Proyecto 1

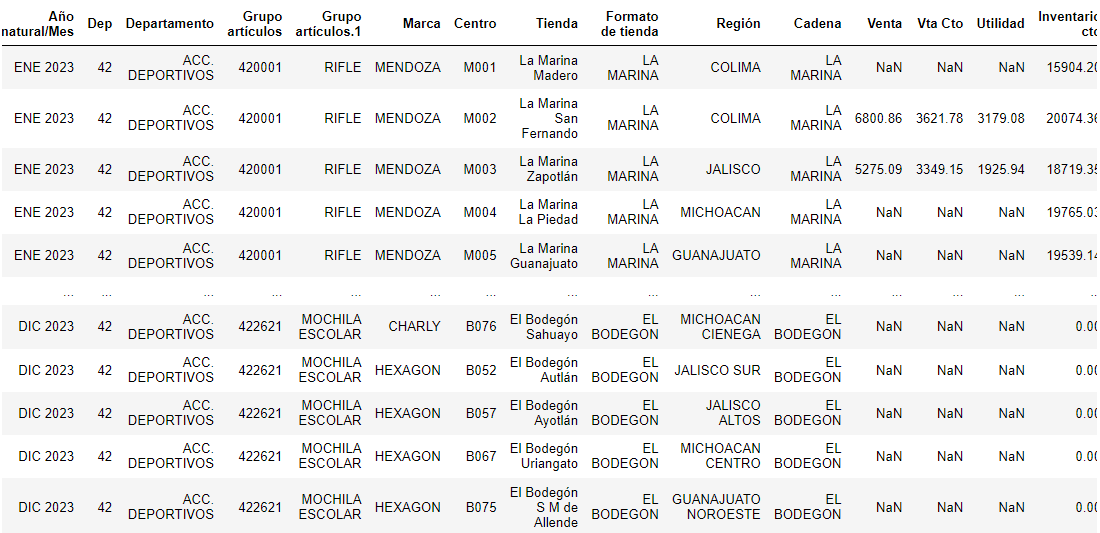
**Parte 1**

Utiliza la base de datos "BASE 1", realiza un análisis descriptivo al mayor detalle posible.

El objetivo entender y explicar el comportamiento del departamento de Acc Deportivos

**Análisis Descriptivo:**

Veamos los datos



Se puede observar que tenemos 17 variables y contiene un total de 124751 datos.

Hacemos una revisión de lo que pueden significar las variables:

𝐀ñ𝐨 𝐧𝐚𝐭𝐮𝐫𝐚𝐥/𝐌𝐞𝐬: Representa el año y mes de los datos.

𝐃𝐞𝐩: Código de departamento.

𝐃𝐞𝐩𝐚𝐫𝐭𝐚𝐦𝐞𝐧𝐭𝐨: Nombre del departamento (ACC. DEPORTIVOS significa accesorios deportivos).

𝐆𝐫𝐮𝐩𝐨 𝐚𝐫𝐭í𝐜𝐮𝐥𝐨𝐬:  Identificador para un grupo específico de artículos.

𝐆𝐫𝐮𝐩𝐨 𝐚𝐫𝐭í𝐜𝐮𝐥𝐨𝐬:  Identificador adicional del grupo de productos como puede ser el nombre/tipo del producto.

𝐌𝐚𝐫𝐜𝐚: Nombre de la marca.

𝐂𝐞𝐧𝐭𝐫𝐨: Código de la tienda.

𝐓𝐢𝐞𝐧𝐝𝐚: Nombre de la tienda específica con su ubicación.

𝐅𝐨𝐫𝐦𝐚𝐭𝐨 𝐝𝐞 𝐭𝐢𝐞𝐧𝐝𝐚: Nombre de la tienda.

𝐑𝐞𝐠𝐢ó𝐧: Estado de donde se encuentra la tienda.

𝐂𝐚𝐝𝐞𝐧𝐚: Cadena de la tienda.

𝐕𝐞𝐧𝐭𝐚: Monto de ventas.

