

MIAD



Maestría
en Inteligencia
Analítica de Datos



Proyecto del curso Aprendizaje no Supervisado

Segmentación del mercado colombiano de cerveza.

A partir de sus propiedades fisicoquímicas y sensoriales.

Realizado por: Santiago Pulido, Liliana Catalina Briceño, Diego Peñaloza, Cristhian Barbosa

Septiembre de 2024



Problema: Mercado monopolizado y uniforme de cerveza en Colombia que representa ingresos anuales de 22 billones.

Objetivo: Encontrar oportunidades de negocio a partir de posibles segmentos de crecimiento para los diferentes tipos de cerveza.



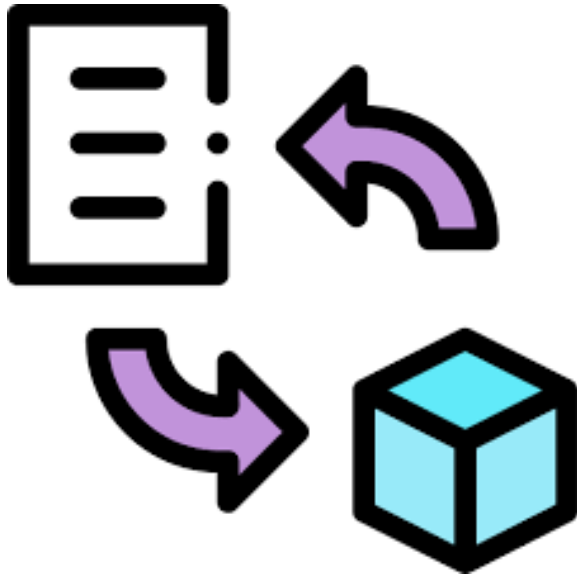
Datos

Se cuenta con las características físico químicas, sensoriales y de producción de 52 marcas de cervezas masivas y artesanales del mercado nacional para los años entre 2020 y 2023.



