

TECNOLÓGICO DE MONTERREY



INTELIGENCIA ARTIFICIAL AVANZADA
PARA LA CIENCIA DE DATOS II

TC3006C

**Reto: Arca Continental - Preguntas
Iniciales**

Autores:

Lautaro Gabriel Coteja - A01571214

Daniela Jiménez Téllez - A01654798

Andrés Villarreal González - A00833915

Héctor Hibran Tapia Fernández - A01661114

1. Planteamiento del problema

El problema central que aborda este reporte es el análisis a los clientes de la empresa Arca Continental para observar sus patrones de compras y lograr predecir su comportamiento. En este reporte se verá el análisis de la fluctuación en las ventas a lo largo de los meses, para descubrir patrones de ventas en diferentes periodos, ver qué meses tienen más o menos ventas, y analizar las primeras compras de los clientes. Estos son fundamentales para la toma de decisiones en estrategias de marketing y gestión de inventario. También se verán productos con baja recurrencia, productos de lanzamiento, y las tipologías de clientes que generen más ventas.

2. Hipótesis

- **Hipótesis 1:** Existen meses específicos que generan un mayor volumen de ventas, lo cual puede deberse a factores estacionales o eventos puntuales.
- **Hipótesis 2:** Los clientes tienden a comprar productos que no han adquirido anteriormente en ciertos periodos específicos del año.
- **Hipótesis 3:** Los productos de lanzamiento, especialmente aquellos con baja recurrencia, requieren mayor promoción o ajustes en las estrategias de venta para mejorar su desempeño.

Estas hipótesis se justifican con base en la necesidad de optimizar las estrategias de ventas y marketing para los productos y periodos con menor rendimiento, basándose en patrones de comportamiento de compra de los clientes.

3. Métodos usados

Se han utilizado métodos de análisis descriptivo y visualización de datos. Estos incluyen la agrupación de ventas por mes y año, normalización de datos de ventas, y el uso de gráficos de barras y líneas para identificar tendencias.

Además, se ha realizado la identificación de las primeras compras de los clientes utilizando las funciones de duplicación en pandas y se han identificado productos con baja recurrencia al calcular la suma total de ventas por producto y seleccionar aquellos que no alcanzan un umbral específico. Se utilizaron librerías como pandas, plotly, y seaborn para la manipulación y visualización de datos. A continuación se muestran las gráficas:

