Análisis Multivariado y Clusterización de Productos para la Optimización de la Producción de Alimento Balanceado de una Empresa del Sector Acuícola

Melany Mayeline Vaca Vaca
Marco Vinicio Chiluiza Pullupaxi
ESAI Business School
Inteligencia de Negocios y Ciencia de Datos
14 de junio 2025
Palabras Clave: Balanceado, Acuícola, Producción, Multivariado, Clusterización

A. Resumen Ejecutivo

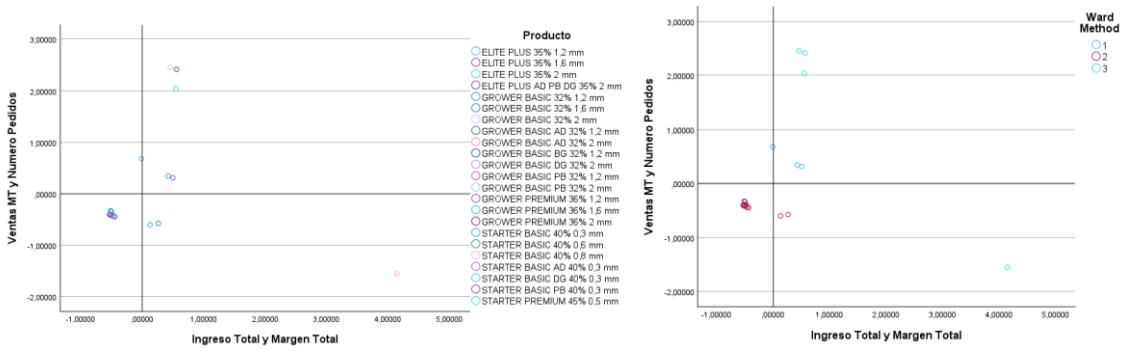
FeedSea es una empresa del sector acuícola dedicada a la producción y comercialización de alimentos balanceados para camarón. Ofrece a sus clientes un amplio portafolio de 23 productos distribuidos en cinco líneas principales, lo que le permite atender todas las fases del cultivo. Sin embargo, esta variedad ha ocasionado un uso ineficiente de recursos operativos, destinando parte de la capacidad de las líneas de producción a productos con baja rotación que no generan valor proporcional para el negocio.

Este estudio busca optimizar el portafolio mediante la identificación de grupos de productos con patrones similares según su desempeño comercial y técnico, que faciliten estrategias de producción más eficientes.

Se realizó un análisis descriptivo y exploratorio de la base de datos del periodo **2020 – 2024** para caracterizar las variables. Los resultados preliminares mostraron patrones trimestrales en el comportamiento de las ventas y rotación de los productos. Posteriormente, se aplicaron técnicas multivariadas como el Análisis Factorial y Componentes Principales para reducir la dimensionalidad de los datos e identificar las variables que más influyen en la demanda de los productos. Se identificaron dos factores principales que explican más del 90% del comportamiento de los datos y mediante la técnica de clúster jerárquico se identificaron tres grupos principales de productos: Grupo 1: Rotación, demanda y margen altos; grupo 2: Rotación, demanda y margen medios, y grupo 3: Rotación, demanda y margen bajos. Con base en los hallazgos, se propone establecer una estrategia comercial para los productos que se encuentran en el grupo tres.

B. Resumen Ejecutivo Gráfico

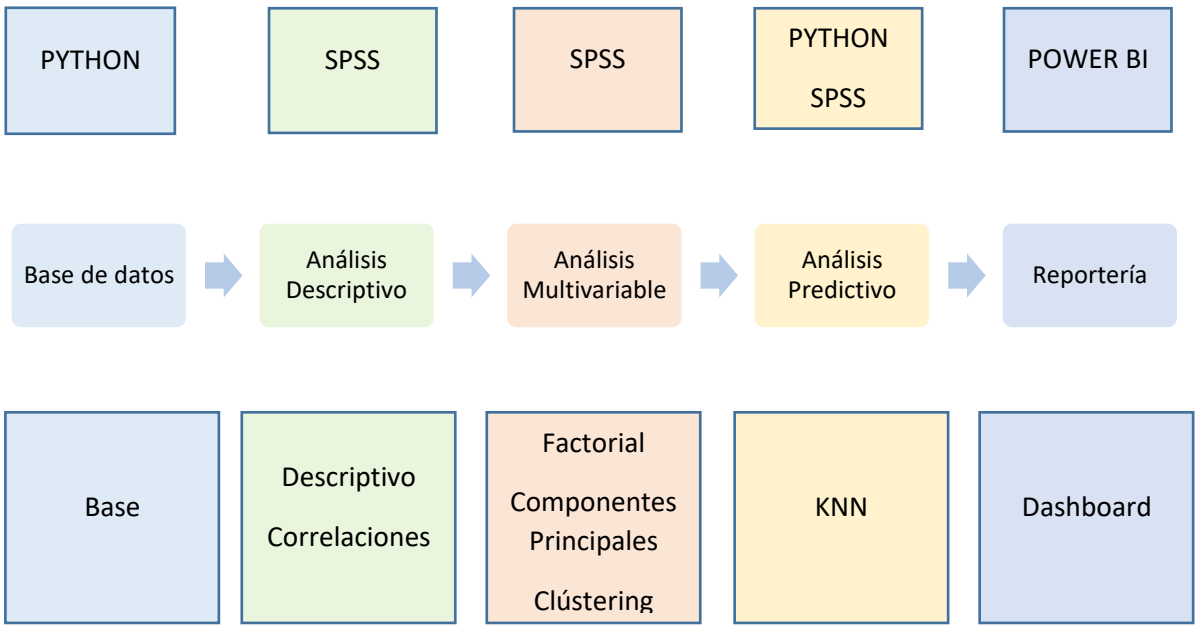
Figura 1. Portafolio de productos (a) y clusterización (b) para el segundo trimestre de cada año



(a) (b)

Nota: A) Eje horizontal ratio de ingreso total y margen, eje vertical ratio de toneladas métricas vendidas y número de pedidos.