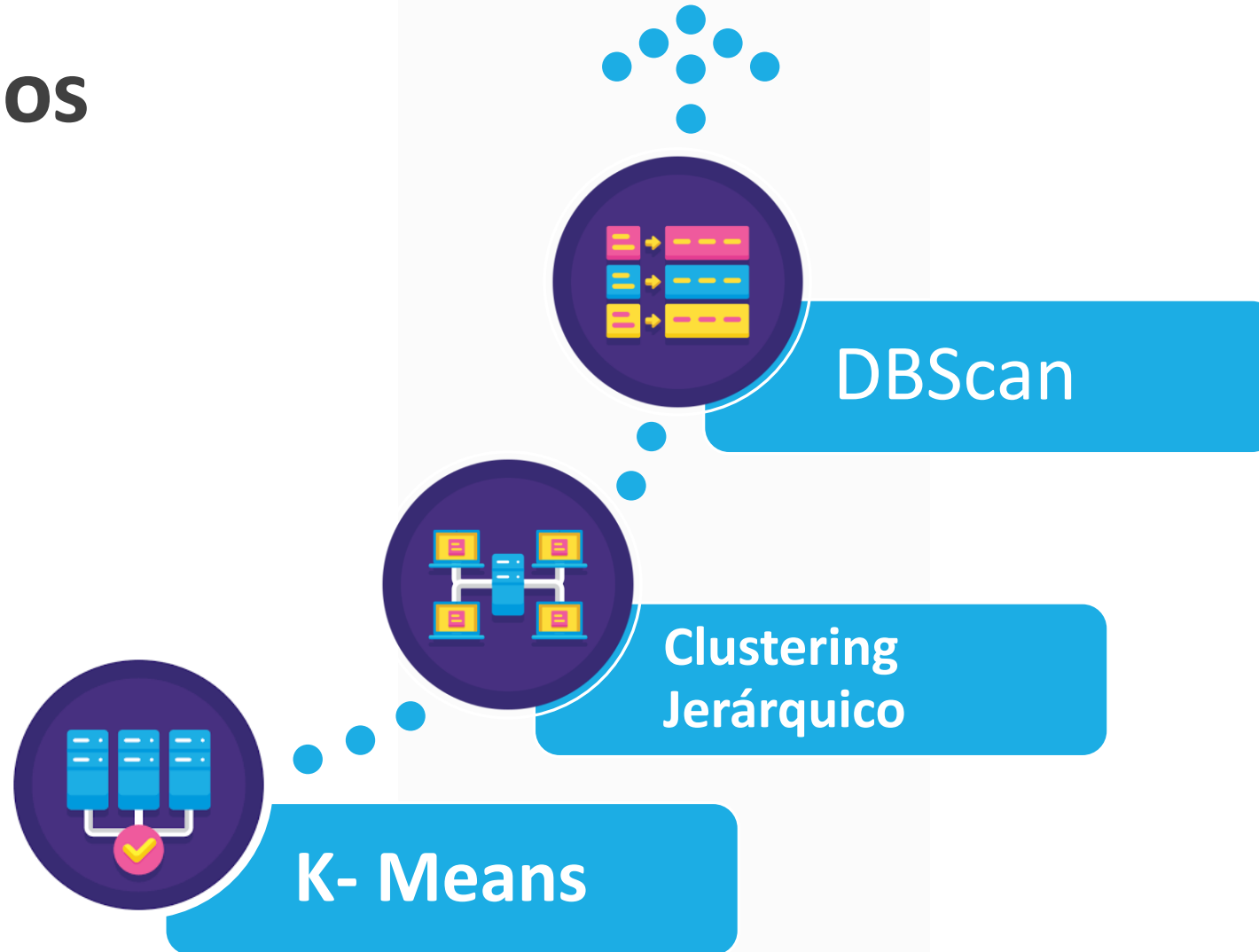
| Variable | Unidad | Definición |
|----------------------------------|--------------------------|--|
| Parámetro Físico químicos | | |
| Extracto Original | °P | Concentración de azúcares en el mosto antes de la fermentación. Determina el potencial de alcohol que la cerveza puede alcanzar después de la fermentación |
| Extracto Aparente | °P | Concentración de azúcares que quedan después de la fermentación. |
| Alcohol | % v/v | Cantidad de alcohol en la cerveza |
| pH | | Medida de acidez o alcalinidad de la cerveza. Asegura la estabilidad del sabor, claridad y conservación de la cerveza. |
| Color | EBC | Se refiere a la intensidad de color de la cerveza. Esto varía dependiendo del tipo de cerveza. |
| Amargo | IBU | Cantidad de compuestos amargos, sobre todo ácidos alfa de lúpulo en la cerveza. El valor de IBU varía dependiendo del tipo de cerveza. |
| Medición Sensorial | | |
| Turbiedad | 1. Claro / cristalino. | Se refiere a la clara u opacidad de la cerveza. Influye en la percepción visual de la bebida |
| | 2. Ligeramente turbio. | |
| | 3. Moderadamente turbio. | |
| | 4. Turbio. | |
| Cuerpo | 1. Cuerpo bajo. | Sensación de textura en boca de la cerveza. Se pueden llegar a percibir cervezas más "pesadas" o más "ligeras". |
| | 2. Cuerpo ligero medio. | |
| | 3. Cuerpo moderado. | |
| | 4. Cuerpo alto. | |
| Dulzor | 1. Nada dulce. | Percepción de los azucars residuales en la cerveza. |
| | 2. Ligeramente dulce. | |
| | 3. Moderadamente dulce. | |
| | 4. Dulce alto. | |
| Astringencia | 1. Ninguna. | Sensación de sequedad o aspereza en la boca. |
| | 2. Leve. | |
| | 3. Moderado. | |
| | 4. Alto. | |
| Ventas 2021 | HI (hectolitros) | Cantidad de hectolitro vendidos por año. |
| Ventas 2022 | | |
| Ventas 2023 | | |