𝐕𝐭𝐚 𝐂𝐭𝐨: Costo de las ventas.

𝐔𝐭𝐢𝐥𝐢𝐝𝐚𝐝: Ganancia, calculada como la diferencia entre los ingresos de Venta y Vta Cto.

𝐈𝐧𝐯𝐞𝐧𝐭𝐚𝐫𝐢𝐨 𝐜𝐭𝐨: Costo del inventario.

𝐕𝐭𝐚 𝐩𝐳𝐚𝐬: Cantidad de piezas vendidas.

𝐈𝐧𝐯𝐞𝐧𝐭𝐚𝐫𝐢𝐨 𝐩𝐳𝐚𝐬: Cantidad de piezas en el inventario.

Primero se hará un análisis de correlación entre las variables o ver sus rangos de valores porque parece que podemos omitir algunas columnas que no nos aportan.

El primer ejemplo sería quitar la columna ‘Departamento’ y ‘Dep’, que solo tienen un valor.

Debido a que hay valores NaN en las columnas de "Venta", "Vta Cto", "Utilidad" y "Vta pzas", estos los podemos llenar con 0 porque justamente cuanto no hay datos en la variable "Vta Pzas" tampoco hay información en las variables "Venta", "Vta Cto", "Utilidad", lo cual tiene sentido ya que si no hay venta de piezas no hay ganancias.

**Aquí mostramos cuáles son las variables categóricas y numéricas**

Columnas Categóricas

--Año natural/Mes: Categorica con 12 valores diferentes. Ejemplo: ['ENE 2023', 'FEB 2023', 'MAR 2023', 'ABR 2023', 'MAY 2023'].

--Grupo artículos.1: Categorica con 229 valores diferentes. Ejemplo: ['RIFLE', 'PISTOLA', 'DIABOLOS MUNICIONES', 'ACCESORIOS PARA CAZA', 'MIRAS TELESCOPICAS'].

--Marca: Categorica con 151 valores diferentes. Ejemplo: ['MENDOZA', 'GAMO', 'CROSMAN', 'XISICO', 'HATSAN'].

--Centro: Categorica con 86 valores diferentes. Ejemplo: ['M001', 'M002', 'M003', 'M004', 'M005'].

--Tienda: Categorica con 86 valores diferentes. Ejemplo: ['La Marina Madero', 'La Marina San Fernando', 'La Marina Zapotlán', 'La Marina La Piedad', 'La Marina Guanajuato'].

--Formato de tienda: Categorica con 2 valores diferentes. Ejemplo: ['LA MARINA', 'EL BODEGON'].

--Región: Categorica con 14 valores diferentes. Ejemplo: ['COLIMA', 'JALISCO', 'MICHOACAN', 'GUANAJUATO', 'JALISCO SUR'].

--Cadena: Categorica con 2 valores diferentes. Ejemplo: ['LA MARINA', 'EL BODEGON'].

--Grupo artículos: Categorica con 249 valores diferentes. Ejemplo: [420001, 420002, 420003, 420004, 420005].

Columnas Numéricas

--Venta: El rango de valores en esta columna es de -29322.93 a 153384.92.

--Vta Cto: El rango de valores en esta columna es de -16983.6 a 93268.51.

--Utilidad: El rango de valores en esta columna es de -26259.32 a 61355.64.

--Inventario cto: El rango de valores en esta columna es de -6033.88 a 246154.77.

--Vta pzas: El rango de valores en esta columna es de -6.0 a 252.0.

--Inventario pzas: El rango de valores en esta columna es de -5 a 639.

**Matriz de correlación para ver cuáles variables se están repitiendo o no están aportando información adicional**

A colorful chart with black and red squares

Description automatically generated with medium confidence

Se puede observar que en las categóricas hay una relación de 1, por ejemplo, Centro tiene relación de 1 con Tienda, Formato de tienda, Región y Cadena. Entonces utilizando esto se podría quitar las demás variables y solo dejar Tienda, ya que esa información no nos podría importar más información. También hay una relación perfecta entre Grupo artículos.1 y Grupo artículos.

