesis del estudio.**

Considerando que la intención analítica fue la prueba de hipótesis, se desarrollaron los pasos correspondientes al ritual de la significancia estadística:

- Planteamiento del sistema de hipótesis:

La hipótesis a dos colas o bilateral quedó planteada de la forma siguiente:

- ✓ H0: El Impacto, la Apertura y la Excelencia en los tres grupos es igual (hipótesis nula o de trabajo).

Las hipótesis alternas o del investigador fueron:

- ✓ H1: El Impacto tiene una diferencia estadísticamente significativa en los tres grupos.
- ✓ H2: La Apertura tiene una diferencia estadísticamente significativa en los tres grupos.
- ✓ H3: La Excelencia tiene una diferencia estadísticamente significativa en los tres grupos.
- Establecimiento del nivel de significancia: se precisó un nivel de significancia del 5% (0,05).
- Elección del estadístico de prueba: el estadístico de prueba seleccionado fue la prueba estadística análisis de la varianza (ANOVA).
- Lectura al p-valor calculado: se analizó la probabilidad de error de acuerdo con el valor final que se obtuvo.
- Decisión estadística: se tuvo como punto de partida el criterio de que si el p-valor es menor que 0,05 se rechaza la hipótesis nula y se acepta la hipótesis del investigador.

### **Resultados.**

Siguiendo el diagrama de flujo de la Figura 1, se obtuvo los resultados siguientes:

- **Elección del número de clúster k.**

La clasificación con el algoritmo de k-medias dividió a la base de datos de 65 universidades en  $K = 3$  grupos, con el método de iterar y clasificar, definiéndose un número máximo de 10 iteraciones.

Para caracterizar a toda la población de estudio, en la Tabla 1 se muestran las variables de las 65 universidades ecuatorianas, de acuerdo con Aguillo (2021a), a las cuales se les asignó un código, que correspondió al lugar en que se ordenaron descendientemente en el ranking mundial, que es con el que se hace referencia a ellas en la Tabla 7.

Tabla 1. Variables de cada universidad ecuatoriana y lugar en el ranking mundial.

<b>Código</b>	<b>Universidad</b>	<b>Impacto (Posición)</b>	<b>Apertura (Posición)</b>	<b>Excelencia (Posición)</b>	<b>Ranking Mundial</b>
1	Escuela Politécnica Nacional	3408	1846	1878	1745
2	Universidad de las Américas Ecuador	1155	2395	3298	1948
3	Universidad Técnica Particular de Loja	4188	1869	2702	2284
4	Universidad San Francisco de Quito	3070	5963	1432	2292
5	Pontificia Universidad Católica del Ecuador	3673	2089	2834	2296
6	Universidad de las Fuerzas Armadas ESPE	4220	1798	2774	2324
7	Universidad Politécnica Salesiana Ecuador	4553	2037	3154	2631
8	Universidad de Cuenca	5160	2103	2991	2676
9	Escuela Superior Politécnica del Litoral	1990	5963	2471	2717
10	Universidad Central del Ecuador	6508	2301	3258	3046
11	Escuela Superior Politécnica de Chimborazo	5412	3463	4441	3760
12	Universidad del Azuay	5613	3590	4675	3988
13	Universidad Técnica del Norte	8511	3129	4269	4036
14	FLACSO Ecuador Facultad Latinoamericana de Ciencias Sociales	7531	2668	4675	4106
15	Yachay Tech	15399	2316	3171	4140
16	Universidad Técnica de Ambato	4000	5963	4359	4255
17	Universidad Internacional del Ecuador	6101	4943	4579	4305
18	Universidad Nacional de Loja	9498	3614	4579	4425
19	Universidad Nacional de Chimborazo	11163	3520	4441	4519
20	Universidad UTE	8476	5963	3690	4537
21	Universidad Nacional de Educación UNAE	3966	3737	5655	4551
22	Universidad Tecnológica Indoamérica	12953	3591	4441	4730
23	Universidad Técnica de Machala	11163	2602	5006	4739
24	Ikiam Universidad Regional Amazónica	13677	3795	4359	4807
25	Universidad del Pacífico Escuela de Negocios	2624	5963	5541	4882
26	Universidad Estatal Amazónica	16900	3399	4032	4893
27	Universidad Católica de Santiago de Guayaquil	8981	5963	4299	4927
28	Universidad de Guayaquil	5863	5963	4934	4948
29	Universidad Estatal Península de Santa Elena	9240	4848	5073	4989
30	Universidad de Especialidades Espíritu Santo	10717	5963	4115	5015
31	Universidad Andina Simón Bolívar Ecuador	4273	5963	5342	5017
32	Instituto de Altos Estudios Nacionales	4015	5963	5442	5039

