

Segmentación de estudiantes universitarios según el nivel socioeconómico, ingesta calórica y composición corporal utilizando el algoritmo K-means

Segmentation of college students according to socioeconomic status, caloric intake and body composition using the K-means algorithm.

Katherine Denisse Suarez Gonzalez^{1[0009-0008-9220-9544]}, Nibia Noemí Novillo Luzuriaga^{2[0000-0001-8919-5040]},
Vanessa Paulina Vargas Olalla^{3[0000-0002-2116-2004]}, Erik Steven Suarez Gonzalez^{4[0009-0005-0232-7386]}

^{1,2,3,4}Universidad Estatal de Milagro (UNEMI), Facultad de Salud y Servicios Sociales. Ciudadela Universitaria “Dr. Rómulo Minchala Murillo”
Km 1.5 vía Milagro – Virgen de Fátima. 091706. Milagro – Guayas - Ecuador

¹ksuarezg@unemi.edu.ec, ²nnovillol@unemi.edu.ec, ³vvargaso@unemi.edu.ec, ⁴esuarezg2@unemi.edu.ec

CITA EN APA:

Suarez Gonzalez, K. D., Novillo Luzuriaga, N. N., Vargas Olalla, V. P., & Suarez Gonzalez, E. S. (2024). Segmentación de estudiantes universitarios según el nivel socioeconómico, ingesta calórica y composición corporal utilizando el algoritmo K-means. *Tesla Revista Científica*, 4(2), e433.
<https://doi.org/10.55204/trc.v4i2.e433>

Recibido: 2024-09-28

Revisado: 2024-10-05 al 2024-10-25

Corregido: 2024-11-08

Aceptado: 2024-11-11

Publicado: 2024-11-14

TESLA
Revista Científica
ISSN: 2796-9320



Los contenidos de este artículo están bajo una licencia de Creative Commons Attribution 4.0 International (CC BY 4.0) Los autores conservan los derechos morales y patrimoniales de sus obras.

The contents of this article are under a Creative Commons Attribution 4.0 International (CC BY 4.0) license. The authors retain the moral and patrimonial rights of their works.

Resumen:

El aumento de problemas de salud como la obesidad y las afecciones metabólicas ha reforzado la necesidad de entender cómo el nivel socioeconómico (NSE) impacta la nutrición, especialmente entre jóvenes adultos. Este estudio utilizó el algoritmo K-means para segmentar a jóvenes adultos universitarios basándose en su NSE, ingesta calórica y valores de índice de masa corporal (IMC). Se recopilieron datos de 150 estudiantes universitarios utilizando cuestionarios validados que permitieron mediar la cantidad de calorías consumidas y la proporción de macronutrientes. El análisis se llevó a cabo a niveles univariado, bivariado y multivariado mediante el algoritmo K-means, identificando tres clústeres dentro del grupo de estudio, se utilizó Python en la plataforma Google Colab para procesar la información. Los resultados identificaron tres clústeres: uno con una ingesta calórica ligeramente superior a lo recomendado, pero con un IMC normal, sugiriendo una mejor calidad de dieta o mayor actividad física; otro con un consumo calórico inferior al recomendado y un IMC que indicaba sobrepeso, posiblemente debido a una dieta de baja calidad; y un tercer clúster con altos niveles de consumo energético, incluyendo potencialmente a deportistas de alto rendimiento. La variabilidad dietética subraya la importancia de personalizar estrategias nutricionales para mejorar la salud y bienestar.

Palabras claves: Análisis cluster, Conducta alimentaria, Obesidad, Estudio nutricional, Estrategias preventivas.

Abstract:

The increase in health problems such as obesity and metabolic conditions has reinforced the need to understand how socioeconomic status (SES) impacts nutrition, especially among young adults. This study used the K-means algorithm to segment young adult college students based on their SES, caloric intake, and body mass index (BMI) values. Data were collected from 150 college students using validated questionnaires that measured the amount of calories consumed and the proportion of macronutrients. The analysis was carried out at univariate, bivariate and multivariate levels using the K-means algorithm, identifying three clusters within the study group; Python on the Google Colab platform was used to process the information. The results identified three clusters: one with a caloric intake slightly higher than recommended, but with a normal BMI, suggesting better diet quality or increased physical activity; another with a caloric intake lower than recommended and a BMI indicating overweight, possibly due to a low-quality diet; and a third cluster with high levels of energy intake, potentially including high-performance athletes. Dietary variability underscores the importance of personalizing nutritional strategies to improve health and well-being.