También hay mucha relación entre la variable Forma de tienda con Marca y Grupo artículos, y entre las variables Venta y Vta Cto, teniendo una correlación muy cercana a 1, y también con utilidad.

Entonces al final de tener 17 columnas pasamos a tener 10 columnas. Esto ayuda a tener menos datos que analizar y los datos que quitamos no aportan mucho o prácticamente no aportan nada

Hacemos una descripción de los datos restantes

A screenshot of a computer

Description automatically generated

Se puede notar que la diferencia 75% a max para las variables "Venta", "Vta Cto", "Utilidad", "Inventario cto", "Vta pzas" y "Inventario pzas" es muy grande, esto se debe a la cantidad de 0's que hay.

**Análisis de cuáles productos no se están vendiendo, que tenga venta de piezas igual a 0**

**Agrupamos por variable**

**A screenshot of a computer

Description automatically generatedA screenshot of a computer

Description automatically generated**

La marca Coleman encabeza la lista con 8,624 ocurrencias de productos sin vender

Las tiendas La Marina San Fernando, La Marina La Piedad, y La Marina Guanajuato son las que presentan el mayor número de productos sin venta, con más de 7,000 ocurrencias cada una.

Los rifles y accesorios para caza son los grupos de artículos con la mayor cantidad de productos que no se vendieron, con más de 5,000 ocurrencias cada uno. Esto podría indicar una baja demanda.

De agosto a octubre 2023 muestran la mayor cantidad de productos sin venta, con más de 9,300 ocurrencias cada uno. Hubo una acumulación significativa de inventario sin movimiento.

**Agrupamos cada combinación de 2 variables**

**A screenshot of a computer

Description automatically generated** **A screenshot of a computer

Description automatically generated**

Los rifles tienen consistentemente el mayor número de ventas cero a lo largo de varios meses, con mayo de 2023 mostrando el pico más alto con 500 instancias.

La marca Coleman presenta el mayor número de ventas cero en diferentes meses, con agosto de 2023 liderando con 766 instancias.

La Marina Guanajuato y La Marina San Fernando son las tiendas con más ventas cero en varios meses, con septiembre de 2023 y octubre de 2023 teniendo los conteos más altos.

Ciertos grupos de productos como "MEDICINA DEL DEPORTE" con la marca LOB y "BICI. MONT.R27.5 C/S" con MERCURIO muestran conteos extremadamente altos de ventas cero (más de 1,000 cada uno).

El grupo "ACC. PARA DEPORTE EN" (presumiblemente accesorios para deportes) tiene las ventas cero más altas en varias tiendas como La Marina Guanajuato y La Marina San Fernando.

La marca NIKE tiene un número abrumador de ventas cero, particularmente en La Marina Guanajuato y La Marina La Piedad (más de 1,000 cada una).

**Agrupamos cada combinación de 3 variables**

A screenshot of a computer

Description automatically generated

FEB 2023 y OCT 2023 muestran los mayores conteos de ventas cero para productos como "MEDICINA DEL DEPORTE" con la marca ACCU-CHECK y "BICI MONT R26" con las marcas BDS y VELOCI, cada una con 86 instancias.

El grupo "ACC. PARA DEPORTE EN" muestra los mayores números de ventas cero en tiendas como La Marina Guanajuato y La Marina San Fernando, particularmente en SEP 2023 y FEB 2023.

La marca NIKE tiene un número elevado de ventas cero en tiendas como La Marina San Fernando y La Marina Guanajuato, especialmente en NOV 2023 y DIC 2023, con hasta 96 instancias.

El grupo "NO UTILIZAR (GUANTES)" con la marca NIKE tiene un número muy alto de ventas cero en varias tiendas, como La Marina San Fernando y La Marina Madero, con hasta 72 instancias.