Transformación de datos



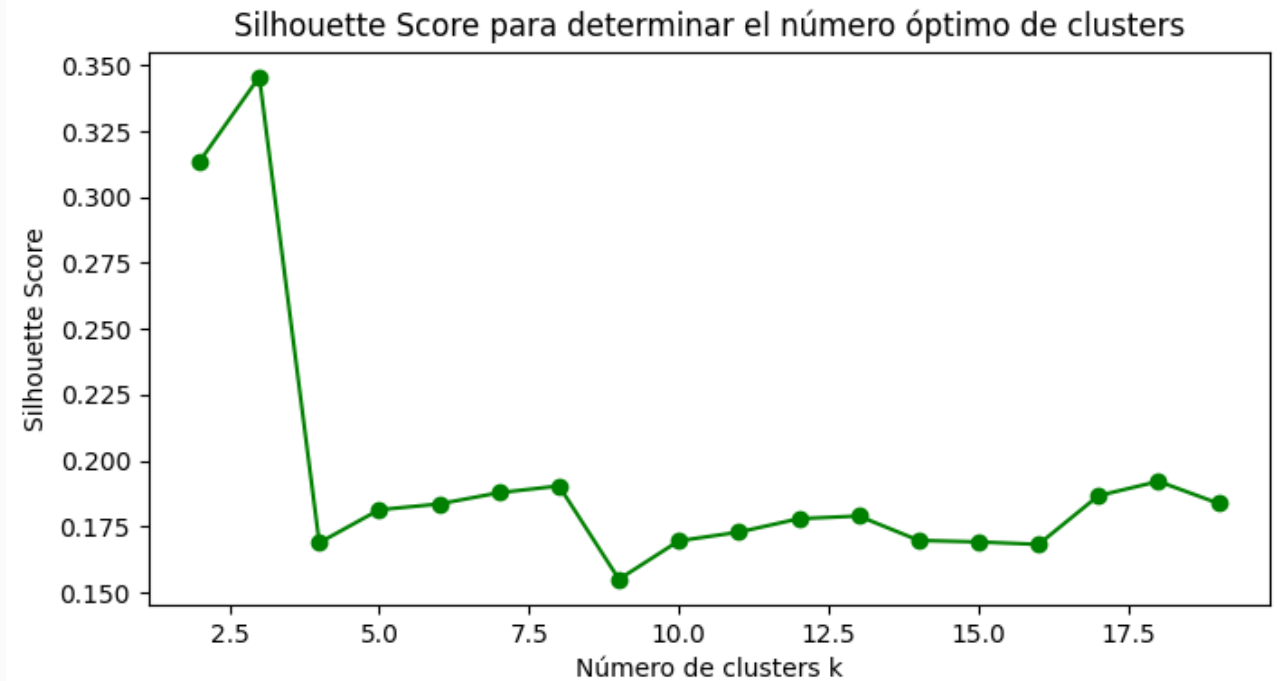
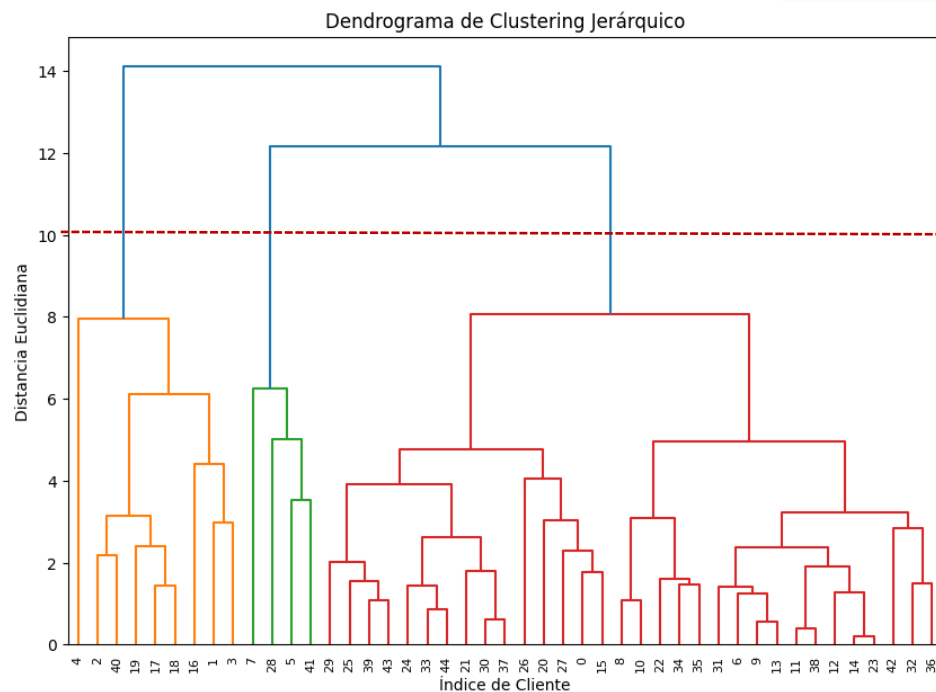
- Se **escalan** las **variables numéricas** (parámetros físico-químicos)
- Se **estandarizan** las **variables ordinales** (variables de medición sensorial), de manera que se re-escalan los valores de cada variable para que estén en un rango proporcional a su valor máximo, centrados alrededor de 0.5.
- Se estima el **incremento anual de ventas**. Sin embargo, **las ventas y el incremento no son incluidas en la estimación de clústeres**, sino en la evaluación de nuevas estrategias de mercado.