Keywords: Cluster analysis, Eating behavior, Obesity, Nutritional study, Preventive strategies

1. INTRODUCCIÓN

Actualmente, la salud y la nutrición son temas de preocupación creciente entre las sociedades modernas debido al aumento alarmante de problemas como la obesidad y las afecciones metabólicas. La asociación entre los niveles socioeconómicos y la salud es bien conocida y documentada, ya que las condiciones sociales y económicas influyen en el comportamiento alimentario y en la situación nutricional de las personas. En consecuencia, los jóvenes adultos son, en muchos aspectos, un grupo de interés particular, ya que varios patrones aprendidos a temprana edad pueden contribuir a sus resultados de salud a largo plazo. A través de la segmentación de los casos de base relevantes, se puede recopilar información valiosa para el desarrollo de intervenciones centradas en solucionar las desigualdades en salud y nutrición. En este sentido, el nivel socioeconómico ha sido identificado como un determinante crítico en lo que respecta a la calidad y cantidad ingerida de alimentos (Mukoma et al. 2023).

Los individuos clasificados en un mayor nivel socioeconómico generalmente pueden acceder a alimentos más nutritivos y, consecuentemente, poseen mejores patrones alimentarios, a diferencia de aquellos pertenecientes a un NSE bajo, quienes se ven forzados a consumir productos ultraprocesados y de alta densidad calórica. Esto no solo modifica la cantidad de calorías ingeridas, sino que también involucra una alteración en la composición corporal, lo cual se manifiesta a nivel del índice de masa corporal y en una mayor prevalencia de sobrepeso y obesidad (Vos et al. 2022).

En general, la ingesta calórica y la composición corporal se consideran indicadores básicos para evaluar el estado nutricional de los jóvenes adultos. La ingesta calórica total es crítica para garantizar un balance energético saludable y, en última instancia, un peso corporal apropiado. No obstante, más allá de la cantidad de calorías consumida, la dieta y la ingesta de proteínas y carbohidratos también son importantes para mantener la salud metabólica y la calidad de la composición corporal. Sin la ingesta de proteínas adecuada, el metabolismo y la propensión a enfermedades crónicas pueden ser más frecuentes, lo que justifica la necesidad de evaluar estos factores de manera integral. La herramienta de segmentación es un instrumento esencial para comprender mejor los mecanismos subyacentes relacionados con la salud y la nutrición (Zaccagni et al. 2020).

El poder de la segmentación radica en su capacidad de identificar grupos homogéneos dentro de una población diversa, permitiendo la implementación efectiva de intervenciones diseñadas para sus necesidades. Este aspecto es esencial para la salud pública, ya que permite la optimización de la estrategia de intervención (Bekbolatova et al. 2024).

El algoritmo K-means es una técnica estadística adecuada que se puede aplicar para segmentar a adultos jóvenes mediante NSE, ingesta calórica y valores de IMC. K-means es un algoritmo de agrupamiento que divide un conjunto de observaciones en “k” clústeres de similitud aproximada en patrones de conducta o jerarquías. Garner. La aplicación de K-means permite identificar claramente a una serie de individuos con

patrones dietéticos de comportamiento similares y a un patrón de consumo que puede ser fundamental para comprender la forma en que los valores de NSE influyen en la ingesta calórica y los IMC (Qasrawi y Abu Al-Halawa 2022).

La capacidad de K-means para identificar clústeres significativos en grandes conjuntos de datos hace que sea una herramienta particularmente útil en la investigación en nutrición y salud pública. A través de este enfoque, los investigadores pueden no solo describir las diferencias en el comportamiento alimentario entre grupos, sino también comprender las características subyacentes que contribuyen a estas diferencias. Esto es esencial para el diseño de programas de intervención más efectivos y dirigidos, que se adapten a las necesidades específicas de cada segmento poblacional (Kirk et al. 2022).

La segmentación de jóvenes adultos en función de su nivel socioeconómico, ingesta calórica y composición corporal es un enfoque prometedor para abordar las disparidades en salud y nutrición. El uso del algoritmo K-means ofrece una metodología robusta para identificar patrones y relaciones que pueden informar el desarrollo de intervenciones dirigidas. Al comprender mejor cómo se relacionan estos factores, los profesionales de la salud y los investigadores pueden colaborar para implementar estrategias que promuevan la salud y el bienestar en esta población crucial, contribuyendo así a mejorar los resultados de salud pública en general (Jones et al. 2020).

2. METODOLOGÍA O MATERIALES Y METODOS

2.1. Enfoque de la investigación

En este estudio transversal no probabilístico por conveniencia, se analizarán estudiantes universitarios utilizando el algoritmo K-means para segmentarlos según su nivel socioeconómico, ingesta calórica y composición corporal. El análisis integrará variables cualitativas y cuantitativas, como el índice de masa corporal y la proporción de macronutrientes en la dieta. La identificación de clústeres con características homogéneas permitirá explorar patrones de comportamiento nutricional adaptados a perfiles específicos. Este enfoque segmentado facilitará el desarrollo de intervenciones en nutrición y salud pública, dirigidas a mejorar la salud de la población estudiantil (Pobee 2022).