**Agrupamos cada combinación de 4 variables**

A screenshot of a computer

Description automatically generated

El grupo "NO UTILIZAR (GUANTES)" bajo la marca NIKE muestra consistentemente ventas cero en varios meses en la tienda La Marina San Fernando. Cada instancia tiene 6 ocurrencias, cubriendo los meses de agosto 2023, noviembre 2023, febrero 2023, marzo 2023, y septiembre 2023.

**Análisis de cuáles productos no se están vendiendo, que tenga venta de piezas diferente a 0**

A screenshot of a computer screen

Description automatically generated A screenshot of a computer

Description automatically generated

La marca NEW ERA encabeza la lista con 4,732 piezas vendidas, seguida por ADIDAS con 4,143 piezas.

La Marina San Fernando es la tienda con más ventas, con 7,996 piezas vendidas, seguida por La Marina Zapotlán con 5,847 piezas.

El grupo GORRA es el líder en ventas con 6,518 piezas vendidas, seguido de DIABOLOS MUNICIONES con 3,112 piezas y estos son particularmente populares entre los consumidores

Diciembre 2023 lidera con 4,339 piezas vendidas lo que podría estar relacionado con navidad

**Agrupamos cada combinación de 2 variables**

**A screenshot of a computer screen

Description automatically generated** **A screenshot of a computer

Description automatically generated**

Diciembre 2023 lidera con la venta de GORRA (1,201 piezas), seguido por MOCHILA ESCOLAR en agosto 2023 con 1,002 piezas.

NEW ERA en diciembre 2023 es la marca más vendida con 899 piezas, seguida por PUMA en Agosto 2023 y Diciembre 2023.

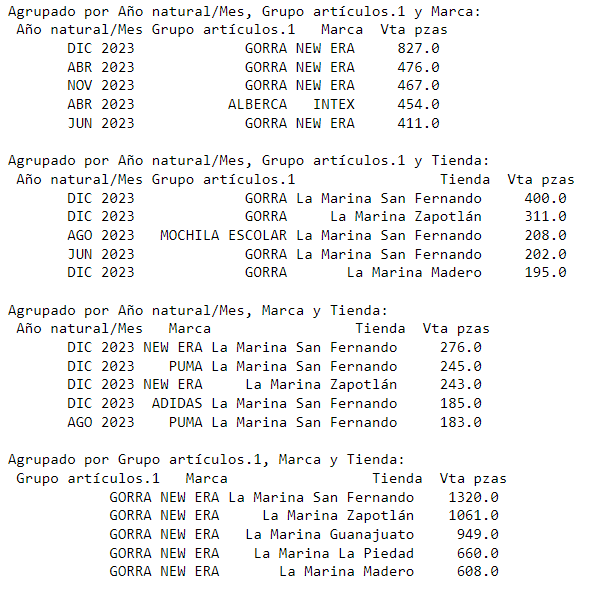
La Marina San Fernando en diciembre 2023 tiene las mayores ventas con 1,185 piezas, seguida por La Marina Zapotlán en el mismo mes con 777 piezas.

GORRA bajo la marca NEW ERA es el producto más vendido, con 4,598 piezas, seguido por DIABOLOS MUNICIONES de la marca GAMO con 1,563 piezas.

GORRA es el artículo más vendido en La Marina San Fernando con 1,962 piezas, seguido por la misma categoría en La Marina Zapotlán con 1,453 piezas.

NEW ERA en La Marina San Fernando lidera las ventas con 1,363 piezas, seguida de cerca por ADIDAS en la misma tienda con 1,357 piezas

**Agrupamos cada combinación de 3 variables**

****

Diciembre 2023 lidera con la venta de GORRA de la marca NEW ERA (827 piezas), seguido por Abril 2023 con la misma combinación (476 piezas).

GORRA en La Marina San Fernando en diciembre 2023 tiene las mayores ventas con 400 piezas, seguida por la misma categoría en La Marina Zapotlán (311 piezas).

NEW ERA en La Marina San Fernando en diciembre 2023 lidera las ventas con 276 piezas, seguido por PUMA en la misma tienda durante el mismo mes (245 piezas).