Métodos





Resultados Clustering Jerárquico



Se seleccionan 3 clústeres según lo definido en el Silhouette Score.

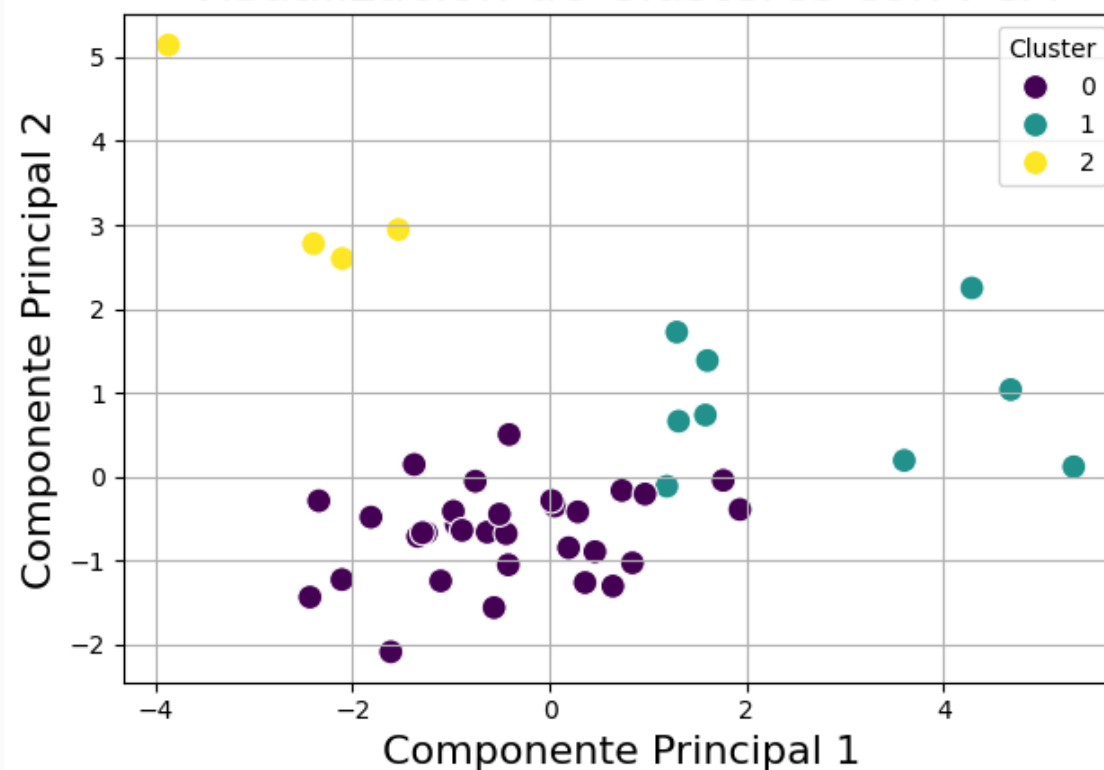


Resultados Clustering Jerárquico

Estadísticas generales clústeres - C. Jerárquico

| Cluster | | 0 | 1 | 2 |
|-------------------|------|-------|-------|------|
| n | | 4 | 17 | 8 |
| Extracto aparente | Mean | 1,86 | 2,92 | 4,55 |
| | Std | 0,87 | 0,8 | 1,59 |
| Alcohol v/v | Mean | 4,28 | 4,85 | 2,38 |
| | Std | 0,54 | 0,65 | 2 |
| pH | Mean | 4,38 | 4,45 | 3,77 |
| | Std | 0,14 | 0,14 | 0,24 |
| Color | Mean | 10,21 | 45,42 | 6,64 |
| | Std | 13,07 | 47,58 | 0,89 |
| Amargo | Mean | 13,9 | 28,18 | 10,8 |
| | Std | 3,8 | 8,22 | 8,35 |
| | Min | 6,11 | 19,6 | 1,5 |
| | Max | 20 | 44,4 | 20,3 |
| Turbiedad | Mode | 1 | 3 | 1 |
| Cuerpo | Mode | 2 | 3 | 1 |
| Dulzor | Mode | 2 | 3 | 4 |
| Astringencia | Mode | 2 | 3 | 3 |

Visualización de Clústeres con PCA





Resultados DB-Scan

Estadísticas generales clústeres – DB Scan