33	Universidad Católica de Cuenca	12452	3043	5201	5096
34	Universidad Laica Eloy Alfaro de Manabí	9489	5963	4781	5307
35	Universidad Técnica de Cotopaxi	3002	5963	5999	5432
36	Universidad Particular Internacional SEK	14166	4577	5442	6038
37	Universidad Técnica de Manabí	10694	5963	5541	6624
38	Universidad Regional Autónoma de los Andes UNIANDES	13443	4063	5999	6913
39	Universidad Laica Vicente Rocafuerte de Guayaquil	8511	5370	6245	7045
40	Universidad Estatal de Milagro	11631	1616	6683	7166
41	Universidad Técnica Estatal de Quevedo	14185	5963	5442	7631
42	Universidad Estatal de Bolívar	13677	5878	5655	7984
43	Universidad Metropolitana del Ecuador	10111	4226	6683	8151
44	Universidad Casa Grande	12414	5963	5999	9163
45	Universidad Agraria del Ecuador	15888	5703	5807	9433
46	Universidad Técnica de Babahoyo	14412	3807	6683	10534
47	Escuela Superior Politécnica Agropecuaria de Manabí	16072	4943	6245	10574
48	Universidad Técnica Luis Vargas Torres de Esmeraldas	17330	4781	6245	11147
49	Universidad Tecnológica ECOTEC	14283	4571	6683	11487
50	Universidad Estatal del Sur de Manabí UNESUM	12414	5543	6683	12161
51	Universidad Politécnica Estatal del Carchi UPEC	16531	5543	6245	12325
52	Universidad Particular San Gregorio de Portoviejo USGP	11237	5963	6683	12856
53	Universidad de los Hemisferios	15717	5963	6245	13369
54	Universidad Tecnológica Israel	15114	5420	6683	13977
55	Pontificia Universidad Católica Sede Ibarra	16126	5213	6683	14142
56	Pontificia Universidad Católica del Ecuador Sede Ambato	15984	5291	6683	14229
57	Universidad de Otavalo	19108	5963	6245	16020
58	Universidad de las Artes de Ecuador	15687	5963	6683	16532
59	Universidad Santa María de Chile Campus Guayaquil	17691	5963	6683	18293
60	Escuela de Dirección de Empresas	18070	5963	6683	18621
61	Universidad Tecnológica Empresarial de Guayaquil	18239	5963	6683	18772
62	Pontificia Universidad Católica de Ecuador Sede Santo Domingo	20131	5751	6683	19251
63	Universidad Iberoamericana del Ecuador	20602	5963	6683	20941
64	Universidad de Especialidades Turística	23018	5963	6683	23221
65	Universidad Internacional de La Rioja Ecuador	23191	5963	6683	23386

Fuente: Aguillo (2021a).

- **Centroide y cálculo de la distancia de los objetos a los centroides.**

Se seleccionaron los k centroides, donde k es el número de grupos deseado, o sea, los tres grupos, por considerarse una cifra adecuada, además, por coincidir con la cifra de variables independientes, tal como se expuso en el apartado Método.

Dado que las variables están en escalas muy diferentes, fue necesario estandarizarlas. Para ello, el software IBM SPSS Statistics ejecutó el procedimiento descriptivo sobre variables tipificadas y se crearon las nuevas variables en la matriz que se utilizaron en el estudio.