GORRA de la marca NEW ERA en La Marina San Fernando lidera las ventas con 1,320 piezas, seguida por la misma combinación en La Marina Zapotlán (1,061 piezas)

**Agrupamos cada combinación de 4 variables**

**A white background with black text

Description automatically generated**

Diciembre 2023 tiene las mayores ventas para la GORRA de NEW ERA en La Marina San Fernando con 252 piezas, seguido por La Marina Zapotlán con 224 piezas y La Marina Guanajuato con 150 piezas, lo que indica una fuerte estacionalidad probablemente impulsada por las compras navideñas.

**Ahora vamos a checar las utilidades**

**Agrupamos cada combinación de 1 variables para ver cuales productos tienen mayor y menor utilidad**

**Mayor Utilidad**

**A screenshot of a computer

Description automatically generated** **A screenshot of a computer

Description automatically generated**

Diciembre 2023 es el mes con la mayor utilidad generada, alcanzando 1,933,336.62 unidades monetarias, seguido por Marzo 2023 con 1,851,015.08

El grupo RIFLE encabeza la lista con una utilidad de 3,120,474.12unidades monetarias, seguido por BICI MONT R29 con 2,502,594.24

La marca MERCURIO lidera en términos de utilidad generada, alcanzando 2,822,463.97 unidades monetarias, seguida por GAMO con 1,787,734.48

La Marina San Fernando es la tienda que más utilidad genera, con 1,867,816.46 unidades monetarias, seguida por La Marina Zapotlán con 1,515,315.77

Productos de alto valor como rifles y bicicletas, así como marcas relacionadas con actividades al aire libre, están entre los más rentables. Las tiendas La Marina San Fernando y La Marina Zapotlán emergen como líderes en generación de utilidades

**Menor Utilidad**

**A screenshot of a computer

Description automatically generated** **A screenshot of a computer

Description automatically generated**

Septiembre 2023 es el mes con la menor utilidad positiva, generando 976,701.28 unidades monetarias, seguido de mayo 2023 con 996,885.65.

BICI. RUTA R700 y MONITORES DEPORTIVOS son los grupos de artículos que generan las menores utilidades, con pérdidas de -34,672.40 y -34,569.80 unidades monetarias, respectivamente.

Las marcas BELFORT y GARMIN son las que generan las menores utilidades, con pérdidas de -43,296.81 y -35,834.09 unidades monetarias, respectivamente.

El Bodegón Manzanillo II es la tienda con la utilidad positiva más baja, con 38,694.92 unidades monetarias, seguida por El Bodegón Teocaltiche con 49,704.06.

Las bicicletas y productos tecnológicos, como monitores deportivos, junto con marcas como BELFORT y GARMIN, están generando pérdidas

**Agrupamos cada combinación de 2 variables para ver cuales productos tienen mayor y menor utilidad**

**Mayor Utilidad**

**A white screen with black text

Description automatically generated** **A white screen with black text

Description automatically generated**

Marzo 2023 lidera con la mayor utilidad en el grupo de ALBERCA (piscinas) con 429,194.69 unidades monetarias, seguido de Abril 2023 con 381,140.48 para el mismo grupo.

Marzo 2023 y Abril 2023 son meses exitosos para la marca INTEX con utilidades superiores a 400,000 unidades monetarias.

La Marina San Fernando genera la mayor utilidad en Diciembre 2023 con 312,284.43 unidades monetarias, seguida por La Marina Zapotlán con 205,296.50.

RIFLES de la marca GAMO generan la mayor utilidad con 1,485,177.22 unidades monetarias. BICI MONT R29 de la marca MERCURIO también es altamente rentable, con más de 1,200,000 unidades monetarias en utilidad.

GORRA en La Marina San Fernando es extremadamente rentable, generando 427,582.19 unidades monetarias.