| Clúster | | -1 | 0 | 1 |
|-------------------|------|-------|-------|------|
| Extracto aparente | Mean | 3,47 | 2,07 | 0,35 |
| | Std | 1,27 | 0,47 | 0,13 |
| Alcohol v/v | Mean | 4,13 | 4,3 | 3,88 |
| | Std | 1,64 | 0,55 | 0,37 |
| pH | Mean | 4,2 | 4041 | 4,26 |
| | Std | 0,37 | 0,09 | 0,14 |
| Color | Mean | 37,33 | 8,93 | 5,92 |
| | Std | 44,75 | 5,75 | 0,7 |
| Amargo | Mean | 21,92 | 14,99 | 7,3 |
| | Std | 11,97 | 3,52 | 1,37 |
| | Min | 1,5 | 7,6 | 6,11 |
| | Max | 44,4 | 20,5 | 8,8 |
| Turbiedad | Mode | 1 | 1 | 1 |
| Cuerpo | Mode | 2 | 2 | 2 |
| Dulzor | Mode | 2 | 2 | 2 |
| Astringencia | Mode | 3 | 2 | 2 |

Resultados por número de vecinos

| Neighbors | Epsilon | Silhouette | D-B Score |
|-----------|---------|------------|-----------|
| 3 | 1.13 | 0.29 | 1.64 |
| 2 | 1.03 | 0.25 | 1.65 |
| 8 | 1.77 | 0.47 | 1.90 |
| 9 | 1.81 | 0.47 | 1.90 |
| 7 | 1.70 | 0.39 | 2.54 |
| 6 | 1.47 | 0.37 | 2.93 |
| 4 | 1.24 | 0.36 | 2.97 |
| 5 | 1.35 | 0.36 | 2.97 |