En la Tabla 2 se exponen los estadísticos descriptivos de cada variable independiente.

Tabla 2. Estadísticos descriptivos.

<b>Variabes</b>	<b>N</b>	<b>Mínimo</b>	<b>Máximo</b>	<b>Media</b>	<b>Desviación Estándar (DE)</b>	<b>Varianza</b>
Impacto	65	1155	23191	10994,17	5677,286	32231576,799
Apertura	65	1616	5963	4616,42	1493,609	2230866,497
Excelencia	65	1432	6683	5151,65	1420,584	2018058,263

Fuente: elaboración propia.

- **Agrupar objetos en función de la distancia mínima.**

En la Tabla 3 se exponen los centros de clústeres iniciales. Se asignó cada observación al grupo cuyo centroide fue el más cercano.

Tabla 3. Centros de clústeres iniciales.

<b>Variabes</b>	<b>1</b>	<b>2</b>	<b>3</b>
Impacto	23191	1155	8511
Apertura	5963	2395	5370
Excelencia	6683	3298	6245

Fuente: elaboración propia.

En la Tabla 4 se muestra el historial de las únicas cuatro iteraciones que fueron necesarias, dentro de un número máximo de 10 iteraciones que se habían permitido, por lo que no fueron necesarias las otras seis.

Tabla 4. Historial de iteraciones <sup>a</sup>.

Iteración	Cambiar en centros de clústeres		
	1	2	3
1	4901,092	2640,800	2493,113
2	625,788	860,771	687,264
3	0,000	177,396	172,302
4	0,000	0,000	0,000

a. Convergencia conseguida debido a que no hay ningún cambio en los centros de clústeres o un cambio pequeño. El cambio de la coordenada máxima absoluta para cualquier centro es 0,000. La iteración actual es 4. La distancia mínima entre los centros iniciales es 8464,406. Fuente:

elaboración propia.

Es oportuno señalar, que hasta que se satisfaga un criterio de parada, como por ejemplo, los centroides casi no cambian o los grupos derivados de dos iteraciones consecutivas son idénticos, el proceso de iteración debe continuar.

- **Respuesta a la pregunta: ¿El centroide se mueve?**

La Tabla 5 destaca los centros de clústeres finales y el número de casos en cada clúster. Con ello, la respuesta a la pregunta fue que a partir de esos resultados los centroides ya no se mueven.

**Tabla 5.** Centros de clústeres finales.

	Clúster		
	1 (N=19)	2 (N=21)	3 (N=25)
Impacto	17726	4301	11500
Apertura	5370	3931	4620
Excelencia	6197	3925	5387

Fuente: elaboración propia.

Respecto a la hipótesis del estudio, en la Tabla 6 se muestra el análisis de la varianza (ANOVA).

Tabla 6. Análisis de la varianza (ANOVA).

	Clúster		Error		F	Sig.
	Media cuadrática	gl	Media cuadrática	gl		
Impacto	904098536,451	2	4106836,165	62	220,145	0,000
Apertura	10331166,833	2	1969566,486	62	5,245	0,008
Excelencia	26866785,017	2	1216486,433	62	22,086	0,000

Fuente: elaboración propia.

En la Tabla 7, se expone la composición final de cada uno de los tres clústeres conformados, mencionándose los códigos de cada universidad mostrados en la Tabla 1.

Tabla 7. Composición final de los tres clústeres conformados.

Clúster 1 (N=19)	Clúster 2 (N=21)	Clúster 3 (N=25)
15; 26; 45; 47; 48; 51; 53; 54; 55; 56; 57; 58; 59; 60; 61; 62; 63; 64; 65	1;2; 3; 4; 5; 6; 7; 8; 9; 10; 11; 12; 14; 16; 17; 21; 25; 28; 31; 32; 35	13; 18; 19; 20; 22; 23; 24; 27; 29; 30; 33; 34; 36; 37; 38; 39; 40; 41; 42; 43; 44; 46; 49; 50; 52

Fuente: elaboración propia.