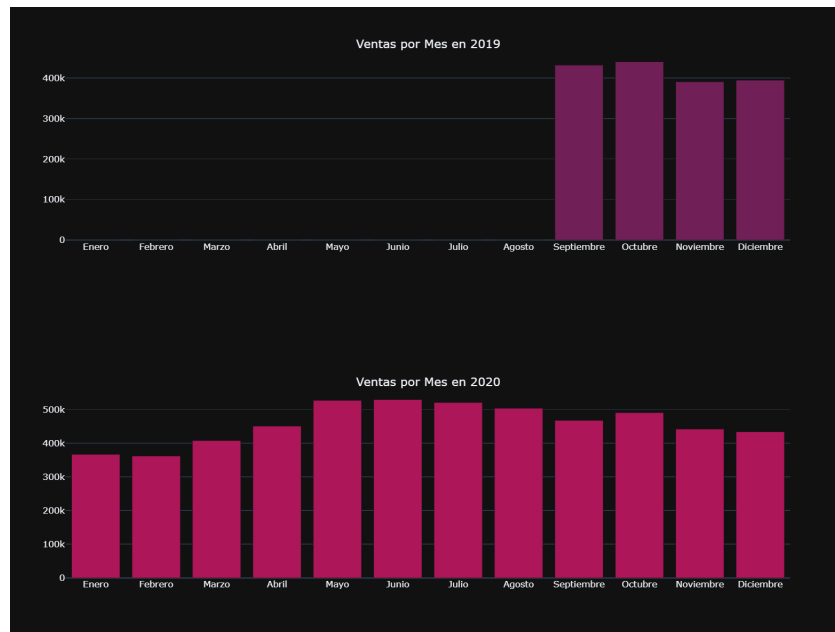


Figura 1: Ventas por Mes en los años 2019 y 2020

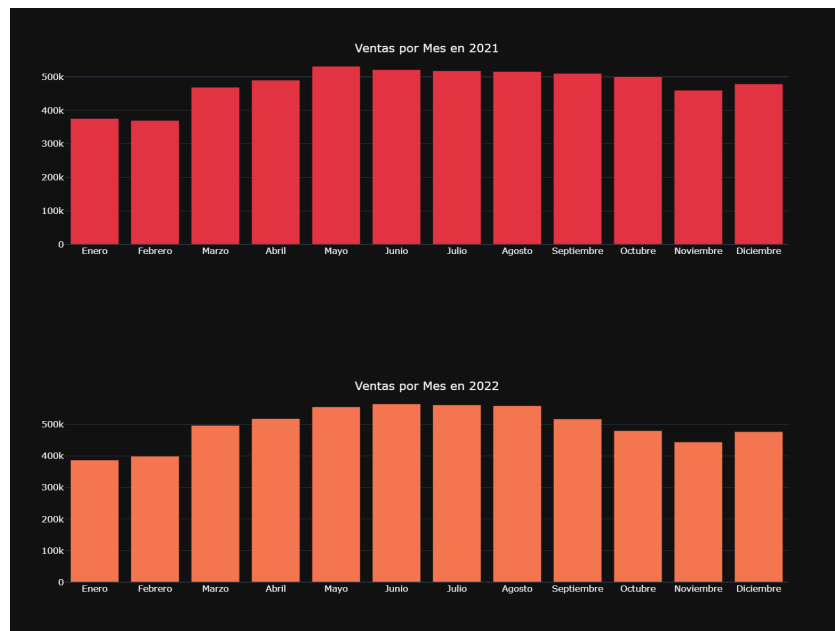


Figura 2: Ventas por Mes en los años 2021 y 2022

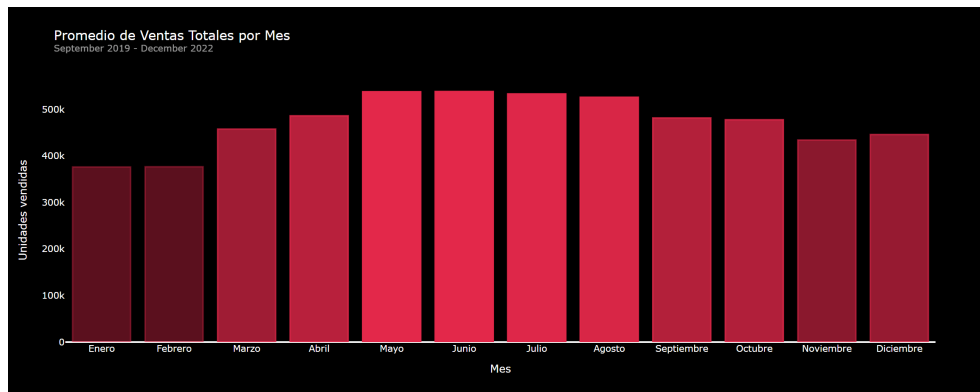


Figura 3: Promedio de Ventas Totales por Mes

Productos con Baja Recurrencia
Aquellos productos que han sido vendidos por primera vez y aún no tienen muchas ventas.

producto	calmonth	total_galones
DEL VALLE DURAZNO NECTAR LATA 335 ML 24B	2022-02-01T00:00:00	9.852528
CIEL EXPRIM MORAS 1 LT NR 6B	2019-09-01T00:00:00	9.510299999999999
SPRITE 1.50 LT RET PET 12 SPLA	2021-06-01T00:00:00	8.9840634
JOYA PONCHE 2.50 L NR PET 4	2022-09-01T00:00:00	8.806000000000001
COCA COLA 500 ML RT VID 6B	2020-10-01T00:00:00	8.3682708
MONSTER E PIPELINE 473 ML NR LAT 4 REFO	2021-04-01T00:00:00	8.33
ADES LACTEO VAINILLA 946ML TPK 12	2020-07-01T00:00:00	7.9972
SPRITE 235ML LATA 12B	2019-09-01T00:00:00	7.5339456
FLORIDA 7 MANZANA 200ML TETRA 40B	2021-08-01T00:00:00	7.0445
DEL VALLE GUAYABA NECTAR LATA 335 ML 6B	2019-09-01T00:00:00	7.020881999999999
COCA COLA LIGHT 355ML NR LATA 12 MPK	2021-11-01T00:00:00	6.7527
FRESCA 235ML LATA 12B	2019-11-01T00:00:00	6.415873899999999
COSTA DESCA TO MOL 500 G NR BOL 1	2022-08-01T00:00:00	6.164
T CH HARD S PINA T 355ML NR LAT 6 MPK	2021-01-01T00:00:00	5.9393334
FLORIDA 7 DURAZNO 200ML TETRA 40B	2021-08-01T00:00:00	5.6356
COCA COLA SIN AZUCAR 235ML NR LATA 24MPK	2021-05-01T00:00:00	4.9665

Figura 4: Productos con Baja Recurrencia

Productos más recientes
Estos son los productos que han tenido ventas más recientes según el 20% superior.