2.2. Unidades de análisis

La unidad de análisis de este estudio está compuesta por 150 estudiantes de la carrera de nutrición de la Universidad Estatal de Milagro. Estos jóvenes adultos, en su mayoría entre 18 y 25 años, representan una muestra diversa en términos de nivel socioeconómico, hábitos alimentarios y composición corporal. La selección de este grupo permite explorar las interrelaciones entre la ingesta calórica, el índice de masa corporal (IMC) y el nivel socioeconómico en una población que está en formación profesional en el ámbito de la nutrición, lo que puede aportar insights relevantes sobre su estado nutricional y su futuro como profesionales en el área (Okeyo et al. 2020). Esta investigación forma parte del proyecto PRIVACE-SALUD, cuyo nombre significa "Proyecto de Investigación para Valorar la Composición y Estructura Corporal", busca

analizar las diferencias en la estructura corporal, estado de salud y dieta de adultos ecuatorianos, según la región geográfica en la que residen.. El proyecto tuvo una duración de 18 meses, llevándose a cabo entre enero de 2022 y diciembre de 2023 entre las Universidad Estatal de Milagro y la Escuela Superior Politécnica de Chimborazo.

2.3. Técnicas de recolección

Para la recolección de datos, se utilizaron cuestionarios validados que evaluaron diferentes aspectos del estado nutricional, incluyendo la ingesta calórica, la composición de macronutrientes y el nivel socioeconómico de los participantes. Estos cuestionarios fueron aplicados para garantizar la precisión y la confiabilidad de las respuestas. La administración de los cuestionarios se llevó a cabo de manera presencial, con la supervisión de investigadores capacitados, lo que permitió aclarar dudas y asegurar la correcta comprensión de las preguntas por parte de los estudiantes. Además, el estudio cuenta con la aprobación del comité de ética institucional, garantizando el cumplimiento de las normativas éticas en la investigación para seres humanos (Washington et al. 2020).

2.4. Procesamiento y análisis de la información

El análisis estadístico se realizó en tres niveles: univariado, bivariado y multivariado. En el análisis univariado, se describieron las características demográficas y nutricionales de la muestra, utilizando medidas de tendencia central y dispersión. Luego, el análisis bivariado exploró las relaciones entre variables, identificando asociaciones significativas entre el nivel socioeconómico, la ingesta calórica y el índice de masa corporal (IMC)(Moyer et al. 2020).

Para el análisis multivariado, se aplicó el algoritmo K-means, segmentando a los jóvenes adultos en grupos homogéneos según sus características nutricionales y socioeconómicas. Se utilizó el criterio del método del codo para seleccionar el número óptimo de clústeres, con una métrica de distancia euclidiana y un proceso de inicialización k-means++ para mejorar la convergencia. Además, se ajustaron parámetros como el número máximo de iteraciones y las condiciones de convergencia. Todo el análisis se realizó utilizando Python en la plataforma Google Colab, lo que facilitó tanto el procesamiento como la visualización de los datos(Baptista 2021).

3. RESULTADOS Y DISCUSIÓN

Resultados:

Los resultados del análisis de la muestra de estudiantes reflejan variaciones significativas en el estado nutricional, el nivel socioeconómico y la ingesta calórica. Al observar el desglose por género, se aprecia que los hombres tienen una mayor proporción de estudiantes en las categorías "por debajo del rango normal" y "superior al rango normal" en comparación con las mujeres, con un notable 44.67% de hombres en la

categoría superior al rango normal. Esto sugiere una mayor prevalencia de sobrepeso u obesidad en hombres, lo que puede ser relevante para el diseño de intervenciones específicas.

En cuanto al nivel socioeconómico, los datos indican que la mayoría de los estudiantes con un nivel socioeconómico bajo tienden a estar "por debajo del rango normal" en comparación con aquellos de nivel medio bajo y medio típico, donde se observa una mayor proporción de estudiantes en el rango normal. La falta de acceso a recursos adecuados y alimentos nutritivos en este grupo puede estar contribuyendo a un peor estado nutricional.

Respecto a la evaluación de la ingesta calórica, se observa que hay un aumento significativo en la tendencia a un exceso de carbohidratos (CHO), con un 27.33% de estudiantes en la categoría superior al rango normal. Esto puede reflejar una alimentación poco equilibrada y el consumo excesivo de alimentos ricos en azúcares y carbohidratos simples. Por el contrario, el déficit de grasas y proteínas se presenta en proporciones menores, lo que sugiere que, aunque hay un exceso de CHO, los estudiantes no están necesariamente en desnutrición por proteínas o grasas.

En el análisis del índice de masa corporal (IMC), la mayoría de los estudiantes se clasifican como normopeso (34.00% en la categoría superior al rango normal), lo que es alentador. Sin embargo, se observa una preocupación en los niveles de sobrepeso y obesidad, especialmente en las categorías de sobrepeso (12.00%) y obesidad I (2.00%), lo que sugiere la necesidad de una atención continua a la salud y la nutrición en esta población. Estos hallazgos resaltan la importancia de intervenciones educativas y programas de nutrición que fomenten hábitos alimentarios más equilibrados y saludables, especialmente entre los grupos más vulnerables. En la Tabla 1 se presenta un resumen general de los resultados de las variables cualitativas frente a la variable de interés.

Tabla 1. Diagnóstico del porcentaje de adecuación y su relación con las variables de interés.