Figura 2. Etapas, herramientas y entregables de la investigación



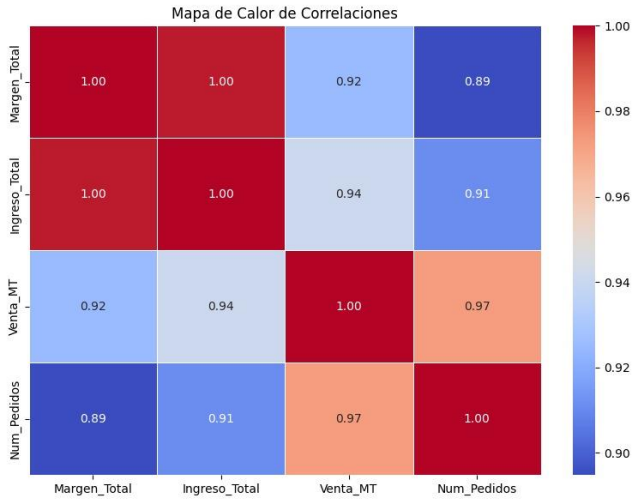
Del análisis llevado a cabo en el presente documento se identificó que con las variables comerciales analizadas de los productos producidos en los trimestres dos, Figura 3,

existe una categorización de tres segmentos: Alto volumen de venta trimestral – alta demanda – alto margen; medio volumen de venta trimestral – media demanda – medio margen y bajo volumen de venta trimestral – baja demanda – bajo margen (4, 3 y 16 Productos respectivamente). Ver Figura 1.

Para obtener el resultado fue necesario realizar un proceso inicial como estructuración de la base de datos, análisis descriptivo y multivariable. Posterior a la clusterización se ejecutó un análisis predictivo mediante la técnica KNN que permitirá determinar a futuras producciones trimestrales si el producto planificado forma parte de uno de los grupos identificados en los clústeres. Finalmente se ejecutó un reporte visual mediante la herramienta Power BI.

La Figura 2 detalla todo el proceso llevado a cabo en el proceso investigativo en el eje central se encuentra las etapas, en el eje superior las herramientas y softwares utilizados mientras que en el eje inferior los entregables o las evidencias de fin de cada etapa.

Figura 3. Mapa de Calor de Correlaciones: 4 Variables Comerciales



C. Detalle Técnico

1. Fuente de datos

1.1. Descripción y justificación

La base de datos fue generada mediante programación en Python, a partir de un análisis preliminar de información pública de las principales empresas del sector acuícola y tendencias del mercado, además de la experiencia directa en una empresa del rubro. Python fue el lenguaje de programación utilizado por ser gratuito y flexible.

Para la simulación de la base de datos se consideró un periodo de 60 meses, desde enero 2020 a diciembre 2024. La simulación abarcó la construcción de una base ficticia de clientes y un portafolio de productos dividido en dos grandes grupos: Estándar y Especiales, con subcategorías según la fase de desarrollo del animal.

Las ventas mensuales se generaron considerando una base anual ajustada por factores estacionales y distribuciones de la demanda de los productos. Los precios y costos fueron simulados dependiendo de la línea de producto, canal y región del cliente. Para lograr un mayor realismo, se incorporaron variaciones aleatorias mediante ruido gaussiano.

La base de datos cuenta con 23 columnas y 6000 registros de datos.

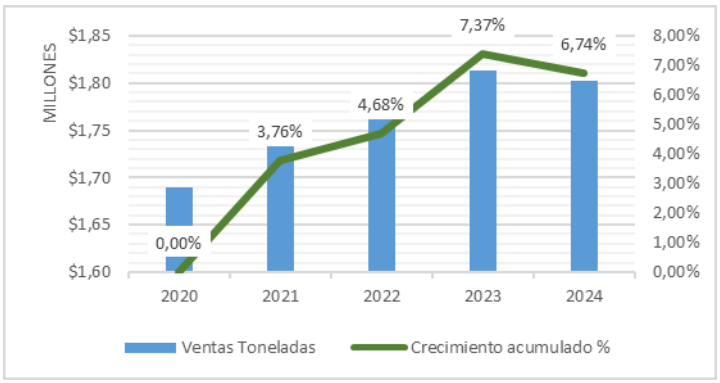
- Orden ID: El campo expone la clave de la transacción de venta; valor alfanumérico; tipo cadena.
- Fecha: Fecha de la transacción; tipo fecha.
- Cliente ID: Clave de identificación de clientes, se tiene una base de 500 Clientes, tipo cadena.
- Producto: Nombre comercial del producto; tipo cadena.
- Línea Fase: Categoría o segmento del producto; tipo cadena.
- Cliente Región: Ubicación geográfica o destino del cliente, tipo cadena.
- Canal: Dos tipos Directo o exportación, tipo cadena.
- Tipo Cliente: Grupo de segmentación del cliente, definido por la empresa, tipo cadena.
- Presentación KG: Se tiene dos presentaciones 25Kg y 15 Kg, tipo numérica.
- Sacos Vendidos: Número de sacos vendidos, tipo numérica.
- Ventas Toneladas: Volumen de toneladas vendidas; tipo numérica.
- Precio Unitario Saco: Precio de venta unitario del saco; tipo numérica.
- Costo Unitario Saco: Costo de fabricación del saco; tipo numérica.
- Ingreso Total: Ingreso de dinero por total de transacción, tipo numérica.
- Margen Total: Margen de ganancia por total de transacción, tipo numérica.

