

ASIGNATURA: Inteligencia de Negocios

NOMBRE:

- Erick Casa
- Diego Cervantes
- Adrian Galeas
- Naobe Ovando
- Saud Quishpe

NIVEL: VII

ACTIVIDAD: Examen Final

FECHA: 06-12-2023**CALIFICACIÓN:**

TEMA: Examen Final

Datos de la empresa “GASOLINERA DEL PUERTO “

[illegible]

Dataset utilizado

	fecha	propietario_nombre	propietario_cedula	tripulante_nombre	tripulante_cedula	embarcacion_nombre	embarcacion_matricula	galones	motores	zarpe	factura	sector
0	3/1/2021	MARIN PARRALES JOHNNY ANIBAL	080206690-2	MARIN PARRALES JOHNNY ANIBAL	080206690-2	ELIZABETH I	B -02-05041	460	75HP / 40HP	1501969- 2021	6001000005107	PUERTO PESQUERO
1	3/1/2021	BENNETT TORRES WASHINGTON	080216148-9	BENNETT TORRES WASHINGTON	080216148-9	BIGLA X	B -02-08860	552	75HP / 75HP	1495877- 2020	6001000005108	PUERTO PESQUERO
2	3/1/2021	PARRALES ALAVA BALTAZAR FRANCISCO	080048207-7	PARRALES ALAVA BALTAZAR FRANCISCO	080048207-7	PAJARITO V	B -02-03699	368	75HP	1497116- 2020	6001000005109	PUERTO PESQUERO
3	3/1/2021	PARRALES QUIJIJE FRANCISCO FELIPE	080237238-3	PARRALES QUIJIJE FRANCISCO FELIPE	080237238-3	CHAVITA	B -02-08240	368	75HP	1497107- 2020	6001000005110	PUERTO PESQUERO
4	3/1/2021	CORTEZ BAGUI CARLOS AIRON	080140596-0	CORTEZ BAGUI CARLOS AIRON	080140596-0	MARIA YOLANDA	B -02-05506	460	75HP	1497149- 2020	6001000005111	PUERTO PESQUERO

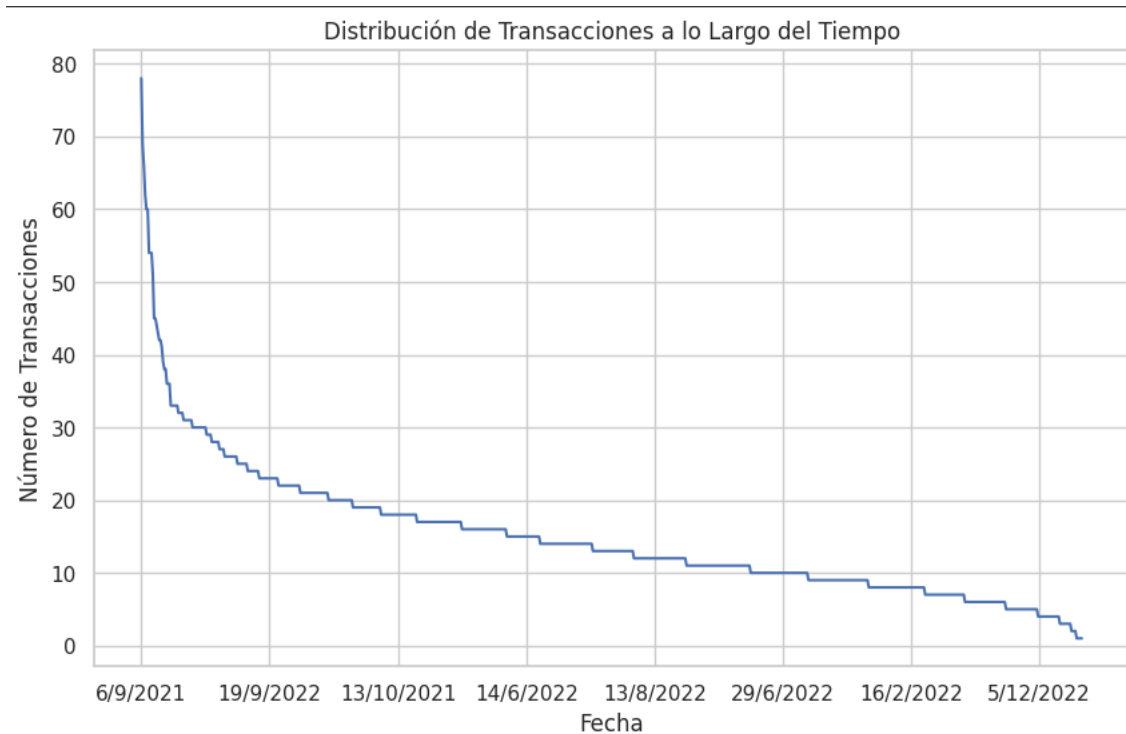
El dataset está relacionado con transacciones en un puerto pesquero, incluyendo detalles sobre las embarcaciones, sus propietarios y tripulantes, y aspectos logísticos como el combustible y las salidas de las embarcaciones. Aquí hay un detalle de los datos contenidos en el dataset:

1. **Fecha:** Representa la fecha en la que se registró la transacción o el evento. El formato parece ser día/mes/año. Este campo es importante para entender la cronología de las transacciones o actividades.
2. **Propietario Nombre:** El nombre completo del propietario de la embarcación. Este campo es clave para identificar a la persona responsable o dueña de la embarcación.