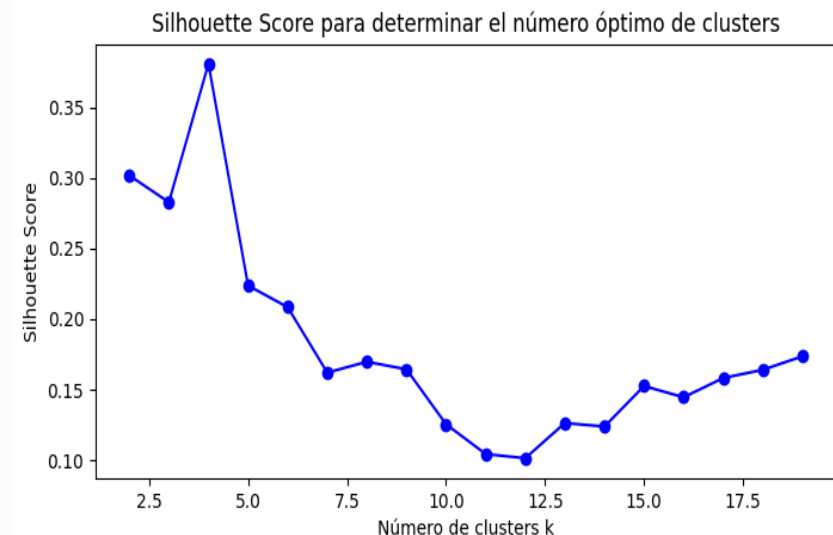


Método seleccionado: K-means

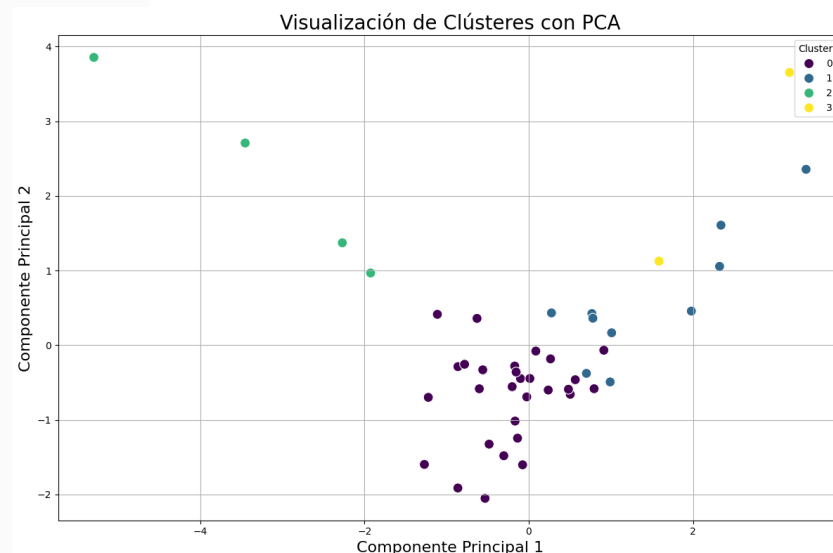
Descripción de los clústeres

Estadísticas generales clústeres - k means

| Clúster | | 0 | 1 | 2 | 3 |
|-------------------|------|-------|-------|------|--------|
| n | | 29 | 10 | 4 | 2 |
| Extracto aparente | Mean | 1,77 | 2,66 | 4,55 | 3,92 |
| | Std | 0,84 | 0,59 | 1,59 | 0,69 |
| Alcohol v/v | Mean | 4,23 | 4,7 | 2,38 | 5,47 |
| | Std | 0,55 | 0,39 | 2 | 1,05 |
| pH | Mean | 4,37 | 4,45 | 3,77 | 4,47 |
| | Std | 0,13 | 0,12 | 0,24 | 0,2 |
| Color | Mean | 7,28 | 30,3 | 6,64 | 110,65 |
| | Std | 2,37 | 29,29 | 0,89 | 47,16 |
| Amargo | Mean | 13,61 | 26,86 | 10,8 | 17,5 |
| | Std | 3,72 | 8,6 | 8,35 | 7,78 |
| | Min | 6,11 | 18 | 1,5 | 12 |
| | Max | 18,6 | 44,4 | 20,3 | 23 |
| Turbiedad | Mode | 1 | 3 | 1 | 1 |
| Cuerpo | Mode | 2 | 3 | 1 | 3 |
| Dulzor | Mode | 2 | 3 | 4 | 1 |
| Astringencia | Mode | 2 | 3 | 3 | 4 |