- Tamaño de Partícula MM: Tamaño del pienso o tamaño de la partícula o pienso, tipo numérica.
- Tipo de Procesamiento: Forma de procesar del producto, dos tipos peletizado o extruido, tipo cadena.
- Porcentaje de Grasa de Producto: Característica nutricional del producto, tipo numérica.
- Porcentaje de Proteína de Producto: Característica nutricional del producto, tipo numérica.
- Porcentaje de Fibra de Producto: Característica nutricional del producto, tipo numérica.
- Porcentaje de Ceniza de Producto: Característica nutricional del producto, tipo numérica.
- Ventas Totales: Volumen total de toneladas vendidas en la transacción, tipo numérica.
- Nombre de Producto Final: Nombre final del producto incluyendo información de calibre, tipo cadena.

1.2. Análisis exploratorio

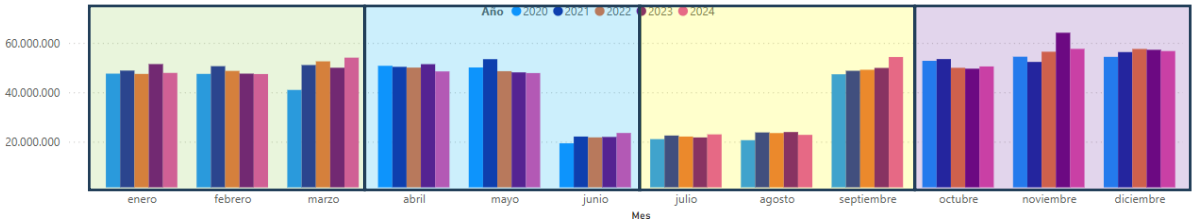
De acuerdo con el análisis de la información histórica, la empresa ha evolucionado a lo largo de los 5 años de operación incrementando sus volúmenes de venta. Como se observa en la Figura 4, las ventas han mostrado una tendencia de crecimiento moderado y sostenido en el tiempo a pesar de la presencia de ciertas fluctuaciones interanuales, siendo el crecimiento acumulado del 6.74% al año 2024.

Figura 4. Toneladas Métricas Vendidas por la Empresa (2020-2024)



Al analizar la evolución mensual del margen de ganancia, segmentada por mes, trimestre y año, se identifican patrones de comportamiento de venta en cada cuartil. De acuerdo con la Figura 5, el segundo trimestre inicia un periodo de venta bajista cuya tendencia se mantiene en el tercer trimestre, con una recuperación notable en el mes de septiembre. En contraste al primer y cuarto trimestre donde la tendencia de la venta es homogénea.

Figura 5. Margen de ganancia por mes y por año



De acuerdo con la Matriz de Correlaciones, existe una correlación fuerte y directa fuerte entre las ganancias totales y las ventas de toneladas métricas. Ver Tabla 1.

Tabla 1. Matriz de Correlación Margen Total vs Ventas Totales

Etiquetas	Detalle	Margen Total	Ventas MT
Margen Total	Correlación de Pearson	1	**0,712**
	Sig. (Bilateral)		<0,001
	N	6000	6000
Ventas MT	Correlación de Pearson	0,712	1
	Sig. (Bilateral)	<0,001	
	N	6000	6000

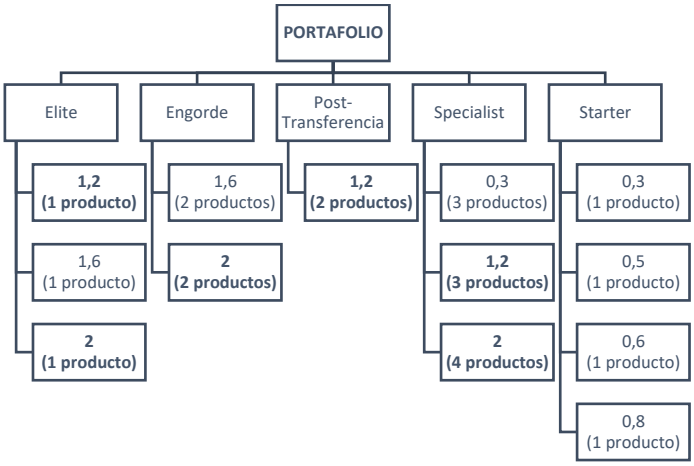
Al realizar el análisis por ubicación geográfica, la venta del producto balanceado acuícola se centra en dos puntos: Guayas y El Oro. La suma de ventas de toneladas métricas a clientes ubicados en estos dos puntos suma más del 50% de venta en los 5 años acorde a la información de la empresa, ver Anexo Figura A 1. Mientras que, el análisis por segmento de clientes identifica que el principal cliente es la camaronera industrial con una participación del 55 % del total de las ventas siendo el segmento predominante, seguido de la Camaronera mediana con el 24.03% de participación. En menor proporción se encuentran los clientes “Pequeño productor” y “Distribuidor”. Ver Anexo Figura A 2.

Por otra parte, de acuerdo con el análisis del portafolio, la empresa cuenta con veintitrés productos segmentados dentro de cinco categorías definidas por sus características nutricionales específicas, siendo las principales aquellas inherentes al producto como la

proteína, la grasa, las cenizas y fibra, al igual que características dimensionales como el tamaño de partícula.