		En el rango normal	Por debajo del rango normal	Superior al rango normal
Género	H	5.33	5.33	6.00
	M	14.00	24.67	44.67
	ALTO	0.67	2.67	0.67
Nivel socioeconómico	BAJO	2.00	5.33	4.67
	MEDIO ALTO	3.33	2.67	8.00
	MEDIO BAJO	6.67	8.67	21.33
	MEDIO TIPICO	6.67	10.67	16.00
	Con tendencia a Deficid de CHO	0.00	0.00	0.67
	Con tendencia a Deficid de P	0.00	0.00	1.33
	Con tendencia a Deficid de G	4.00	6.67	6.00
Evaluación de la ingesta calórica	Con tendencia a exceso de CHO	10.67	12.00	27.33
	Con tendencia a Exceso de G	0.00	0.67	2.00
	Con tendencia a Exceso de P	2.67	8.67	9.33
	Equilibrio	2.00	2.00	4.00

	BAJO PESO	0.67	0.67	2.67
	NORMO PESO	12.00	10.00	34.00
IMC	OBESIDAD I	0.00	6.67	2.00
	OBESIDAD II	0.00	2.00	0.00
	SOBREPESO	6.67	10.67	12.00

El análisis descriptivo de las variables relacionadas con el perfil nutricional de los participantes (n=150). La edad promedio fue de 22.10 años (± 2.38), con un rango que va desde los 19 hasta los 35 años. El promedio de GET ingerido (Gasto Energético Total) fue de 2405.70 kcal (± 4616.97), lo que refleja una variabilidad considerable, evidenciada en un rango amplio de 562.59 a 46916.71 kcal. El GET recomendado mostró un promedio de 1723.73 kcal (± 302.22). El porcentaje de adecuación, que indica el cumplimiento del GET ingerido respecto al recomendado, tuvo un valor medio de 1.40 (± 2.51), con valores que varían entre 0.27 y 26.40.

En cuanto a la ingesta de macronutrientes, el consumo de proteínas promedio fue de 139.33 g (± 269.56), mientras que el porcentaje de aporte proteico fue del 0.29 (± 0.18). La ingesta de grasas totales fue de 67.70 g (± 91.71), y el porcentaje de aporte de grasa al total energético fue del 0.15 (± 0.09). Los carbohidratos tuvieron un promedio de ingesta de 257.17 g (± 306.82), y su aporte porcentual al total energético fue del 0.56 (± 0.23).

El peso promedio de los participantes fue de 61.45 kg (± 13.90), y la talla media fue de 1.59 m (± 0.09). El Índice de Masa Corporal (IMC) promedio fue de 24.22 kg/m² (± 4.51), lo cual sugiere que la mayoría de los participantes se encuentran en un rango de IMC considerado como peso normal, aunque con algunos casos que podrían clasificarse en categorías de sobrepeso y obesidad (IMC máximo de 38.17 kg/m²). Estos resultados reflejan una variabilidad considerable en los perfiles de ingesta y estado nutricional de los participantes. En la Tabla 2, se presenta las medidas de tendencia central y variabilidad de las variables de estudio.

Tabla 2. Medias de tendencia central y variabilidad de las variables que intervienen en el estudio.

	Media	Std	Min	25%	50%	75%	Max
EDAD	22.10	2.38	19.00	20.25	22.00	23.00	35.00
GET INGERIDO	2405.70	4616.97	562.59	1506.81	1838.02	2219.29	46916.71
GET RECOMENDADO	1723.73	302.22	1233.77	1482.49	1674.77	1869.25	2834.40
% DE ADECUACION	1.40	2.51	0.27	0.85	1.12	1.34	26.40
PROT	139.33	269.56	27.23	65.58	85.08	122.62	3197.06
% de P	0.29	0.18	0.09	0.18	0.22	0.29	0.84
GRASA T	67.70	91.71	8.34	35.27	50.01	78.69	1031.39
% de G	0.15	0.09	0.02	0.10	0.14	0.19	0.71
CARBOH	257.17	306.82	12.70	126.39	220.29	298.98	3529.30
% de CHO	0.56	0.23	0.06	0.53	0.63	0.71	1.00
TOTAL	462.77	620.16	119.64	316.00	390.49	499.12	7757.75
PESO	61.45	13.90	36.20	50.30	60.15	69.15	110.00
TALLA	1.59	0.09	1.39	1.53	1.58	1.63	1.82

IMC	24.22	4.51	15.20	20.54	23.47	27.09	38.17
-----	-------	------	-------	-------	-------	-------	-------

El análisis de clústeres realizado mediante K-Means destaca significativas divergencias en los patrones nutricionales entre estudiantes universitarios, subrayando cómo las diferencias socioeconómicas impactan en el gasto energético total (GET). Una elevada desviación estándar en el GET ingerido (4616.97 kcal) revela una distribución heterogénea en el consumo energético, con estudiantes de estratos socioeconómicos altos consumiendo más calorías, posiblemente debido a una disponibilidad mayor de alimentos densamente calóricos y estilos de vida que implican un consumo energético elevado. Este fenómeno sugiere una correlación entre el nivel socioeconómico y el acceso a dietas variadas y nutricionalmente ricas.