Se seleccionan 4 clústeres según lo definido en el Silhouette Score.



Método seleccionado: K-means

Análisis de resultados



| Modelo | Clúster | Descripción |
|-------------------|-----------|--|
| K-Means | 0 | Cervezas tipo lager que presentan cuerpo, dulzor y astringencia relativamente bajos, con un color dorado y un nivel de amargor moderadamente bajo. |
| | 1 | Cervezas artesanales o industriales con cuerpo, dulzor y astringencia moderadamente altos. |
| | 2 | Cervezas predominantemente dulces con un contenido alcohólico bajo. |
| | 3 | Cervezas oscuras, con cuerpo y astringencia moderadamente altos. |
| Jerárquico | 0 | Cervezas tipo lager y ale, con un rango de color bastante amplio desde cervezas doradas hasta rojizas. Cuentan con turbiedad clara, cuerpo, dulzor y astringencia moderada. Alcohol moderado con un alto rango de amargor. |
| | 1 | Cervezas artesanales o industriales con un cuerpo, astringencia y dulzor alto. Color moderadamente oscuro y rango de amargo relativamente elevado. |
| | 2 | Cervezas color dorado, con un bajo nivel de alcohol, dulzor elevado y astringencia moderadamente alta. Segmento enfocado en cervezas de alto consumo. |
| DBSCAN | -1 | Cervezas oscuras y amargas comparadas con las demás. Son cervezas que hacen parte de sector masivo industrial. |
| | 0 | Corresponde a cervezas con un grado de alcohol medio, relativamente claras y de un amargor medio bajo. |
| | 1 | Agrupar Cervezas de bajo nivel de alcohol, color más claro y un amargor bajo. |