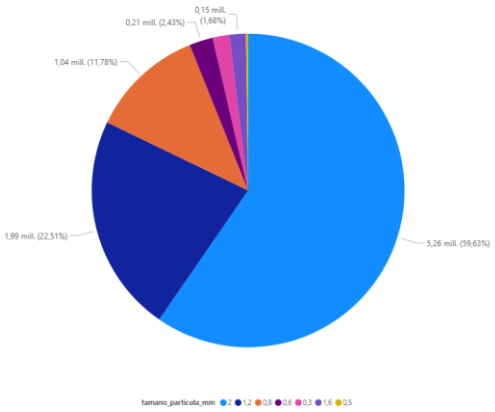
En la Figura 6 se presenta la segmentación del portafolio de productos, la disponibilidad de tamaños de partícula por segmento y la cantidad de productos correspondientes a cada uno de ellos. El análisis revela que los tamaños de partícula con más productos son los tamaños 2.0 mm y 1.2 mm con 7 y 6 productos en total, respectivamente.

Figura 6. Portafolio: Segmento y Tamaños de Partículas Disponibles por Segmento



Los productos con un tamaño de partícula 2.0 representan más del 50% de los pedidos de venta, siendo así el calibre predominante. Mientras que el segundo ítem con mayor participación es el calibre 1.2 mm que representa el 22,51%, seguido por el calibre 0.8 mm que alcanza el 11,78% de participación. Ver Figura 7

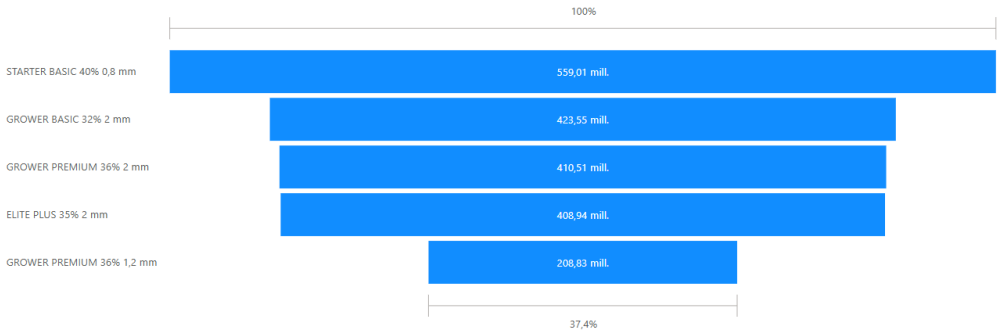
Figura 7. Volumen y % de Participación de Ventas 2020 - 2024



Sin embargo, al analizar por producto se identifica que para la partícula 0.8 mm se tiene una ventaja de planificación, debido a que con un solo producto es la tercera categoría

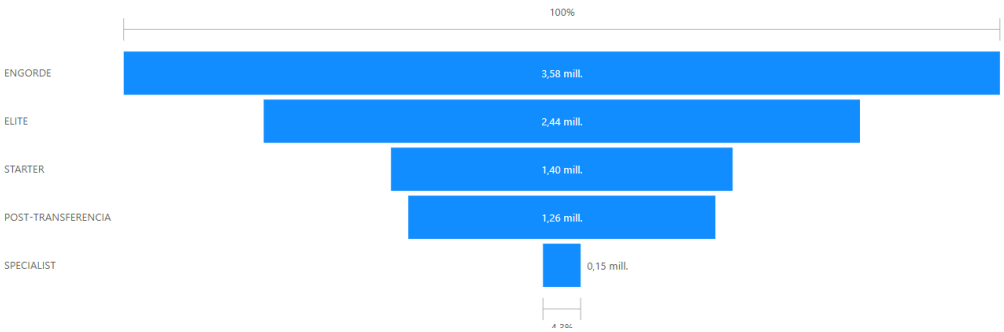
más vendida por la compañía y llega a ser el producto número uno según el margen de utilidad. A diferencia del pienso con tamaño de partícula de 2.0 mm, el cual, a pesar de tener siete productos, menos del 50% de sus productos se encuentran dentro del top 5 de productos con mayor margen de ganancia desde 2020 a 2024. Ver Figura 8.

Figura 8. TOP 5 de Productos con Mayor Margen desde 2020 a 2024



Finalmente, al analizar la distribución de las toneladas métricas vendidas por segmento del portafolio, se determina que las dietas de ENGORDE y ELITE concentran la mayor demanda, como se visualiza en la Figura 9. Esto sugiere una mayor preferencia de los clientes por productos destinados a las etapas finales del cultivo de camarón y aquellas que mejoran su rendimiento.

Figura 9. Ventas en MT de acuerdo con el segmento del producto.



Con base en los resultados del análisis, a nivel de producto se identifican dos oportunidades principales:

1. La identificación de los productos de mayor demanda y rentabilidad para el segundo y tercer trimestre dado que son periodos que presentan fluctuaciones estacionales, lo cual facilitará la programación de productos que mayor valor agregado generen a la empresa.
2. La optimización del portafolio de productos del tamaño de partículas de 2.0 mm, considerando la posibilidad de reducir la cantidad de productos disponibles

acorde a sus características técnicas. Estos dos análisis serán desarrollados en el presente trabajo de titulación

Estas oportunidades constituyen la de base para los análisis que se desarrollarán en los siguientes capítulos.