Producto	Primera Fecha de Venta
BT CIEL AP BIDON 5 LTS NR	2022-12-01T00:00:00
T CH AGUA MINERAL 355 ML NR VID 4B	2022-11-01T00:00:00
PULPY DURAZNO 400ML NRP 6B	2022-11-01T00:00:00
JOYA PONCHE 2.50 L NR PET 4	2022-09-01T00:00:00
FZT NG DZO 453 LATA 6B INK ART	2022-09-01T00:00:00
COSTA DESCA TO MOL 1 KG NR BOL 1	2022-09-01T00:00:00
TCH HARD S FRESA 355 ML NR LAT 6 REFO	2022-08-01T00:00:00
COSTA DESCA TO MOL 500 G NR BOL 1	2022-08-01T00:00:00
CIEL MINERALIZADA 355 ML NR LSL 12	2022-07-01T00:00:00
TCH HARD S PINA T 355 ML NR LAT 6 REFO	2022-07-01T00:00:00
TCH HARD S LIMA I 355 ML NR LAT 6 REFO	2022-07-01T00:00:00
FRESCA 250 ML NR 12B	2022-07-01T00:00:00
TCH HARDS MGO TROP 355 ML NR LAT 6 REFO	2022-07-01T00:00:00
COCA-COLA 3.00 L NR PET 9	2022-07-01T00:00:00
COCA-COLA 2.00 L NR PET 9	2022-07-01T00:00:00
MIX 473 ML LAT 4(1MPI1MMGO1MLOW1MORIG)	2022-07-01T00:00:00

Figura 5: Productos más recientes

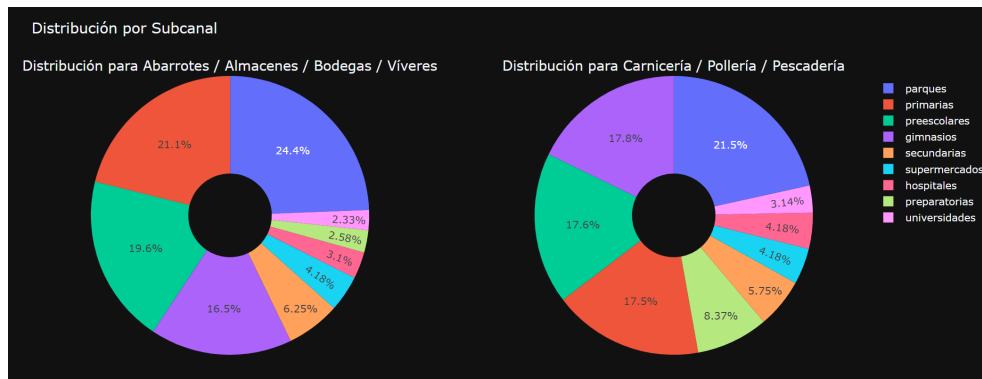


Figura 6: Gráfica de dona de la distribución de subcanal de diferentes clientes

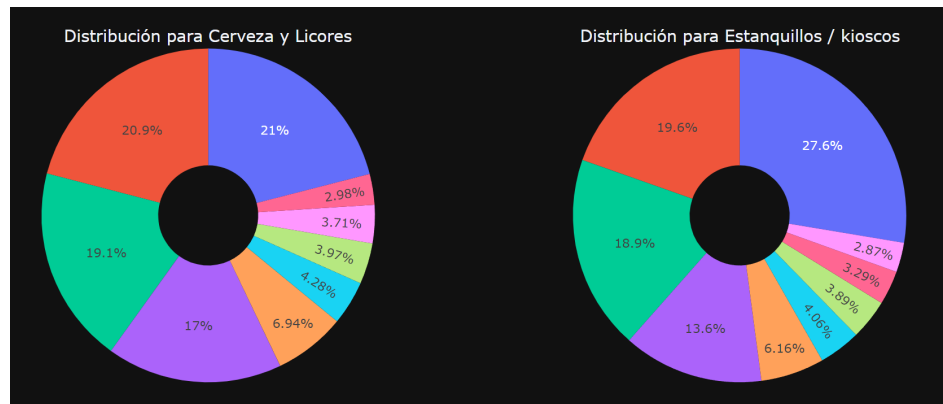


Figura 7: Gráfica de dona de la distribución de subcanal de diferentes clientes

4. Discusión

Los métodos de agrupación y análisis visual han revelado que los meses con mayores ventas corresponden a periodos específicos, posiblemente asociados a factores externos como estacionalidad o promociones. Sin embargo, es necesario eliminar años con datos incompletos como 2019, para evitar sesgos en los resultados, como se señaló al recalcular las ventas desde 2020 en adelante. Los productos de baja recurrencia y los patrones de primeras compras también sugieren oportunidades para refinar las estrategias de ventas y marketing, enfocándose en la retención de clientes y la promoción efectiva de nuevos productos.