Por otro lado, el índice de masa corporal (IMC) promedio de 24.22 cae dentro del rango normal, pero su variabilidad (std 4.51) y el máximo observado de 38.17 indican casos de obesidad, especialmente en clústeres con baja adecuación calórica donde la calidad de la dieta puede ser deficiente. El análisis de macronutrientes refleja una alta variabilidad en la ingesta de proteínas y grasas, asociada a diferentes niveles de conocimiento nutricional y acceso a alimentos saludables. Este detalle enfatiza la importancia de implementar políticas nutricionales que promuevan el acceso equitativo a alimentos saludables y educación nutricional, especialmente dirigidas a estratos socioeconómicos bajos para reducir la disparidad en la calidad de la dieta y mejorar la salud general de la población estudiantil.

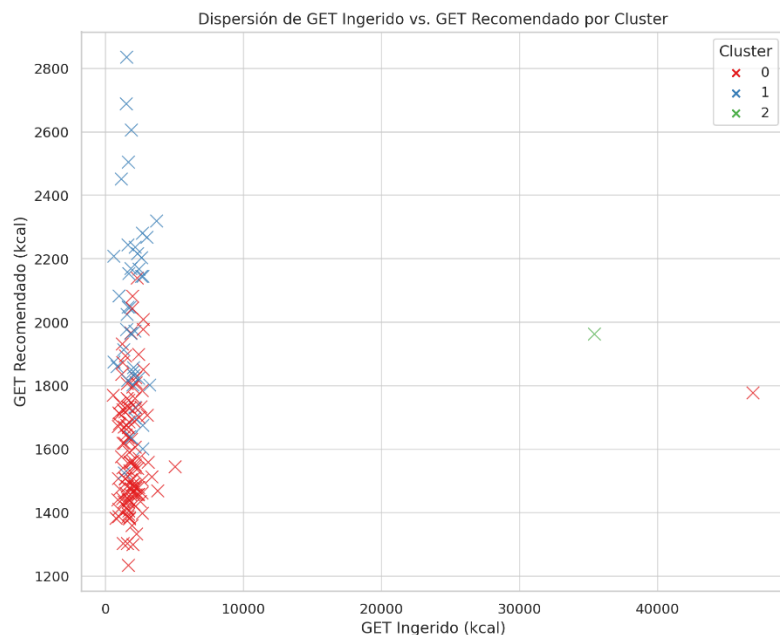
El análisis de K-Means con 3 clusters ha revelado patrones distintos en los hábitos alimenticios y el estado nutricional de los participantes:

- Cluster 0: Representa un grupo con un consumo energético ligeramente superior al recomendado, con un % de adecuación promedio de 1.41, indicando un ligero exceso. La composición de macronutrientes muestra un balance con predominancia de carbohidratos. Este grupo tiene un IMC promedio en el rango normal, sugiriendo que el ligero exceso energético no se traduce necesariamente en sobrepeso.
- Cluster 1: Este grupo consume menos energía de la recomendada, con un % de adecuación de 0.97. Sin embargo, tienen un IMC promedio en el rango de sobrepeso, lo que podría indicar ineficiencias en el metabolismo o en la calidad de los alimentos consumidos, dado que también presentan un alto consumo de proteínas.
- Cluster 2: Este grupo se destaca por un consumo extremadamente alto de energía y macronutrientes, particularmente proteínas y grasas, con un porcentaje de adecuación de 18.05, indicativo de la presencia de deportistas de alto rendimiento o individuos con condiciones médicas especiales que requieren una mayor ingesta calórica. Este patrón, combinado con un IMC promedio inusualmente bajo para el alto consumo energético, sugiere un grupo de individuos con alta actividad física.

La técnica de K-Means permite segmentar la población en grupos con características nutricionales claramente diferenciadas, lo que facilita la identificación de comportamientos alimenticios específicos y

potenciales áreas de intervención nutricional. Este enfoque permite ajustar recomendaciones y políticas de salud pública a las necesidades de subgrupos dentro de una población más grande (Ikotun et al. 2023).

Figura 1. Representación gráfica de la segmentación mediante K-Means



Discusiones:

Los resultados obtenidos en este estudio coinciden con hallazgos previos que destacan la relación entre el nivel socioeconómico, la ingesta calórica y el estado nutricional en poblaciones jóvenes. Se ha observado que los estudiantes universitarios de menores niveles socioeconómicos presentan una mayor prevalencia de obesidad y sobrepeso, sugiriendo que las limitaciones en el acceso a alimentos saludables están directamente relacionadas con un deterioro en su estado nutricional, según un estudio realizado por (Telleria-Aramburu y Arroyo-Izaga 2022).