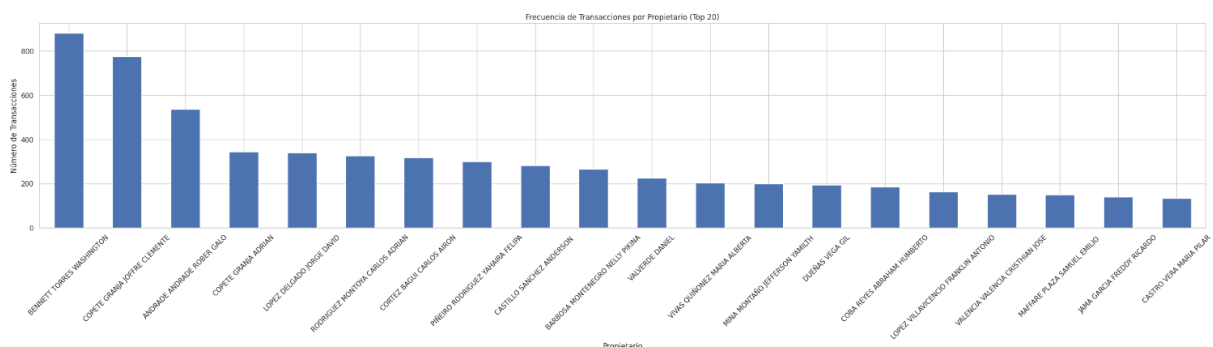
3. **Propietario Cédula:** El número de cédula de identidad del propietario. Este es un documento oficial en muchos países que sirve como identificación personal. Es útil para fines de registro y verificación legal.
4. **Tripulante Nombre:** Nombre del tripulante de la embarcación. Puede ser la misma persona que el propietario o alguien diferente. Este campo es relevante para identificar quién estaba a cargo o trabajando en la embarcación en el momento de la transacción.
5. **Tripulante Cédula:** Número de cédula de identidad del tripulante. Al igual que con el propietario, proporciona un medio de identificación oficial para la persona a bordo de la embarcación.
6. **Embarcación Nombre:** El nombre asignado a la embarcación. Este es un identificador único y personalizado para cada barco, útil para su identificación y registro.
7. **Embarcación Matrícula:** Número oficial de registro o matrícula de la embarcación. Este número es crítico para el registro legal de la embarcación y su identificación en actividades oficiales o legales.
8. **Galones:** La cantidad de galones de gasolina de pesca artesanal que fueron adquiridos. Este campo es esencial para entender el consumo o compra de combustible para las operaciones de la embarcación.
9. **Motores:** Detalles sobre los motores de la embarcación, incluyendo su potencia o modelo. Este campo es importante para comprender la capacidad y el tipo de embarcación, así como sus necesidades operativas.
10. **Zarpe:** Código del permiso proporcionado por la capitanía costera que respalda las operaciones de la lancha registrada.
11. **Factura:** El número de factura relacionado con la transacción. Este es un documento crucial para los propósitos contables y de seguimiento de las transacciones financieras.
12. **Sector:** El sector o área específica relacionada con la transacción, en este caso, parece ser un puerto pesquero. Este campo ayuda a identificar el lugar o contexto de la transacción o actividad.

Visualización de cada posible variable del dataset

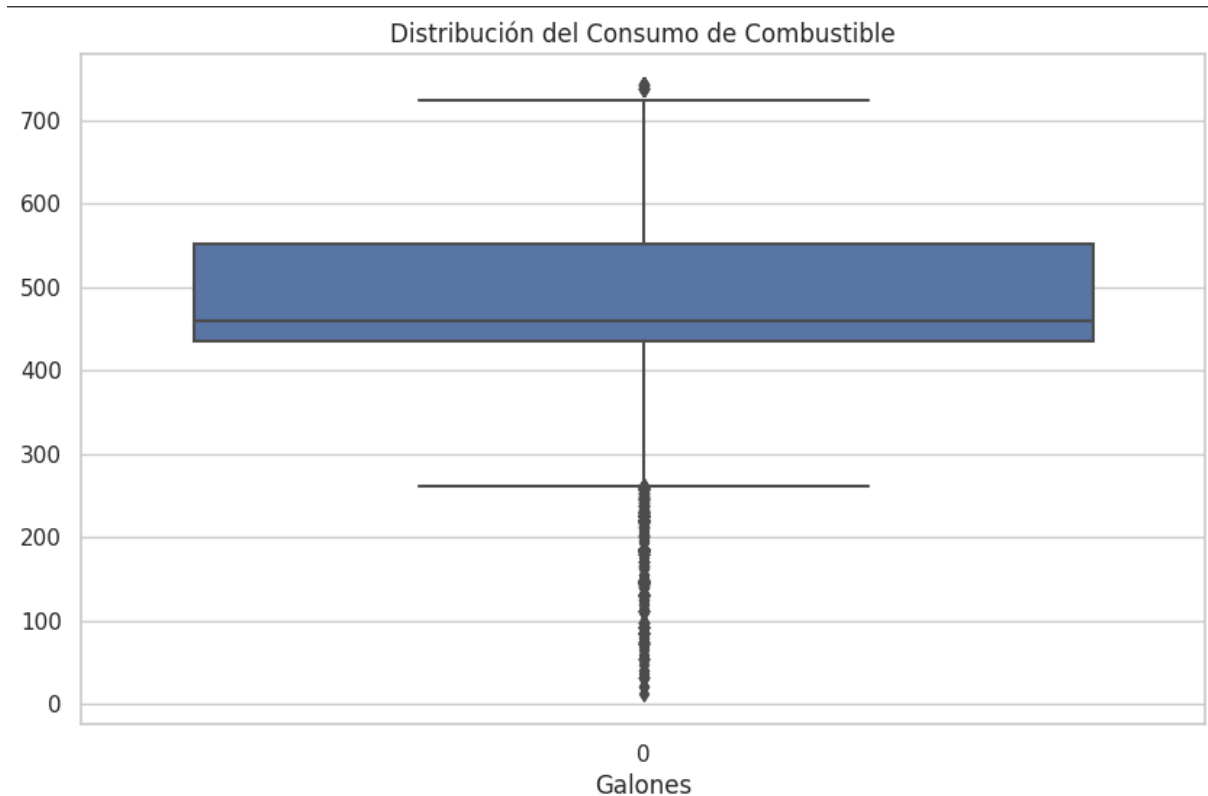
Fecha (Distribución de Transacciones a lo Largo del Tiempo): Un gráfico mostrando el número de transacciones a lo largo del tiempo para identificar tendencias temporales y patrones estacionales.