2. Técnica

Para realizar el análisis de las variables, se agruparon los valores por cada trimestre por los cinco años de análisis. Para fines de este estudio se consideró solo el Segundo Trimestre ya que corresponde a los meses con los niveles de ventas más bajos.

De acuerdo con el mapa de calor de las correlaciones, las variables que presentan una fuerte correlación son Venta Toneladas Métricas, Ingreso Total, Margen Total y Número de Pedidos, las cuales fueron seleccionadas para el análisis multivariado y de clusterización. Ver Figura 3.

2.1. Análisis Factorial

El análisis factorial es una técnica estadística multivariada utilizada para identificar relaciones subyacentes en un conjunto de variables observadas. Su principal objetivo es descubrir los factores latentes que explican las correlaciones entre dichas variables, conservando la mayor cantidad de información posible (Mainardi & Bidoia, 2022).

Antes de proceder con la aplicación de esta técnica, es fundamental evaluar la idoneidad del conjunto de datos. Para ello se emplea la medida de Kaiser-Meyer-Olkin (KMO) en la cual los valores cercanos a 1 indican que los datos son muy adecuados para el análisis factorial. Un valor de KMO superior a 0.6 se considera aceptable.

De acuerdo con los resultados obtenidos, el índice de El KMO fue igual a 0.697 y la Prueba de esfericidad de Barret menor a 0.001 lo cual confirma que los datos estudiados sí pueden ser analizados con la prueba de factorial. Ver Tabla 2

Tabla 2. Trimestre 2 - Análisis Factorial - KMO

Ensayo	Valor	Limites	Resultado	Conclusión
KMO	**0,697**	>0,6	Verdadero	Se puede analizar con la técnica de Factorial
Prueba de esfericidad	< 0,001	< 0,05	Verdadero	

La tabla de comunalidades permite conocer el grado que aporta cada una de las variables a la dimensión y determinar si existe una variable sospechosa a ser eliminada. En este caso ninguna variable es sospechosa de ser eliminada. Ver Tabla 3.

Tabla 3. Trimestre 2 - Análisis factorial - Comunalidad

Variable	Extracción	Límites	Resultado	Conclusión
Ventas de Toneladas Métricas	0,967	>0,5	Verdadero	No sospechosa de ser eliminada
Ingreso total	0,972	>0,5	Verdadero	No sospechosa de ser eliminada
Margen total	0,958	>0,5	Verdadero	No sospechosa de ser eliminada
Número de pedidos	0,965	>0,5	Verdadero	No sospechosa de ser eliminada

La tabla de varianza total explicada muestra cuanto explica el modelo acorde a la cantidad de componentes, el límite inferior es 60%. De acuerdo con la información estudiada, el modelo determinó que las variables explican 1 componente y el mismo explica un 96% de la variabilidad de los datos, ver Tabla 4.

Tabla 4. Trimestre 2 - Análisis factorial - Varianza total explicada

	Autovalores iniciales			Sumas de cargas al cuadrado de la extracción		
Componente	**Total**	**% de varianza**	**% acumulado**	**Total**	**% de varianza**	**% acumulado**
1	3,862	96,542	96,542	3,862	96,542	96,542
2	,136	3,406	99,947			
3	,005	,043	99,990			
4	,000	,010	1000			

Método de extracción: análisis de componentes principales.

La matriz de componentes permite corroborar que tanto aporta las variables a la dimensión. De acuerdo con la variable analizada, solo tenemos un componente explicado por las 4 cuatro variables, donde todas las variables presentan cargas superiores a 0.90 lo que indica una correlación muy alta con el Componente 1. Ver Tabla 5

Tabla 5. Trimestre 2 - Análisis factorial - Matriz de componentes

	Componente
Variables	**1**
Venta_MT	,984
Ingreso_Total	,986
Margen_Total	,979
Num_Pedidos	,982

Método de extracción: análisis de componentes principales.

a. 1 componentes extraídos.

La matriz de análisis de fiabilidad permite verificar la consistencia interna entre las variables. Cuando existe un valor inferior a 0,35 y tiene sospecha de ser eliminada acorde a la Tabla 3, la variable ya no se considera para el estudio. Dado que todas las variables tienen un valor de correlación mayor a 0.35 ninguna fue descartada. Ver Tabla 6.

Tabla 6. Trimestre 2 - Análisis de fiabilidad - Matriz de correlación corregida.

	Correlación total de elementos corregida	**Limites**	**Resultado**	**Conclusión**
Venta_MT	,938	>0,35	Verdadero	No se elimina
Ingreso_Total	,998	>0,35	Verdadero	No se elimina
Margen_Total	,998	>0,35	Verdadero	No se elimina
Num_Pedidos	,935	>0,35	Verdadero	No se elimina

2.2. Análisis de Componentes Principales

"El análisis de componentes principales (PCA) es una técnica multivariante que analiza datos descritos por variables cuantitativas intercorrelacionadas, con el objetivo de extraer información significativa y representarla como nuevas variables ortogonales llamadas componentes principales." (Mishra, 2017)

El análisis de componentes principales es una técnica multivariada cuyo objetivo es reducir la dimensionalidad de un conjunto de datos. A diferencia del Análisis Factorial, esta se centra en transformar las variables originales en un nuevo conjunto de variables no correlacionadas llamadas componentes principales, que expliquen la mayor parte de la variabilidad total.

Para este análisis de componentes principales se utilizaron las cuatro variables que no fueron descartadas en el Análisis Factorial. Se definió un modelo con dos componentes principales con el fin de capturar la mayor variabilidad de la información de los productos.

Los resultados muestran que los dos componentes explican en conjunto el 99,47% de la varianza total de los datos, lo que demuestra la eficacia del modelo para condensar la información presente en las variables originales.