Por otra parte, el trabajo de (Wan et al. 2023) también destaca que la ingesta calórica excesiva, particularmente de carbohidratos, se asocia con el aumento de peso y obesidad en jóvenes adultos. Esto es consistente con nuestros hallazgos, donde un 27.33% de los estudiantes mostró una tendencia hacia el exceso de CHO. Este consumo inadecuado podría atribuirse a una alta disponibilidad de alimentos ultraprocesados y azucarados en entornos urbanos, como indican los estudios de Yki-Järvinen et al. (2021), quienes encontraron que la dieta de los jóvenes adultos a menudo se caracteriza por una alta proporción de carbohidratos simples y grasas saturadas.

Adicionalmente, la investigación de Al-Awwad et al. (2021) señala que la educación nutricional es relevante para mejorar los hábitos alimentarios entre los estudiantes universitarios. En este estudio, el 34% de los participantes se clasificó como normopeso, lo que sugiere un potencial para fomentar prácticas de alimentación saludables, pero también indica que se requiere atención a los estudiantes con sobrepeso y obesidad.

Los resultados muestran variabilidad en la ingesta energética y de macronutrientes, sugiriendo patrones alimentarios diversos entre los participantes. Según Westphalen Palma et al. (2021), un alto % de adecuación podría relacionarse con un riesgo aumentado de sobrepeso, reflejado en el IMC máximo de 38.17 kg/m². Además, la dispersión en la ingesta de proteínas, grasas y carbohidratos concuerda con lo descrito por San-Cristobal et al. (2020), quienes destacan la influencia de la dieta en la composición corporal. Estos hallazgos resaltan la necesidad de intervenciones nutricionales personalizadas para mejorar la calidad de la dieta y prevenir el exceso de peso.

El análisis de K-Means con 3 clusters ha revelado patrones distintos en los hábitos alimenticios y el estado nutricional de los participantes:

- Cluster 0: Representa un grupo con un consumo energético ligeramente superior al recomendado, con un % de adecuación promedio de 1.41, indicando un ligero exceso. La composición de macronutrientes muestra un balance con predominancia de carbohidratos. Este grupo tiene un IMC promedio en el rango normal, sugiriendo que el ligero exceso energético no se traduce necesariamente en sobrepeso.
- Cluster 1: Este grupo consume menos energía de la recomendada, con un % de adecuación de 0.97. Sin embargo, tienen un IMC promedio en el rango de sobrepeso, lo que podría indicar ineficiencias en el metabolismo o en la calidad de los alimentos consumidos, dado que también presentan un alto consumo de proteínas.
- Cluster 2: Este cluster parece contener datos atípicos, con un consumo extremadamente alto de energía y macronutrientes, especialmente proteínas y grasas. El % de adecuación es muy elevado (18.05), y el IMC promedio es inesperadamente bajo considerando el alto consumo energético, lo que podría deberse a errores en los datos o casos muy específicos como deportistas de alto rendimiento o condiciones médicas (Wells et al. 2023).

La técnica de K-Means facilita la segmentación de la población en grupos con características nutricionales distintivas, permitiendo identificar comportamientos alimenticios específicos y áreas potenciales para intervenciones nutricionales. Este enfoque posibilita la adaptación de recomendaciones y políticas de salud pública a las necesidades particulares de subgrupos dentro de una población más amplia. (Kwon et al. 2020).

El análisis de K-Means realizado sobre los datos nutricionales ha revelado tres clusters distintos, cada uno caracterizado por patrones de consumo y estados nutricionales específicos. El Cluster 0 se caracteriza por un ligero exceso en la ingesta energética con respecto a lo recomendado, mientras que el Cluster 1 muestra un déficit, y el Cluster 2 destaca por extremos inusuales en todas las métricas evaluadas, sugiriendo posibles errores en los datos o presencia de casos atípicos como deportistas de alto rendimiento (Ru et al. 2021).

El Cluster 0, con una adecuación energética ligeramente por encima de lo recomendado (1.41), incluye individuos que, aunque consumen más energía de la necesaria, mantienen un IMC dentro del rango normal. Este fenómeno podría estar relacionado con una mayor actividad física o diferencias metabólicas que

permiten un manejo más eficiente del exceso energético, una observación que concuerda con los estudios de Afonso et al. (2021), quienes indicaron que la relación entre el consumo energético y el IMC no es directa y está influenciada por múltiples factores como la composición de la dieta y la actividad física.

Por otro lado, el Cluster 1 muestra un consumo energético por debajo del recomendado con un IMC promedio que cae en el rango de sobrepeso. Este aparente contrasentido puede explicarse por una calidad nutricional deficiente, como sugieren O'Sullivan, Schmidt, y Kratz (2020), destacando que no solo la cantidad sino también la calidad de los macronutrientes es fundamental para el manejo adecuado del peso y la salud general.

El Cluster 2, con su notable alta ingesta de energía y macronutrientes, plantea la posibilidad de errores de medición o la inclusión de individuos con necesidades energéticas excepcionales, como atletas de alto rendimiento o personas con condiciones metabólicas únicas. Esta anomalía destaca la importancia crítica de verificar la calidad de los datos en investigaciones nutricionales y sugiere la necesidad de considerar la exclusión de valores atípicos para mejorar la precisión analítica (Posma et al. 2020).