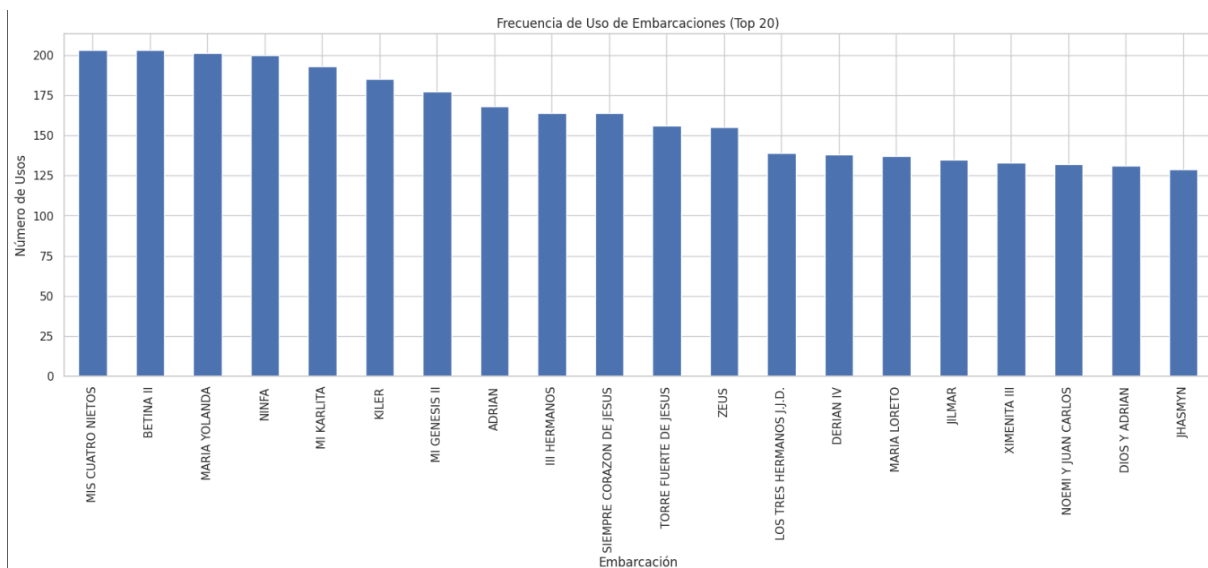
Propietario Nombre (Frecuencia de Transacciones por Propietario): Un histograma que muestra la frecuencia de transacciones por propietario para destacar los propietarios más activos y su distribución.



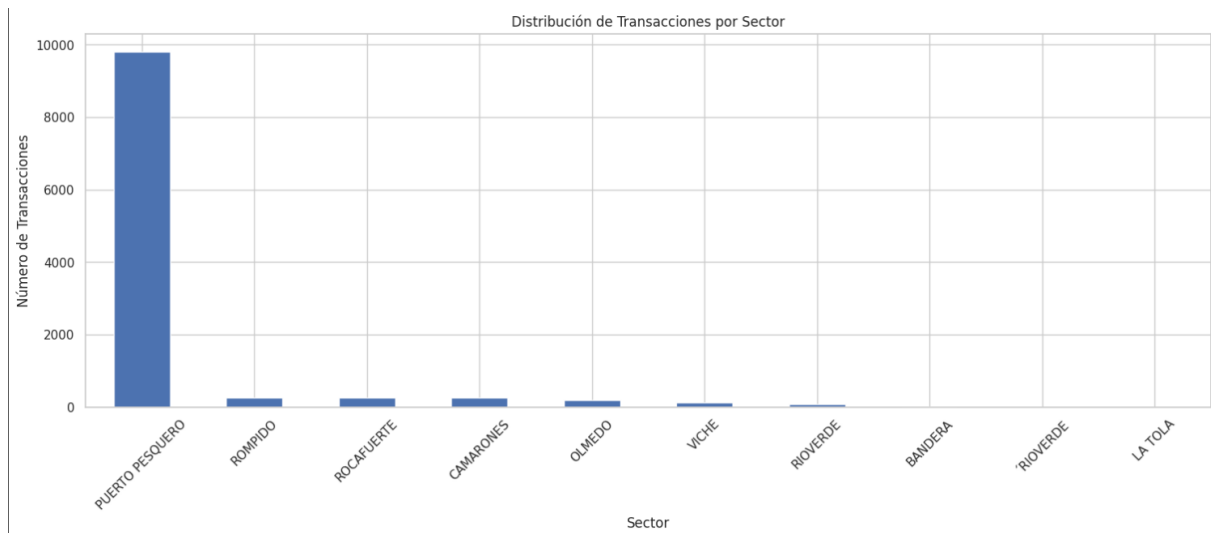
Galones (Distribución del Consumo de Combustible): Un gráfico de cajas (boxplot) mostrando la distribución del consumo de combustible para evaluar la distribución del consumo de combustible e identificar outliers.