NEW ERA en La Marina San Fernando lidera las utilidades, generando 331,522.97 unidades monetaria

**Menor Utilidad**

**A screenshot of a computer

Description automatically generated** **A screenshot of a computer

Description automatically generated**

Marzo 2023 lidera con la menor utilidad en el grupo de MONT.R27.5 C/S con -47,170.93 unidades monetarias y BICI. RUTA R700 -40,191.58

Abril 2023 y Mayo 2023 son meses menos exitosos para la marca BELFORT y GARMIN con -47,803.63 y -28,201.67 respectivamente de unidades monetarias.

Mayo 2023 y el bodegón Guzman II y Bodegon Ahualulco es el mes menos existoso con el bodegon ya que super las pérdidas de -9,000.16 de utilidades

Las Bicicletas de la marca BELFORT y MERCURIO superan las pérdidas de -35,000 de utilidades

Monitores deportivos en la marina san fernando y guanajuato lideran la tabla con pérdidas de -20,602.85 y -13,086.02 de unidades monetarias

Las marcas BELFORT y GARMIN en la Marina Zapotlán tiene las perdidas más altas arriba de -20,000 de utilidades

**Agrupamos cada combinación de 3 variables para ver cuales productos tienen mayor y menor utilidad**

**Mayor Utilidad**

**A screenshot of a computer

Description automatically generated**

En marzo y abril de 2023, las albercas de la marca INTEX generaron una utilidad de 422,997.44 y 380,009.50 respectivamente

En diciembre de 2023, las gorras en la tienda La Marina San Fernando y Zapotlán generaron utilidades mayores de 70,000.

En diciembre de 2023, las marcas NEW ERA en la tienda La Marina San Fernando y Zapotlán generaron utilidades mayores de 65,000.

Las gorrar de NEW ERA en la tienda La Marina San Fernando y Zapotlán generaron utilidades mayores de 28,000.

**Menor Utilidad**

**A screenshot of a computer

Description automatically generated**

En abril y Mayo de 2023, las bicicletas de montaña de la marca BELFORT y BDS generaron una pérdida arriba de 42,000.00.

En abril y Mayo de 2023, las bicicletas de montaña tienda El Bodegón Ahualulco y Marina Guanajuato generaron una pérdida de 13,000.00.

En abril de 2023, la marca BELFORT en la tienda La Marina Zapotlán y Guanajuato generó una pérdida de 19,000.

Las bicicletas y los monitores de la marca BELFORT en la tienda La Marina Zapotlán y San Fernando generaron una pérdida arriba 19,000 de utilidades

**Agrupamos cada combinación de4 variables para ver cuales productos tienen mayor y menor utilidad**

**Mayor Utilidad**

**A white background with black text

Description automatically generated**

Lo que más generó fue en diciembre las gorras marca NEW ERA en la marina Zapotlán y San Fernando superando las utilidades de 60,000

**Menor Utilidad**

**A white background with black text

Description automatically generated**

Lo que más generó perdidas fue en el mes de Abril 2023 la bici de montaña de la marca BELFORD en la marina Zapotlán y Guanajuato haciendo que la suma de las pérdidas llegua a más de 40,000

Parte 2

Utiliza la base de datos "BASE 2", realiza una proyeccion de la demanda para el año 2024

Para la proyeccion de la demanda puedes utilizar uno o varios metodos de pronosticos,

**Análisis Descriptivo:**

Veamos los datos

A table with numbers and symbols

Description automatically generated

Aquí podemos eliminar varias columnas que no nos proporcionan más información. Las cuáles son "Dep", "Departamento", "Grupo Artículos"

También hacemos una matriz de correlación como se muestra abajo:

Año natural/Mes está muy poco relacionado con las demás variables. Las que están muy relacionadas son:Compras Cto y Compras Pzas, Inv cto y Inv pza, Vta Cto y Vta Pza, Venta y Vta pzas y Ventas y Utilidad

A screenshot of a chart

Description automatically generated

A table with numbers and symbols

Description automatically generated

Los cuartiles indican que no hay outlayers en estos datos

Hacemos los pronósticos para el año 2024 para ventas y utilidades

A graph with blue and orange lines

Description automatically generated

A graph with blue lines and orange lines

Description automatically generated