4. CONCLUSIONES

En conclusión, este estudio resalta la necesidad de implementar intervenciones educativas dirigidas a los jóvenes adultos, especialmente en aquellos de niveles socioeconómicos bajos, para promover hábitos alimentarios equilibrados y combatir la obesidad. Las políticas de salud pública deben enfocarse en mejorar el acceso a alimentos nutritivos y en la promoción de estilos de vida saludables, teniendo en cuenta las diferencias socioeconómicas que influyen en la alimentación y el estado nutricional.

Los resultados evidencian una variabilidad significativa en la ingesta calórica y de macronutrientes entre los participantes, lo cual podría estar vinculado a diferentes riesgos de salud, como el sobrepeso. Es fundamental considerar la individualidad en los hábitos alimenticios para desarrollar estrategias nutricionales personalizadas que promuevan una alimentación equilibrada y adecuada a las necesidades de cada persona, mejorando su estado nutricional y bienestar.

El uso de K-Means ha permitido identificar patrones nutricionales que podrían ser cruciales para el diseño de intervenciones dirigidas específicamente a cada grupo, optimizando así las estrategias de salud pública y nutricional. Esta técnica demuestra ser una herramienta valiosa en la epidemiología nutricional para segmentar poblaciones basadas en características de ingesta y metabólicas, facilitando enfoques más personalizados en nutrición.

FINANCIACIÓN

Los autores deben declarar la fuente de financiación de su investigación, estas pueden ser organizaciones gubernamentales, universidades, centros de investigación, becas, proyectos aprobados entre otros. Por otra parte, si la investigación no tuvo financiamiento también debe indicarlo

CONFLICTO DE INTERESES

Los Autores declaran si existen o no conflicto de intereses con su investigación

CONTRIBUCIÓN DE AUTORÍA

En concordancia con la taxonomía establecida internacionalmente para la asignación de créditos a autores de artículos científicos (<https://credit.niso.org/>). Los autores declaran sus contribuciones en la siguiente matriz:

Participar activamente en:	Autor 1.	Autor 2	Autor 3	Autor 3
Conceptualización	X	X	X	X
Análisis formal	X	X	X	X
Adquisición de fondos	X	X		
Investigación	X	X	X	X
Metodología	X			
Administración del proyecto	X	X	X	X
Recursos	X	X	X	X
Redacción –borrador original	X	X	X	X
Redacción –revisión y edición	X	X	X	X
La discusión de los resultados	X	X	X	X
Revisión y aprobación de la versión final del trabajo.	X	X	X	X

REFERENCIAS

- Afonso, Cláudia, Ana Rita Sousa-Santos, Alejandro Santos, Nuno Borges, Patrícia Padrão, Pedro Moreira, y Teresa F. Amaral. 2021. «Frailty status is related to general and abdominal obesity in older adults». *Nutrition Research* 85:21-30.
- Al-Awwad, Narmeen Jamal, Hiba Fathi Al-Sayyed, Zeinah Abu Zeinah, y Reema Fayez Tayyem. 2021. «Dietary and lifestyle habits among university students at different academic years». *Clinical Nutrition ESPEN* 44:236-42.
- Baptista, Leonardo. 2021. «Using Python and Google Colab to Teach Physical Chemistry During Pandemic».
- Bekbolatova, Molly, Jonathan Mayer, Chi Wei Ong, y Milan Toma. 2024. «Transformative potential of AI in Healthcare: definitions, applications, and navigating the ethical Landscape and Public perspectives». P. 125 en *Healthcare*. Vol. 12. MDPI.
- Ikotun, Abiodun M., Absalom E. Ezugwu, Laith Abualigah, Belal Abuhajja, y Jia Heming. 2023. «K-means clustering algorithms: A comprehensive review, variants analysis, and advances in the era of big data». *Information Sciences* 622:178-210.
- Jones, Petra J., Matthew K. James, Melanie J. Davies, Kamlesh Khunti, Mike Catt, Tom Yates, Alex V. Rowlands, y Evgeny M. Mirkes. 2020. «FilterK: A new outlier detection method for k-means clustering of physical activity». *Journal of biomedical informatics* 104:103397.
- Kirk, Daniel, Esther Kok, Michele Tufano, Bedir Tekinerdogan, Edith JM Feskens, y Guido Camps. 2022. «Machine learning in nutrition research». *Advances in Nutrition* 13(6):2573-89.
- Kwon, Yu-Jin, Hyoung Sik Kim, Dong-Hyuk Jung, y Jong-Koo Kim. 2020. «Cluster analysis of nutritional factors associated with low muscle mass index in middle-aged and older adults». *Clinical Nutrition* 39(11):3369-76.
- Moyer, Cheryl A., Sarah D. Compton, Elizabeth Kaselitz, y Maria Muzik. 2020. «Pregnancy-Related Anxiety during COVID-19: A Nationwide Survey of 2740 Pregnant Women». *Archives of Women's Mental Health* 23(6):757-65. doi: 10.1007/s00737-020-01073-5.
- Mukoma, Gudani, Stephanie V. Wrottesley, Juliana Kagura, Tolu Oni, Lisa Micklesfield, y Shane A. Norris. 2023. «The Relationships between Socioeconomic Status, Dietary Knowledge and Patterns, and Physical Activity with Adiposity in Urban South African Women». *South African Journal of Clinical Nutrition* 36(2):56-62. doi: 10.1080/16070658.2022.2076374.
- Okeyo, Alice P., Eunice Seekoe, Anniza de Villiers, Mieke Faber, Johanna H. Nel, y Nelia P. Steyn. 2020. «The food and nutrition environment at secondary schools in the Eastern Cape, South Africa as reported by learners». *International journal of environmental research and public health* 17(11):4038.
- O'Sullivan, Therese A., Kelsey A. Schmidt, y Mario Kratz. 2020. «Whole-fat or reduced-fat dairy product intake, adiposity, and cardiometabolic health in children: a systematic review». *Advances in Nutrition* 11(4):928-50.
- Pobee, Frederick. 2022. «Non-Probabilistic Approach to e-Banking Adoption: The Moderating Impact of Trialability». *Management and Labour Studies* 47(2):183-98. doi: 10.1177/0258042X211054248.