Embarcación Nombre (Frecuencia de Uso de Embarcaciones): Un histograma que muestra la frecuencia de uso de cada embarcación para ver qué embarcaciones son las más utilizadas.



Sector (Distribución de Transacciones por Sector): Un gráfico de barras mostrando el número de transacciones por sector para comparar la actividad entre diferentes sectores.

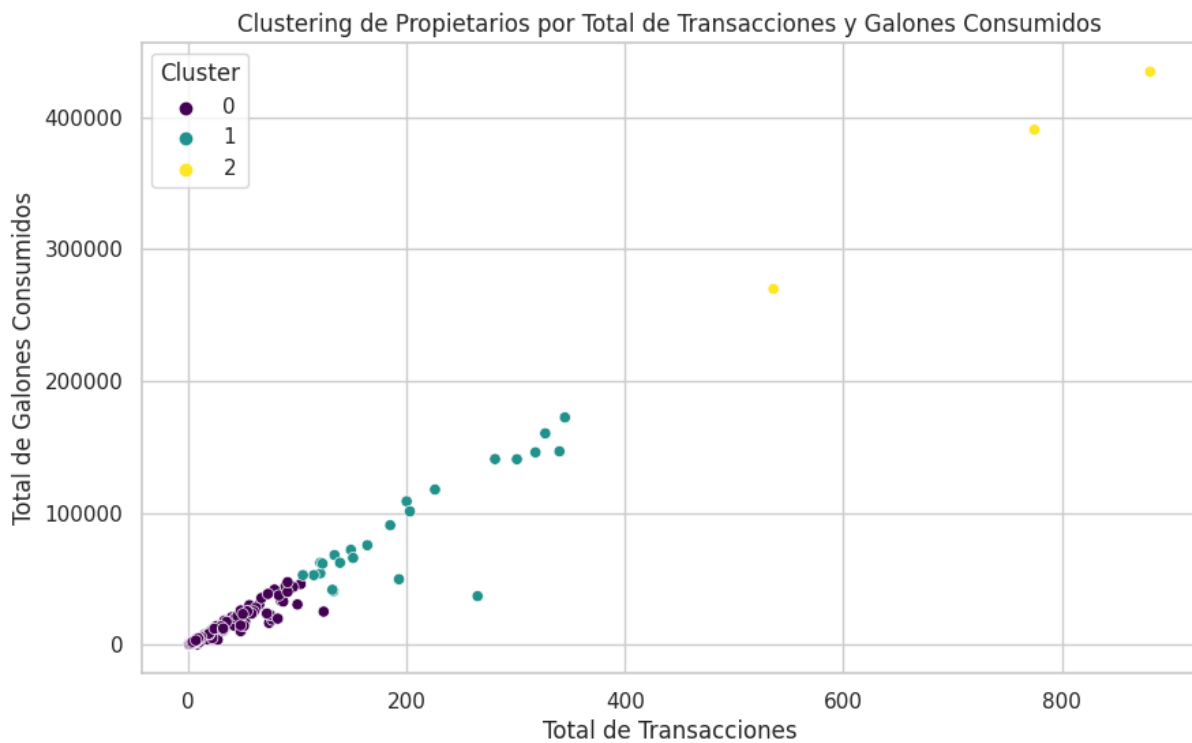


Algoritmos Utilizados

Hipótesis: "Las embarcaciones propiedad de individuos con múltiples transacciones tienden a tener un mayor consumo de combustible y actividad (medido por registros de zarpe)."

Pasos:

1. Agrupar a los propietarios por la cantidad de transacciones realizadas.
2. Calcular el consumo total de combustible para cada propietario.
3. Normalización de Datos: Normalizar los datos para asegurar que cada característica contribuya equitativamente al análisis.
4. Aplicar un algoritmo de clustering, K-means, para identificar grupos de propietarios con patrones similares de transacciones y consumo de combustible.



Análisis de resultados:

Cluster 0 (morado): Este grupo mayoritario incluye propietarios con una cantidad relativamente baja a moderada de transacciones y consumo de combustible. La mayoría de los propietarios caen en este grupo.

Cluster 1 (verde): No se observa claramente en el gráfico, lo que sugiere que podría haber pocos o ningún propietario en este grupo, o que sus características son muy similares a las del Cluster 0.