### Valores atípicos (Outliers).

Teniendo en cuenta, que la agrupación puede no funcionar bien si los datos están llenos de valores atípicos, se valoró este aspecto en el estudio, llegándose a la conclusión de que las variables Impacto y Apertura no presentaron ningún valor atípico, mientras que la variable Excelencia solo tuvo un valor atípico, pero poco significativo, en la Universidad San Francisco de Quito, con una cifra de -2,61839, muy cercano al punto de corte tradicionalmente asumido de 2,5.

### Discusión.

En relación a la Figura 1 (Saxena et al., 2017), donde se expone el diagrama de flujo del algoritmo k-means, debe indicarse que esta secuencia de pasos resulta una adecuada guía para conducir un estudio de esta naturaleza, por la simplicidad que brinda, siempre que sea acompañada por una potente herramienta como lo es el software IBM SPSS Statistics, tal como ocurre en este estudio.

La Tabla 1 expone ordenadamente las variables de las 65 universidades ecuatorianas, especificando el lugar de cada una en el ranking mundial.

Es preocupante que solo dos universidades se encuentren entre las 2000 mejores del planeta: Escuela Politécnica Nacional (lugar 1745) y Universidad de las Américas Ecuador (lugar 1948), lo cual debe poner en alerta a las autoridades de la educación superior ecuatoriana. Además, más de la mitad se encuentran ubicadas en lugares posteriores al 5000.

Resulta también preocupante, que en las tres variables clasificadoras, ninguna universidad de Ecuador se ubica en el top 1000. Respecto al Impacto, la mejor ubicada es la Universidad de las Américas Ecuador, ocupando el lugar 1155, mucho mejor ubicada que el resto. En relación a la variable Excelencia, la mejor situada es la Universidad San Francisco de Quito en el puesto 1432, también superando claramente al resto de las instituciones ecuatorianas. Finalmente, respecto a la variable Apertura, la mejor posicionada es la Universidad de las Fuerzas Armadas ESPE, ubicada en el lugar 1798.

Es llamativo que solo dos universidades se ubican en el top 10 de las tres variables dentro de Ecuador: Escuela Politécnica Nacional y Pontificia Universidad Católica del Ecuador.

Igualmente, es interesante analizar que hay universidades mal ubicadas en el ranking general dentro de Ecuador, pero bien ubicadas en variables individuales. Por ejemplo, la Universidad Estatal de Milagro que integralmente ocupa la posición 40, en la variable Apertura es la mejor institución ecuatoriana, a pesar de su discreto lugar 1616 en el contexto mundial en dicha variable. Igualmente, en la variable Impacto, la Universidad Técnica de Cotopaxi, ocupa el lugar cuatro dentro de Ecuador, a pesar de que en el ranking integral ecuatoriano se ubica en la posición 35.

En la Tabla 2 se muestran los estadísticos descriptivos, resaltándose la alta desviación estándar en la variable Impacto (5677,286), lo que evidencia que en relación al lugar que ocupa la universidad, según el número de redes externas afines a las páginas web de la institución, hay una notable dispersión de resultados en las universidades ecuatorianas.

La Tabla 3 refleja los centros de clústeres iniciales, mostrando que respecto a los valores de las variables de las  $k$  observaciones, las mayores aproximaciones se evidencian en la variable Excelencia, entre los grupos 1 (6683) y 3 (6245), aunque estos resultados son normales en esta etapa del proceso de iteración. Por su parte, en la Tabla 4 se detalla el historial de iteraciones, apreciándose los cambios en los centros de clústeres, que en el caso del grupo 1, en la iteración tres ya quedó en un valor de 0,000, lo cual ocurre en la iteración cuatro para los grupos 2 y 3; por ello, en ningún caso son necesarias las 10 iteraciones permitidas, lo cual manifiesta que no resulta muy complejo este proceso. La Tabla 5 presenta los centros de clústeres finales, percibiéndose que los grupos quedan conformados por cantidades similares de universidades: grupo 1 (N=19); grupo 2 (N=21) y grupo 3 (N=25). Se destaca que la mayor diferencia entre todas las variables respecto a las distancias entre clúster diferentes corresponde a la variable Impacto, con un valor de 17726 para el grupo 1 y de 4301 en el grupo 2.