- Posma, Joram M., Isabel Garcia-Perez, Gary Frost, Ghadeer S. Aljuraiban, Queenie Chan, Linda Van Horn, Martha Daviglius, Jeremiah Stamler, Elaine Holmes, y Paul Elliott. 2020. «Nutriome–metabolome relationships provide insights into dietary intake and metabolism». *Nature food* 1(7):426-36.
- Qasrawi, Radwan, y Diala Abu Al-Halawa. 2022. «Cluster analysis and classification model of nutritional anemia associated risk factors among Palestinian schoolchildren, 2014». *Frontiers in Nutrition* 9:838937.
- Ru, Yuan, Ninglin Wang, Yan Min, Xuemiao Wang, Valerie McGurie, Meng Duan, Xiaochen Xu, Xueyin Zhao, Yi-Hsuan Wu, y Ying Lu. 2021. «Characterization of dietary patterns and assessment of their relationships with metabolomic profiles: A community-based study». *Clinical Nutrition* 40(5):3531-41.
- San-Cristobal, Rodrigo, Santiago Navas-Carretero, Miguel Ángel Martínez-González, José María Ordovas, y José Alfredo Martínez. 2020. «Contribution of macronutrients to obesity: implications for precision nutrition». *Nature Reviews Endocrinology* 16(6):305-20.
- Telleria-Aramburu, Nerea, y Marta Arroyo-Izaga. 2022. «Risk factors of overweight/obesity-related lifestyles in university students: Results from the EHU12/24 study». *British Journal of Nutrition* 127(6):914-26.
- Vos, Marjolijn, Benedicte Deforche, Anneleen Van Kerckhove, Nathalie Michels, Maartje Poelman, Maggie Geuens, y Wendy Van Lippevelde. 2022. «Determinants of healthy and sustainable food choices in parents with a higher and lower socioeconomic status: A qualitative study». *Appetite* 178:106180.
- Wan, Yi, Deirdre K. Tobias, Kristine K. Dennis, Marta Guasch-Ferré, Qi Sun, Eric B. Rimm, Frank B. Hu, David S. Ludwig, Orrin Devinsky, y Walter C. Willett. 2023. «Association between changes in carbohydrate intake and long term weight changes: prospective cohort study». *bmj* 382.
- Washington, Simon, Matthew G. Karlaftis, Fred Mannering, y Panagiotis Anastasopoulos. 2020. *Statistical and econometric methods for transportation data analysis*. Chapman and Hall/CRC.
- Wells, Kimberley, Nikki Jeacocke, Renee Appaneal, Hilary Smith, N. Vlahovich, L. Burke, y D. Hughes. 2023. «DISORDERED EATING IN HIGH PERFORMANCE SPORT POSITION STATEMENT».
- Westphalen Palma, Shelly, Greisse Viero da Silva Leal, Ângelo José Gonçalves Bós, y Loiva Beatriz Dallepiane. 2021. «Factors associated with eating behavior of young-old and oldest-old people from Southern Brazil.» *Saúde e Pesquisa* 14.
- Yki-Järvinen, Hannele, Panu K. Luukkonen, Leanne Hodson, y J. Bernadette Moore. 2021. «Dietary carbohydrates and fats in nonalcoholic fatty liver disease». *Nature reviews Gastroenterology & hepatology* 18(11):770-86.
- Zaccagni, Luciana, Natascia Rinaldo, Barbara Bramanti, Jessica Mongillo, y Emanuela Gualdi-Russo. 2020. «Body Image Perception and Body Composition: Assessment of Perception Inconsistency by a New Index». *Journal of Translational Medicine* 18(1):20. doi: 10.1186/s12967-019-02201-1.