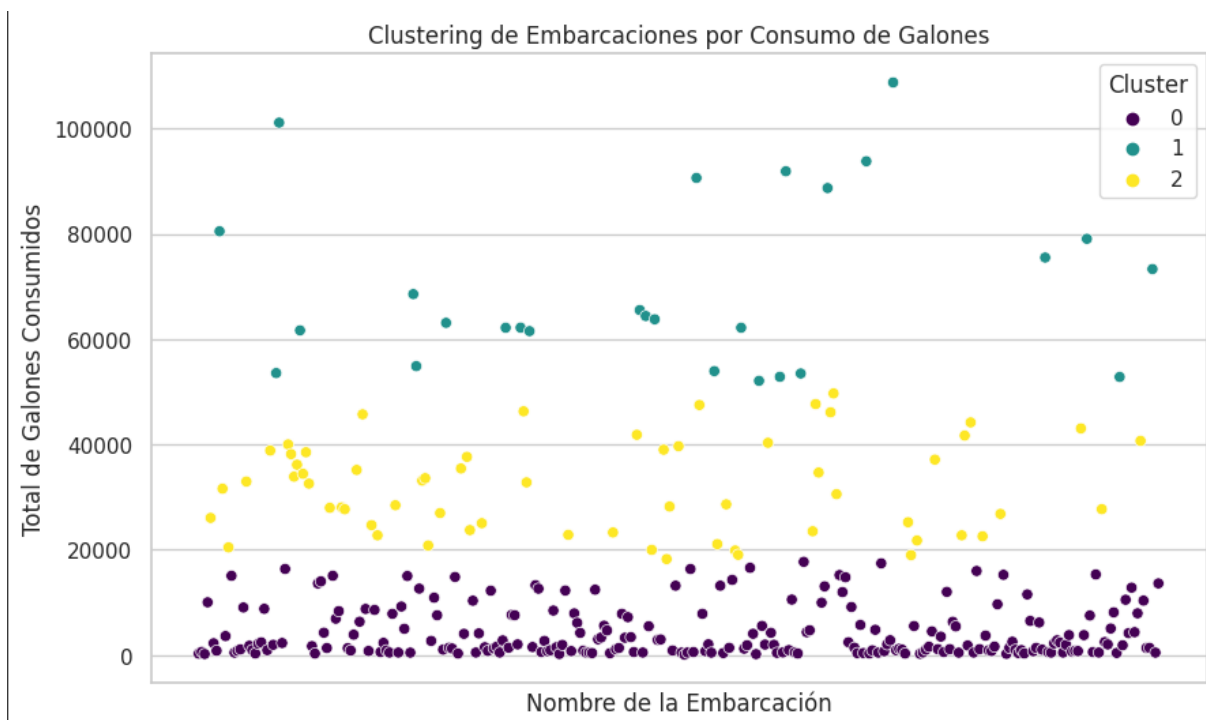
Cluster 2 (amarillo): Este grupo, aunque más pequeño, es distintivo. Incluye propietarios con un alto número de transacciones y un consumo de combustible significativamente mayor. Estos son probablemente los usuarios más activos y de mayor volumen en términos de operaciones portuarias.

La presencia de estos grupos distintos apoya la hipótesis de que existen diferencias significativas en el comportamiento de consumo y actividad entre los propietarios, en función de la frecuencia de sus transacciones y el consumo de combustible asociado.

Hipótesis: "El tipo de embarcación influye en el volumen de combustible consumido."

Pasos:

1. Agrupar las transacciones por embarcación y calcular el total de galones de combustible consumidos para cada embarcación.
2. Normalización de Datos: Normalizar los datos para asegurar que cada característica contribuya equitativamente al análisis.
3. Aplicar un algoritmo de clustering, K-means, para identificar patrones de consumo de combustible en relación con las embarcaciones específicas.



Análisis de resultados:

Cluster 0 (morado): Bajo Consumo de Combustible las embarcaciones que regularmente consumen menos combustible.

Cluster 1 (verde): Consumo Moderado de Combustible las embarcaciones con un consumo de combustible intermedio.

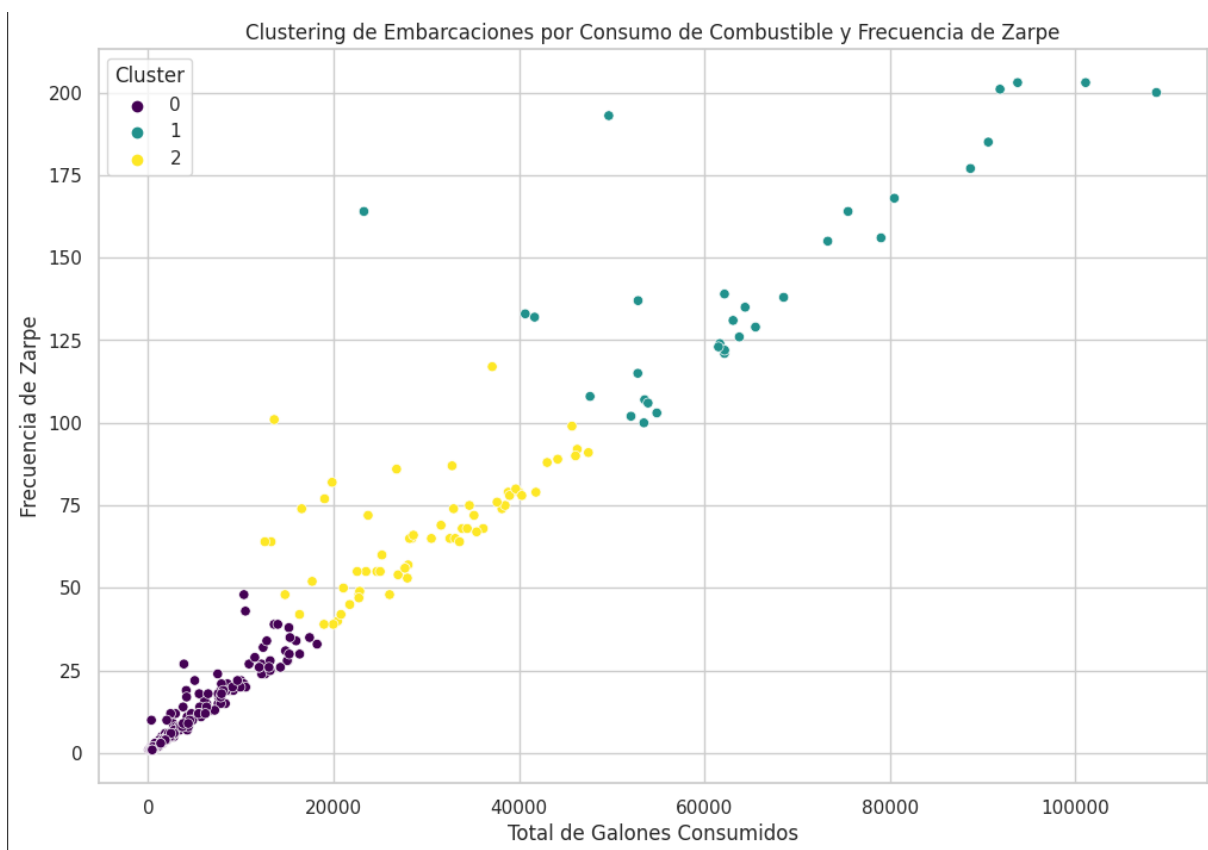
Cluster 2 (amarillo): Alto Consumo de Combustible las embarcaciones que consumen cantidades significativamente mayores de combustible, posiblemente debido a un uso más intensivo, mayores capacidades de carga o viajes más largos.