Tabla 7. Trimestre 2 - Análisis de componentes - Varianza total explicada

	Autovalores iniciales			Sumas de cargas al cuadrado de la extracción			Sumas de cargas al cuadrado de la rotación		
Componente	**Total**	**% de varianza**	**% acumulado**	**Total**	**% de varianza**	**% acumulado**	**Total**	**% de varianza**	**% acumulado**
1	3,862	96,542	96,542	3,868	96,542	96,542	2,004	50,107	50,107
2	,130	3,406	99,947	,136	3,406	99,947	1,994	49,840	99,947
3	,002	,043	99,990						
4	,000	,010	100,000						

Método de extracción: análisis de componentes principales.

La matriz de componentes rotados, mediante rotación Varimax, permite una mejor interpretación de los resultados. El Componente 1 está conformado por las variables por Ingreso Total y Margen Total con cargas de 0.819 y 0.835. Mientras que, el Componente 2 está conformado por las variables Venta MT y Número de Pedidos que presentan cargas de 0.824 y 0.828 respectivamente. Ver Tabla 8.

Tabla 8. Trimestre 2 - Análisis de componentes - Matriz de componentes rotados.

	Componente	
Variables	**1**	**2**
Venta_MT	,824	,567
Ingreso_Total	,581	,813
Margen_Total	,550	,835
Num_Pedidos	,828	,560

Método de extracción: análisis de componentes principales.

Método de rotación: Varimax con normalización Kaiser.

a. La rotación ha convergido en 3 iteraciones.

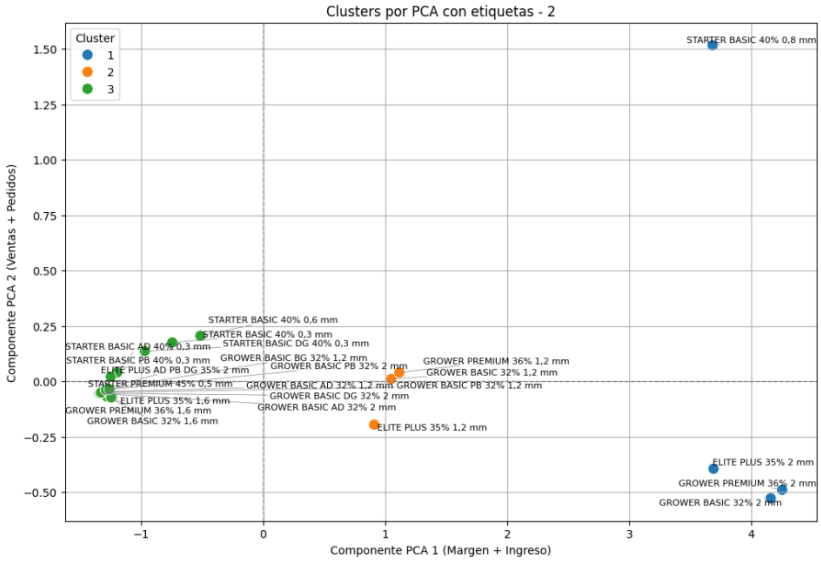
Finalmente, al representar gráficamente los dos componentes en un diagrama de dispersión, ubicando el Componente 1 en el eje de las X y el Componente 2 en el eje Y, se logró identificar 3 zonas en la distribución de los productos de acuerdo con sus características comerciales. En el cuadrante correspondiente a menor ingreso, margen, venta y cantidad de pedidos se identificaron 14 productos. En el cuadrante de rotación media, demanda y margen medios se identificaron 3 productos. Y en el cuadrante de alta rotación, demanda y margen se identificaron 4 productos para el Segundo Trimestre. Ver Figura 1.

2.3. Técnicas de Agrupamiento

Las técnicas de agrupamiento son métodos de aprendizaje automático no supervisado que permiten agrupar los puntos de datos en función de sus atributos similares con el objetivo de revelar conexiones ocultas sin etiquetas predefinidas. Existen diferentes técnicas de agrupamientos como clustering difuso, k-medias, jerárquica, de desplazamiento de la media, de maximización de expectativas y DBSCAN (Sonawane & Pati, 2020).

Para el presente análisis se utilizó la técnica K-means. “K-means es un algoritmo de agrupamiento que caracteriza puntos de datos en K grupos distintos según sus distancias desde los centroides.” (Chao-quin, 2014). Esta técnica se utilizó con el fin de identificar clústeres para los productos de la base de datos. Para ello, se utilizaron dos herramientas de análisis: Python y SPSS, lo que permitió validar la consistencia de los resultados obtenidos. En la Figura 10 y Figura A 3. ubicada en los Anexos, se visualiza la similitud en las agrupaciones.

Figura 10. K-means desarrollado en Python para productos del trimestre 2



La segmentación de los productos mediante la técnica de K-means permitió identificar tres grupos diferenciados según sus características comerciales como nivel de rotación, demanda o número de pedidos, margen e ingresos.

El Clúster 1 está conformando por aquellos productos con mayor desempeño caracterizados por altos niveles de rotación, demanda y margen. En promedio, registran ventas trimestrales de aproximadamente 95 mil toneladas métricas, con una desviación estándar de 18 mil toneladas métricas. La rotación promedio es de 53 pedidos por trimestre, lo que genera un margen total de $28 millones de dólares y un ingreso total por $120 millones de dólares por trimestre.