Por su parte, la Tabla 6 exhibe el análisis de la varianza (ANOVA), que refleja que en las tres variables el p-valor es menor a 0,05, por lo cual la decisión estadística, dentro del ritual de significancia es que se rechaza la hipótesis nula (El Impacto, la Apertura y la Excelencia en los tres grupos es igual) y se acepta la hipótesis del investigador, es decir, el Impacto, la Apertura y la Excelencia tienen una diferencia estadísticamente significativa en los tres grupos, o sea, se corrobora que los tres clústeres son distintos en cada caso.

Debe agregarse que las pruebas F en la prueba de ANOVA (Impacto 220,145; Apertura 5,245; y Excelencia 22,086) se realiza para determinar si la variabilidad entre las medias de los clústeres es mayor que la variabilidad de las universidades dentro de los grupos, por lo que en este caso, sólo se deben utilizar con fines descriptivos porque los clústeres se han designado para maximizar los contrastes entre las universidades de distintos agrupamientos; lo cual no ocurre así cuando ANOVA se utiliza desde el principio con otros fines diferentes al de clustering,

Finalmente, en la Tabla 7 se expone la composición final de los tres clústeres, donde se evidencia que en el clúster dos están las universidades mejor ubicadas en el ranking integral, donde se encuentra 21 de las mejores 35 universidades. Por su parte, en el clúster uno aparece las peor ubicadas, aunque la Yachay Tech, que integralmente ocupa el lugar 15, queda enmarcada en este grupo, por su pésima posición en la variable Impacto (lugar 15399 universal), algo similar a lo que ocurre con la Universidad Estatal Amazónica, que ocupando el lugar 26 entre las universidades ecuatorianas, su mal Impacto (lugar 16900 mundialmente) le impide escalar a mejores posiciones.

### **Valores atípicos.**

Tal como se comenta en el apartado Resultados, dos de las variables clasificadores (Impacto y Apertura) no evidencian valores atípicos, y la variable Excelencia solo tiene como valor atípico, el correspondiente a la Universidad San Francisco de Quito (-2,61839), lo cual resulta casi insignificante por ubicarse muy cercano al punto de corte asumido de 2,5. Por ello, sobre este aspecto se llega a la conclusión de que los valores atípicos no afectan el clustering que se realiza.

En relación con el valor de corte de los valores atípicos, los autores consideran que las puntuaciones pueden ser outliers cuando tiene una puntuación típica, en muestras pequeñas (<80), cuando son menores de 2,5 o cuando tienen una puntuación típica en muestras grandes (>80) de 3 puntos.

Lo anterior tiene en consideración que en una variable que sigue una distribución normal se espera que el 95% de los datos se alejen de la media como máximo 1,96 desviaciones típicas. Además, se espera que un 5% de los casos se pueda alejar más de 1,96 desviaciones típicas y solo el 1% de los casos se aleja más de 2,58 desviaciones típicas de la media; esto quiere decir, que se escoge el valor de 2,5 para detectar casos extremos, porque son casos que se alejan demasiado de la media. Si en la variable hay menos del 1% de casos que se alejan de la media 2,5 desviaciones típicas, en ese caso no ocurre nada, porque es lo esperado, pero si en la variable hay más del 1% de casos que se alejan

2,5 desviaciones típicas, entonces puede ser que esos valores no sean correctos por algún motivo. Entonces se elige el valor de 2,5 para detectar casos e investigar, en la medida que se pueda, si son casos esperados o no; por lo tanto, el valor de corte de 2,5 responde al modelo de probabilidad de la distribución normal.