Este análisis podría proporcionar información valiosa sobre cómo diferentes tipos de embarcaciones contribuyen al consumo total de combustible, y podría ser útil para la planificación de recursos y la gestión ambiental en el sector marítimo.

Hipótesis: "Las embarcaciones pueden agruparse en diferentes tipos de actividad basándose en su consumo de combustible y frecuencia de zarpe."

Pasos:

1. Preparación de Datos: Crear un nuevo DataFrame que contenga el total de galones consumidos y la frecuencia de zarpe para cada embarcación.
2. Normalización de Datos: Normalizar los datos para asegurar que cada característica contribuya equitativamente al análisis.
3. Aplicación de K-means: Aplicar el algoritmo K-means para agrupar las embarcaciones.
4. Análisis de los Clusters Resultantes: Evaluar los clusters resultantes para interpretar los tipos de actividad de las embarcaciones.



















Análisis de resultados:

Cada punto representaría una embarcación, ubicado en el gráfico según su consumo total de combustible (eje x) y su frecuencia de zarpe (eje y).

Los colores de los puntos indicarían a qué cluster pertenece cada embarcación, permitiendo visualizar cómo se agrupan las embarcaciones con patrones similares de consumo y actividad.

Este análisis de clustering ayuda a identificar patrones de comportamiento entre las embarcaciones y podría ser utilizado para tomar decisiones operativas o estratégicas relacionadas con la gestión de flotas, la planificación de recursos o la implementación de políticas de eficiencia energética.

Data Warehouse

 PuertoPesquero  	
 Ventas	
 embarcacion_matricula	
 embarcacion_nombre	
 factura	
 fecha	
 galones	
 motores	
 propietario_cedula	
 propietario_nombre	
 sector	
 tripulante_cedula	
 tripulante_nombre	
 zarpe	

sector	
<u>Storage size:</u> 4.10 kB	<u>Documents:</u> 11

tripulante_cedula	
<u>Storage size:</u> 4.10 kB	<u>Documents:</u> 248

tripulante_nombre	
<u>Storage size:</u> 4.10 kB	<u>Documents:</u> 234

Ventas	
<u>Storage size:</u> 4.10 kB	<u>Documents:</u> 11 K

zarpe	
<u>Storage size:</u> 4.10 kB	<u>Documents:</u> 1.9 K

fecha

Storage size:
4.10 kB

Documents:
734

galones

Storage size:
4.10 kB

Documents:
394

motores

Storage size:
4.10 kB

Documents:
29

propietario_cedula

Storage size:
4.10 kB

Documents:
249

propietario_nombre

Storage size:
4.10 kB

Documents:
232

embarcacion_matricula

Storage size:
4.10 kB

Documents:
310

embarcacion_nombre

Storage size:
4.10 kB

Documents:
323

factura

Storage size:
4.10 kB

Documents:
10 K

Descripción de cada Dato

1. fecha: La fecha en la que se realizó la venta o transacción dividida en día, mes, año.
2. propietario_nombre: Nombre del propietario de la embarcación.
3. propietario_cedula: Cédula de identidad del propietario de la embarcación.
4. tripulante_nombre: Nombre del tripulante de la embarcación.
5. tripulante_cedula: Cédula de identidad del tripulante de la embarcación.
6. embarcacion_nombre: Nombre de la embarcación.
7. embarcacion_matricula: Número de matrícula de la embarcación.
8. galones: Cantidad de galones de combustible u otro producto vendido.
9. motores: Especificaciones de los motores de la embarcación.
10. zarpe: Número de registro o código de zarpe (salida de la embarcación).

11. factura: Número de la factura emitida por la venta.

12. sector: El sector o ubicación donde se realizó la venta, por ejemplo, un puerto pesquero.

Se desea descubrir las necesidades del negocio utilizando el Power BI para realizar un análisis estadístico como, por ejemplo:

- La tendencia de ventas (cantidad de galones) a lo largo del tiempo.
- Cantidad de galones vendidos por embarcación.
- Relación entre la cantidad de motores (como un indicador de tamaño de la embarcación) y la cantidad de galones vendidos.

Paso 1: Actividad a Modelar

En el caso del archivo **VENTAS.csv**, la actividad a modelar sería el análisis estadístico descrito anteriormente).

Paso 2: Identificación de Entidades Clave

- **Usuarios:** Aquí se podrían considerar los propietarios y los tripulantes de las embarcaciones como los "usuarios" del servicio de venta.
- **Navegadores:** En este contexto, no se trata de navegadores web, sino de las embarcaciones mismas.
- **Ubicación:** El sector donde se realiza la venta (por ejemplo, "PUERTO PESQUERO").

Paso 3: Definición de Dimensiones

- **Dimensión de Usuario:** Información detallada de los propietarios y tripulantes, como nombre y cédula.
- **Dimensión de Embarcación:** Detalles de la embarcación, incluyendo nombre, matrícula y especificaciones de motores.
- **Dimensión de Transacción:** Información sobre la transacción realizada, como fecha, cantidad de galones vendidos y número de factura.