Método seleccionado: K-means

Análisis de resultados



Gráfico de Incremento para Amargo

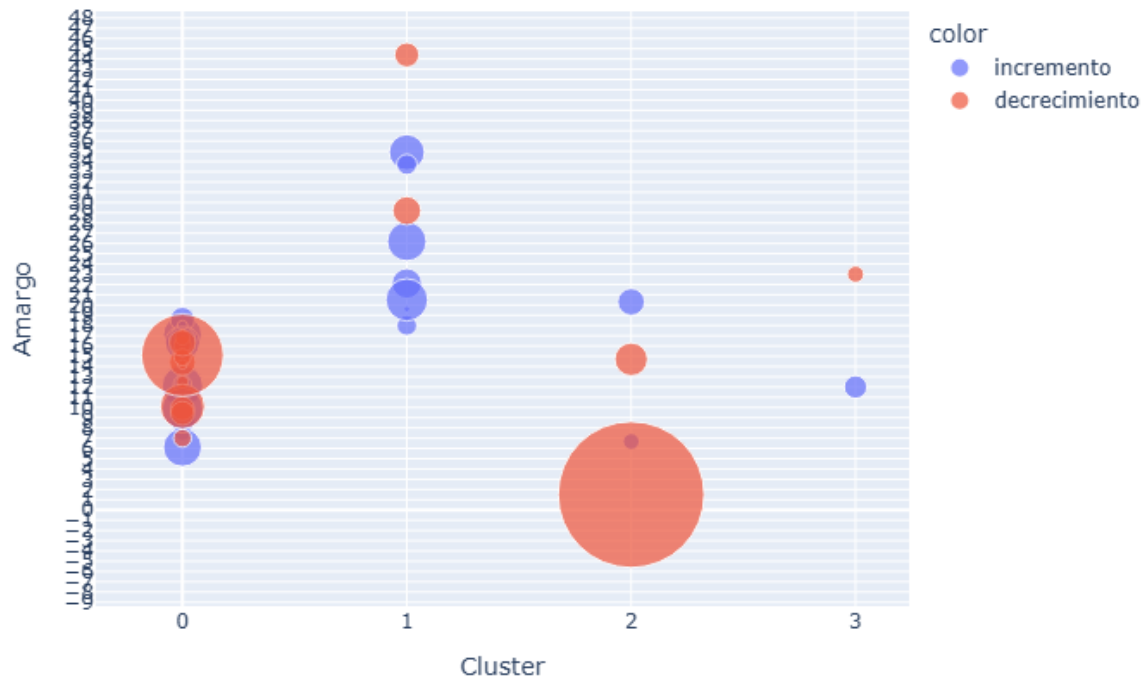
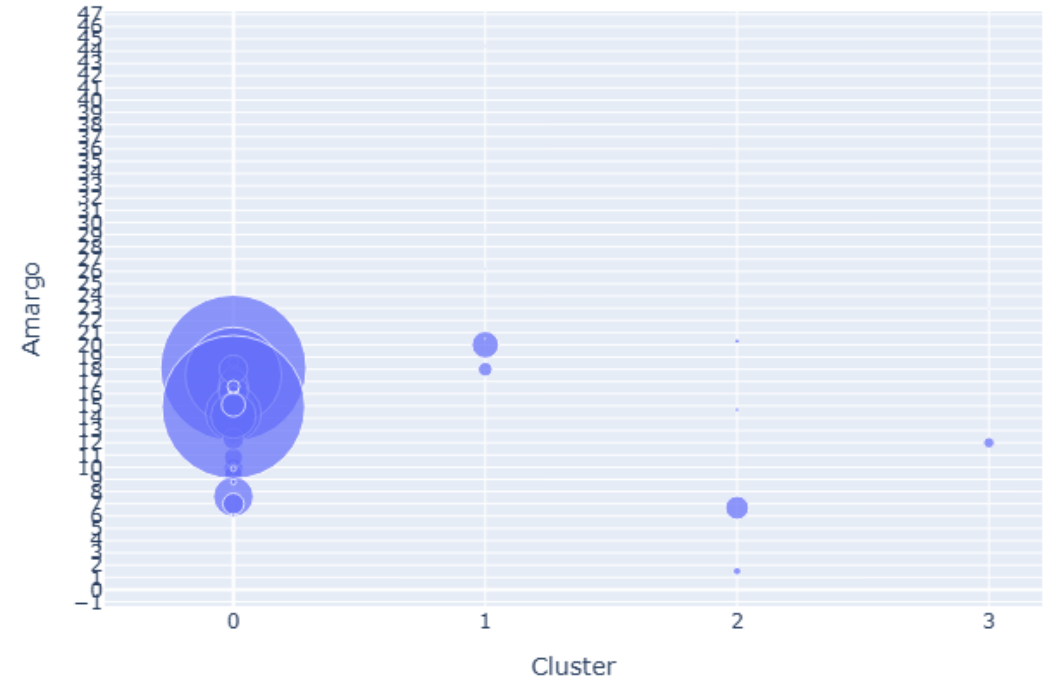


Gráfico de Ventas para Amargo





Conclusiones y recomendaciones

Oportunidad de negocio en el mercado de la cerveza artesanal o premium, con el desarrollo de una nueva marca.

Para invertir en cervezas de consumo masivo es necesario de un capital bastante alto dado que es un segmento muy competitivo y estancado.

Existe una oportunidad de negocio en las cervezas negras a largo plazo, dado su crecimiento lento y baja participación en el mercado.



Gracias