El Clúster 2 por su parte, está compuesto por productos con un desempeño moderado caracterizados por tener niveles intermedios de rotación, demanda y margen. Las ventas trimestrales en promedio alcanzan las 39 mil toneladas métricas con una desviación estándar de 5 mil toneladas métricas. En promedio, se generan 22 pedidos por trimestre, generando un margen total de $11 millones de dólares e ingresos por $46 millones de dólares por trimestre.

Finalmente, el Clúster 3 está conformado por los productos de bajo desempeño, definidos por tener bajos niveles de rotación, demanda y margen. Su volumen de ventas trimestrales en promedio es de 4 mil toneladas métricas con una desviación estándar de 3 mil toneladas métricas. La rotación es muy baja, con apenas de 2 pedidos por trimestre en promedio, lo que genera 1 millón de dólares de margen y 7 millones de dólares de ingresos. No obstante, la alta variabilidad observada mediante su desviación estándar en variables como el margen, ingreso y número de pedidos, sugiere una inestabilidad comercial significativa que podría periodos con resultados negativos o sin ventas.

Tabla 9. Detalle de medias de los clústeres: Ventas, Número de pedidos, margen e Ingreso.

Clúster Identificados	Ventas MT	Número Pedidos	Margen Total	Ingreso Total
Alta rotación, Alta demanda y Alto margen	95K MT ± 18K	53 ± 9	$ 28M USD ± 6M	120M USD ± 21M
Media rotación, media demanda, medio margen	39K MT ± 5K	22 ± 2	$ 11M USD ± 2M	46M USD ± 6M
Baja rotación, baja demanda y bajo margen	4K MT ± 3K	2 ± 2	$ 1M USD ± 2M	7M USD ± 8M

Con base en los resultados obtenidos, en la Tabla 10 se detallan los planes de acción y recomendaciones según el tipo de clúster.

Tabla 10. Trimestre 2 - Resumen de cantidad de productos por Clúster

Clúster Identificados	Ward Method	Productos	Plan de acción
Alta rotación, alto demanda y alto margen	1	4	Mantener. Fortalecer la producción y distribución
Media rotación, media demanda, medio margen	3	3	Impulsar mediante mejoras en formulación o posicionamiento
Baja rotación, baja demanda y bajo margen	2	16	Estrategia comercial para asignar demanda o cambiar a SKUS de media o alta rotación

2.4. Análisis predictivo - KNN

Para complementar el análisis, se aplicaron técnicas de clasificación predictiva. El algoritmo KNN fue la metodología seleccionada para realizar un análisis predictivo para clasificar futuras observaciones. En esta metodología "las predicciones se realizan identificando los K vecinos más cercanos de todo el espacio de muestras de entrenamiento basándose en un punto de datos de prueba" (Zhang & Li, 2020)

Se evaluaron dos métodos: KNN y árboles de decisión. Finalmente se decidió seleccionar como método de clasificación predictivo al método KNN por las siguientes razones:

- Repetitividad de precisión en Softwares: Python, RStudio y SPSS.
- Se identifica que el K seleccionado de 3 faculta una precisión optima sin sacrificar rendimiento. Optimizando el análisis y el procesamiento de la información.
- Los modelos de árbol de decisión determinaban las decisiones en función 1 o dos variables sin tomar en cuenta todo el espectro.

Tabla 11. Resumen de precisiones de métodos de clasificación predictivos evaluados

Trimestre	Método	Precisión	Calificación
Trimestres 1	KNN	1	Muy bueno
	Árbol de Decisión	0.857	Bueno
Trimestre 2	KNN	1	Muy bueno
	Árbol de Decisión	1	Muy bueno
Trimestre 3	KNN	1	Muy bueno
	Árbol de Decisión	1	Muy bueno
Trimestre 4	KNN	0.857	Bueno
	Árbol de Decisión	1	Muy bueno

D. Resultados

En el segundo trimestre de cada año se han identificado tres clústeres de SKUs. La información obtenida brinda la oportunidad al equipo de planificación de desarrollar una estrategia de producción para los productos del clúster de mayor y mediana venta rotación, margen e ingreso. Mientras que aquellos SKUS con baja rotación, venta, margen e ingreso son comunicados al área de comercial para que se defina una estrategia más alineada hacia las producciones de los otros dos clústeres.

Tabla 12. Trimestre 2 - Productos en Clúster baja rotación, demanda y margen

Clúster	Producto
Baja rotación, Baja demanda y Bajo margen	ELITE PLUS 35% 1,6 mm
	ELITE PLUS AD PB DG 35% 2 mm
	GROWER BASIC 32% 1,6 mm
	GROWER BASIC AD 32% 1,2 mm
	GROWER BASIC AD 32% 2 mm
	GROWER BASIC BG 32% 1,2 mm
	GROWER BASIC DG 32% 2 mm
	GROWER BASIC PB 32% 1,2 mm
	GROWER BASIC PB 32% 2 mm
	GROWER PREMIUM 36% 1,6 mm
	STARTER BASIC 40% 0,3 mm
	STARTER BASIC 40% 0,6 mm
	STARTER BASIC AD 40% 0,3 mm
	STARTER BASIC DG 40% 0,3 mm
	STARTER BASIC PB 40% 0,3 mm
	STARTER PREMIUM 45% 0,5 mm

Tabla 13. Trimestre 2 - Productos en Clúster media rotación, demanda y margen

Clúster	Producto
Media rotación, Media demanda y Medio margen	ELITE PLUS 35% 1,2 mm
	GROWER BASIC 32% 1,2 mm
	GROWER PREMIUM 36% 1,2 mm

Tabla 14. Trimestre 2 - Productos en Clúster alta rotación, demanda y margen

Clúster	Producto
Alta rotación, Alta demanda y Alto margen	ELITE PLUS 35% 2 mm
	GROWER BASIC 32% 2 mm
	GROWER PREMIUM 36% 2 mm
	STARTER BASIC 40% 0,8 mm

Los hallazgos obtenidos proporcionan una visión estratégica del portafolio al diferenciar y agrupar los productos con base en su desempeño comercial. Esto permite, por un lado, planificar de forma más eficiente la producción y distribución de los productos. Mientras que, por otro lado, permite diseñar y establecer estrategias comerciales orientadas al fortalecimiento de los productos con alto rendimiento, así como replantear la formulación o posicionamiento de aquellos productos con bajo desempeño.