El estudio que se realiza en las 65 universidades ecuatorianas posibilita que a partir de las tres variables que conforman el ranking (Impacto. Excelencia y Apertura), se construyan tres grupos, de forma tal que las que pertenezcan a un mismo grupo sean muy homogéneas entre sí y al mismo tiempo se obtiene la máxima heterogeneidad posible entre grupos, o sea, la investigación intenta minimizar la variabilidad intra-grupo y simultáneamente maximizar la variabilidad inter-grupos.

Los resultados demuestran que los perfiles de las universidades dentro de un mismo clúster son análogos entre sí; es decir, tiene una cohesión interna, mientras que los de clúster diferentes son desiguales, o sea, tienen un aislamiento externo. Esto permite que la dirección de la educación superior en Ecuador, a partir de organismos rectores como SENESCYT que tiene entre sus objetivos institucionales desarrollar el sistema de educación superior universitaria, técnica y tecnológica, con discernimientos de inclusión, calidad, pertinencia, y democracia; así como elevar los resultados investigativos, la innovación, la investigación y la transferencia tecnológica, manteniendo un nexo entre los sectores productivo y académico, disponga ahora de esta segmentación válida para hacer distinciones en la gestión de presupuestos, planes estratégicos y desarrollo científico e informático, entre otros (Secretaría de Educación Superior, Ciencia, Tecnología e Innovación, 2021).

Este comentario también es válido para el CES, que tiene como misión "...la planificación, regulación y coordinación interna del Sistema de Educación Superior del Ecuador, y la relación entre sus distintos actores con la Función Ejecutiva y la sociedad ecuatoriana." (Consejo de Educación Superior, 2021).

Para ambas instituciones, la sola intención de precisar escenarios futuros de actuación para transformar la información en acción (Vega y Comas, 2017), posibilita prever una mejor evolución de las universidades ecuatorianas, accionando a partir de la segmentación realizada en este estudio.

### **Limitaciones del estudio.**

Para futuras investigaciones se invita a la comunidad académica y científica a replicar en otros contextos los procedimientos seguidos en este estudio, valorando la posible utilización de la combinación de métodos jerárquico (para eliminar outliers, determinar el número de clústeres y sus centros iniciales), con el método k-medias para obtener así una solución conglomerativa final, ya que no fue utilizado así en el presente estudio, lo cual pudiera considerarse como una posible limitación del mismo.

### **CONCLUSIONES.**

Se conformaron tres clústeres de todas las universidades ecuatorianas, de acuerdo al posicionamiento de las variables que conforman el Ranking Web de Universidades, popularizado como Webometrics. Se siguió dentro del análisis de clúster, su principio rector, que consiste en la maximización de la homogeneidad intra-grupos (conglomerados, segmentos) y simultáneamente la maximización de la heterogeneidad inter-grupos.

Se brinda a las entidades rectoras de la educación en Ecuador, una segmentación de las universidades ecuatorianas, que posibilitará futuras acciones encaminadas a una toma de decisiones no solo de forma global, sino de acuerdo a las peculiaridades de la segmentación propuesta, a partir de las variables Impacto, Excelencia y Apertura, siendo necesario actualizar el presente algoritmo de clustering, en la medida en que se desarrollen futuras actualizaciones de Webometrics.