Paso 4: Datos Específicos y Métodos de Análisis

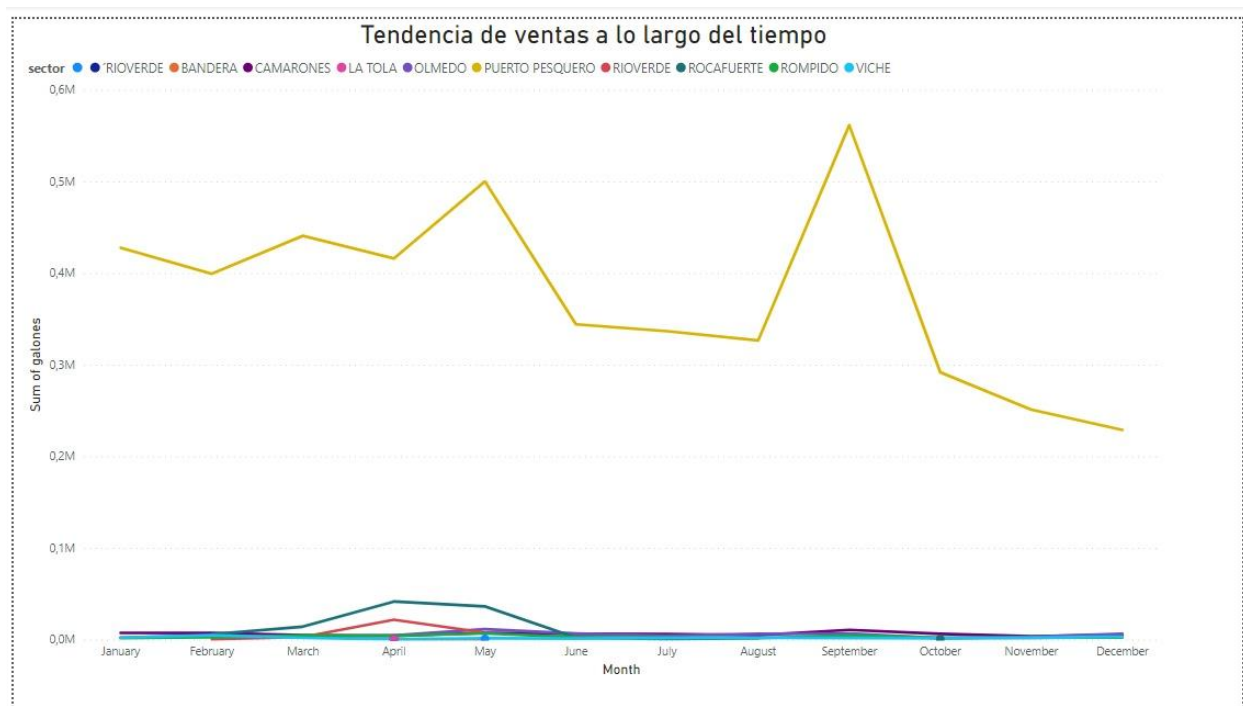
- **Transacción:** Datos como fecha (día, mes, año), número de factura, cantidad de galones.
- **Usuario:** ID (podría ser la cédula del propietario o tripulante), tipo de embarcación (interpretado como "navegador"), y otros detalles relevantes.
- **Ubicación:** Sector o puerto donde se realiza la venta.

- **Métodos de Análisis:** En este caso, podrían ser métodos estadísticos y de minería de datos para analizar patrones de venta, preferencias de clientes, eficiencia de tipos de embarcaciones, etc.