Para enriquecer los resultados obtenidos y fortalecer la toma de decisiones estratégicas, se recomienda complementar este estudio con un análisis de carácter estacional que permita monitorear la evolución de los clústeres a lo largo del tiempo. Este enfoque permitiría identificar si hay productos que migran entre grupos de desempeño en función del trimestre, además de factores estacionales que puedan afectar su comportamiento comercial.

E. Referencias

Chao-quin, H. (2014). A K-means algorithm. *Journal of Changchun University of Technology*. Obtenido de http://en.cnki.com.cn/Article_en/CJFDTOTAL-JLGX201402006.htm

Mainardi, P., & Bidoia, E. (2022). Fundamental concepts and recent applications of factorial statistical designs. *Brazilian Journal of Biometrics, 40*(1). doi:https://doi.org/10.28951/bjb.v40i1.552

Mishra, S. P. (2017). *Multivariate Statistical Data Analysis- Principal Component Analysis (PCA).* International Journal of Livestock Research, 7(5), 60–78. Obtenido de https://www.bibliomed.org/?mno=261590

Sonawane, R. C., & Pati, H. D. (2020). Clustering Techniques and Research Challenages in Machine Learning. *2020 Fourth International Conference on Computing Methodologies and Communication (ICCMC)* (págs. 290-293). Erode: IEEE. doi:10.1109/ICCMC48092.2020.ICCMC-00054

Zhang, S., & Li, J. (2020). KNN Classification with One-step Computation. *arXiv: Learning*. Obtenido de https://arxiv.org/pdf/2012.06047

ANEXOS A

Figura A 1. Ventas de toneladas acorde a la ubicación geográfica

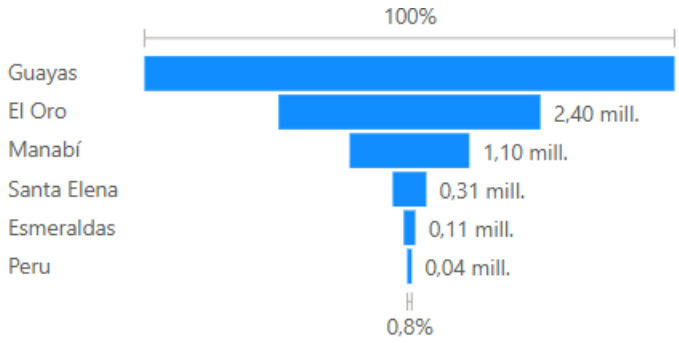


Figura A 2. Ventas toneladas por segmento de cliente de los últimos 5 años

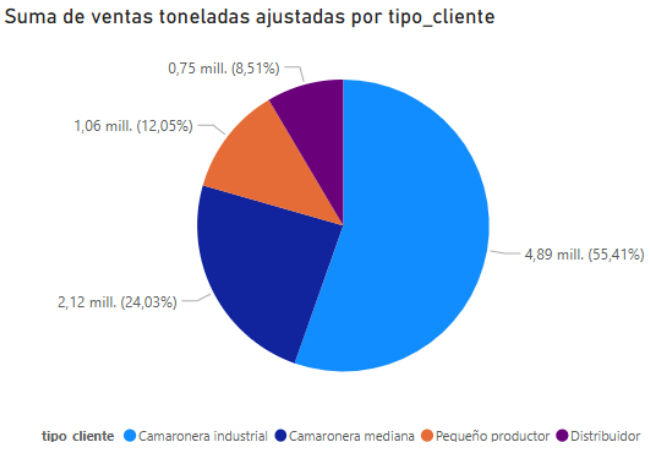


Figura A 3. Trimestre 2 – Clúster de Productos

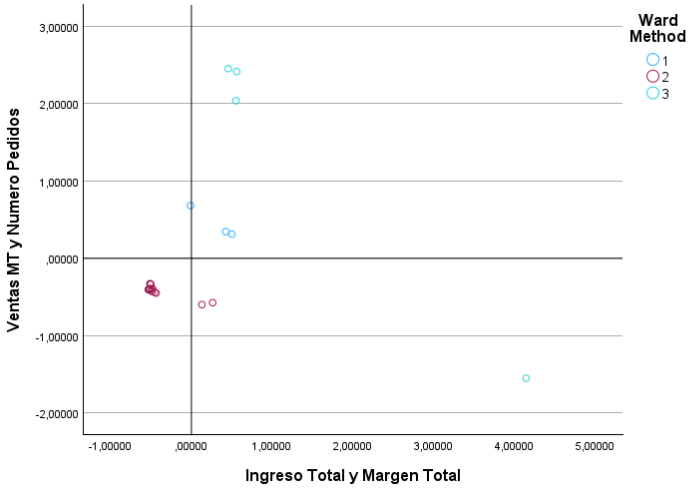


Figura A 4. Trimestre 2 – Clúster para nuevos Productos en el Trimestre 2.