**REFERENCIAS BIBLIOGRÁFICAS.**

1. Aguillo Caño, Isidro F. (2021a). Ranking Web de Universidades. Madrid. España: Laboratorio de Cibermetría. Recuperado de: [https://www.webometrics.info/es/Latin\\_America\\_es/Ecuador](https://www.webometrics.info/es/Latin_America_es/Ecuador)
2. Aguillo Caño, Isidro F. (2021b). Ranking Web de Universidades. Madrid. España: Laboratorio de Cibermetría. Recuperado de: <http://www.webometrics.info/en/Methodology>
3. Alayon Ricardo A., Moncada Diego A., Medina Victor H., y Rodríguez Jorge E. (2020). Estimación de rendimiento académico a través de técnicas para minería de datos. *Sistemas, Cibernética e Informática*. 17(1), 90-95.
4. Consejo de Educación Superior. (2021). Misión, visión y objetivos. (Internet). Recuperado de: [https://www.ces.gob.ec/?page\\_id=44](https://www.ces.gob.ec/?page_id=44)
5. Fontalvo-Herrera, T. J., y Hoz-Granadillo, E. D. L. (2020). Conglomerate method-discriminant analysis-data envelopment analysis to classify and evaluate business efficient. *Entramado*, 16(2), 46-55.
6. Gallego, L., & Araque, O. (2019). Variables de influencia en la capacidad de aprendizaje. un análisis por conglomerados y componentes principales. *Información tecnológica*, 30(2), 257-264.
7. Kakushadze, Z. and Yu, W. (2016b) Factor Models for Cancer Signatures. *Physica A* 462: 527-559.
8. Kakushadze, Z. and Yu, W. (2017a) How to Combine a Billion Alphas. *Journal of Asset Management* 18(1): 64-80.
9. Kakushadze, Z. and Yu, W. (2017b) Statistical Risk Models. *The Journal of Investment Strategies* 6(2): 1-40.
10. Levy Mangin, J. & Varela Mallou, Jesús (2005). *Análisis multivariable para las ciencias sociales*. Madrid-España. Editorial Pearson Educación.

11. Martínez-Pérez, Martín Agustín. (2017). Agrupación de patrones en series de tiempo usando una red neuronal autoorganizativa (SOM) para el análisis de concentraciones de SO<sub>2</sub>. *Revista de Ciencias Ambientales y Recursos Naturales*. 3(7): 16-24.
12. Sánchez, P. P. I., Maldonado, C. J., & Velasco, A. P. (2012). Caracterización de las Spin-Off universitarias como mecanismo de transferencia de tecnología a través de un análisis clúster. *Revista Europea de Dirección y Economía de la empresa*, 21(3), 240-254.
13. Saxena, A., Prasad, M., Gupta, A., Bharill, N., Patel, O. P., Tiwari, A., ... y Lin, C. T. (2017). A review of clustering techniques and developments. *Neurocomputing*, 267, 664-681.
14. Secretaría de Educación Superior, Ciencia, Tecnología e Innovación. (2021). Misión / Visión/ Valores/Objetivos. (Internet): Recuperado de: <https://www.educacionsuperior.gob.ec/valores-mision-vision/>
15. Supo, J. y Zacarías, H. (2020). Metodología de la investigación científica: Para las Ciencias de la Salud y las Ciencias Sociales. (3ra. Edición). Perú: Editorial Bioestadístico EEDU EIRL.
16. Vega, V., y Comas, R. (2017). Gestión del cambio y Dirección Estratégica. (1era. Edición). Quito. Ecuador: Editorial El Siglo.
17. Vintimilla, C., Astudillo-Salinas, F., Severeyn, E., Encalada, L., & Wong, S. (2017). Agrupamiento de K-medias para estimación de insulino-resistencia en adultos mayores de Cuenca. *Maskana*, 8, 31–39.

#### **DATOS DE LOS AUTORES.**

1. **Vladimir Vega Falcón.** Doctor en Ciencias Económicas. Docente de la Universidad Regional Autónoma de Los Andes, Ecuador. E-mail: [ua.vladimirvega@uniandes.edu.ec](mailto:ua.vladimirvega@uniandes.edu.ec)
2. **Fernando Castro Sánchez.** Doctor en Ciencias de la Educación. Docente de la Universidad Regional Autónoma de Los Andes, Ecuador. E-mail: [ua.fernandocastro@uniandes.edu.ec](mailto:ua.fernandocastro@uniandes.edu.ec)

**3. Belkis Sánchez Martínez.** Doctora en Medicina. Docente de la Universidad Regional Autónoma de Los Andes, Ecuador. E-mail: [ua.belkissanchez@uniandes.edu.ec](mailto:ua.belkissanchez@uniandes.edu.ec)

**RECIBIDO:** 26 de marzo del 2021.

**APROBADO:** 19 de abril del 2021.