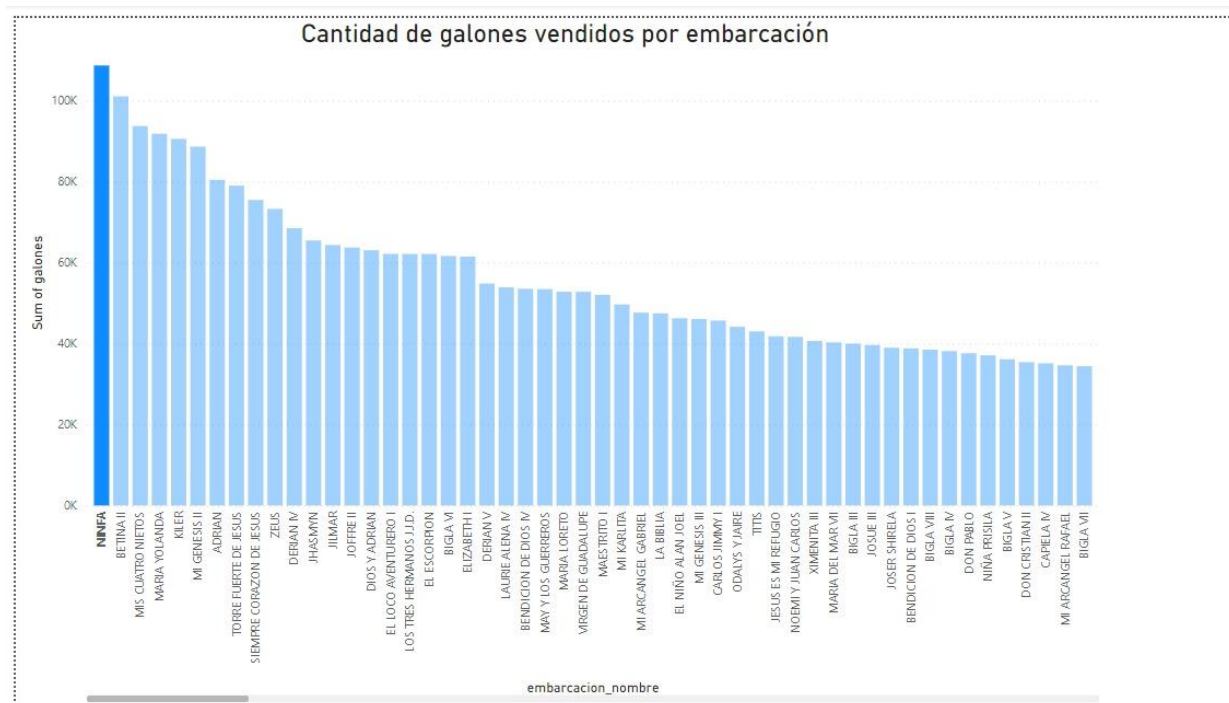
Interpretación de los Datos

- Se trabajó con el dataset completo para obtener una comprensión más amplia.

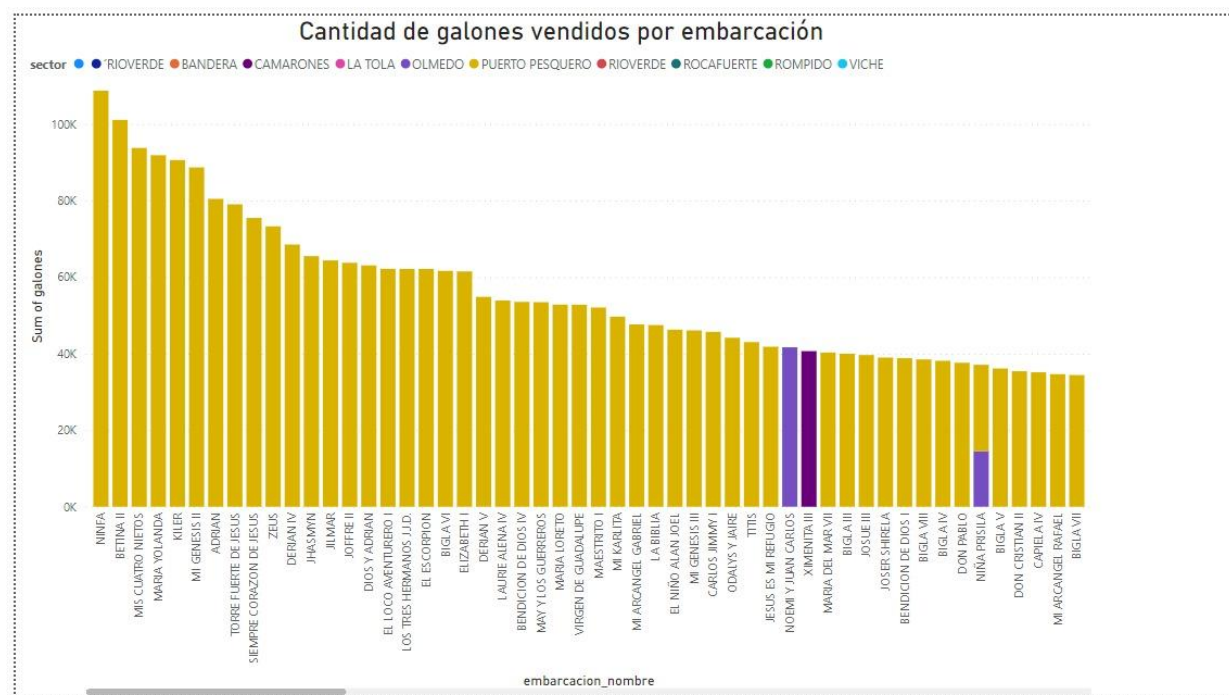
Tendencia de ventas (cantidad de galones) a lo largo del tiempo.



Cantidad de galones vendidos por embarcación.



Relación entre la cantidad de motores (como un indicador de tamaño de la embarcación) y la cantidad de galones vendidos.



Conclusiones:

1. Los gráficos de frecuencia de uso de embarcaciones revelan una diversidad significativa en cómo se utilizan las embarcaciones, con algunas siendo mucho más activas que otras.
2. Los histogramas y boxplots del consumo de combustible muestran una amplia gama de comportamientos en cuanto al consumo, desde muy bajo hasta excepcionalmente alto.
3. Al visualizar los datos, especialmente en gráficos de dispersión o boxplots, se pueden identificar outliers o anomalías que podrían indicar errores en los datos, prácticas inusuales o áreas para una investigación más profunda.
4. Las embarcaciones con propietarios que realizan múltiples transacciones muestran patrones de consumo de combustible más altos, lo que podría indicar un uso más intensivo o frecuente.
5. El análisis de consumo de combustible puede ofrecer insights sobre el impacto ambiental de diferentes tipos de embarcaciones y operaciones.
6. Estos datos pueden informar a las políticas y regulaciones relacionadas con la actividad marítima, promoviendo prácticas más sostenibles y eficientes en el uso de combustible.
7. Los datos muestran una amplia variedad de usuarios (propietarios y tripulantes) y embarcaciones, lo que indica una clientela diversa y posiblemente necesidades variadas. Esto sugiere que el servicio de ventas podría beneficiarse de estrategias de personalización y segmentación para atender mejor a cada grupo de clientes.
8. La cantidad de galones vendidos y los tipos de motores de las embarcaciones ofrecen insights sobre los patrones de compra y preferencias de combustible. Estos datos pueden ser cruciales para gestionar el inventario, optimizar la oferta de combustible y planificar promociones o descuentos específicos.