

Reporte de Análisis de Datos de Instacart

Introducción

Este informe presenta los hallazgos clave derivados del análisis de los datos de compras en Instacart. Se detalla el proceso de limpieza de datos, la justificación del diseño del star-schema, las consideraciones de negocio que influyeron en las transformaciones de datos y los insights obtenidos a partir del análisis.

1. Proceso de Limpieza de Datos

Principales Acciones Realizadas:

1. Manejo de valores nulos:

- *add_to_cart_order*: Se rellenaron valores nulos con -1 para indicar productos sin una secuencia específica de compra.
- *days_since_prior_order*: Valores nulos fueron reemplazados por 0, asumiendo que es la primera compra del usuario.
- *product_name*: Se reemplazaron valores nulos con "Producto desconocido".

2. Justificación del Star-Schema

Diseño de Tablas:

Se optó por un esquema en estrella para mejorar la eficiencia de las consultas analíticas. Las tablas están estructuradas de la siguiente manera:

- **Tabla de Hechos:** FactOrderProducts
- **Dimensiones:**
 - DimTime
 - DimUser
 - DimProduct

Granularidad: nivel de producto por orden

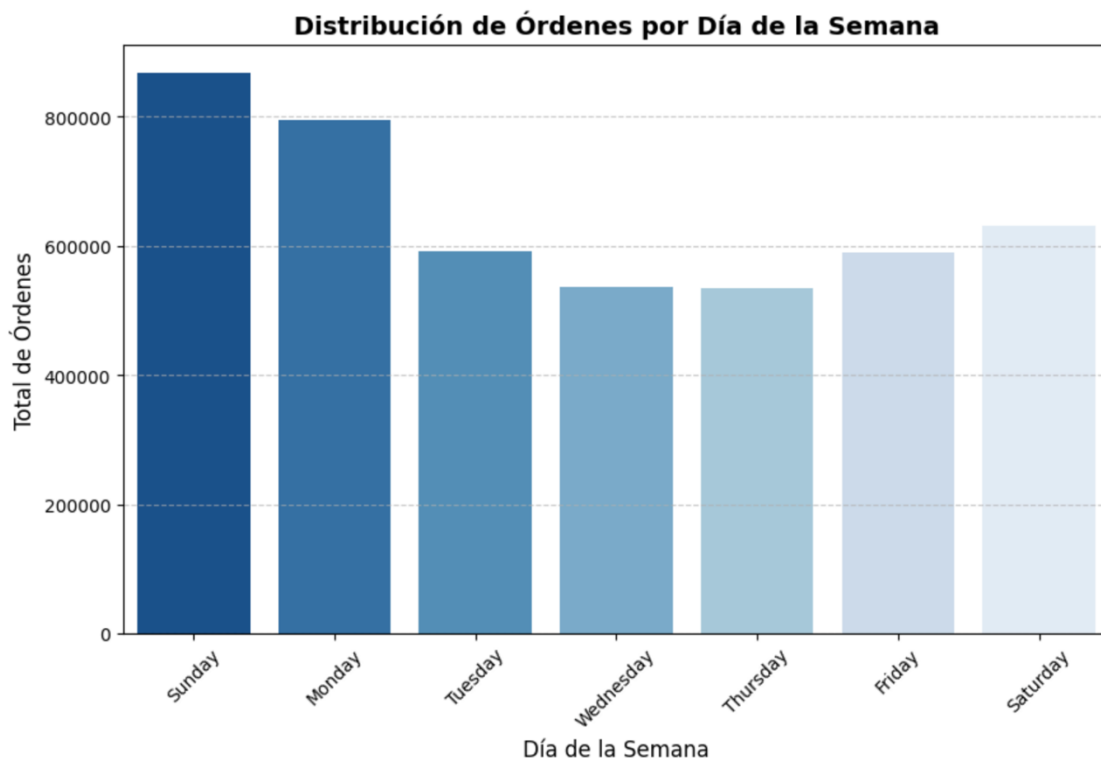
3. Consideraciones de Negocio

Algunas transformaciones de datos fueron influenciadas por factores clave de negocio:

- La limpieza de valores nulos se realizó para evitar errores en los reportes.
- Se agregaron claves primarias y foráneas para optimizar la integridad de los datos.
- La estructura del modelo permite realizar análisis rápidos sobre frecuencia de compras, tendencias de reordenación y patrones de consumo.

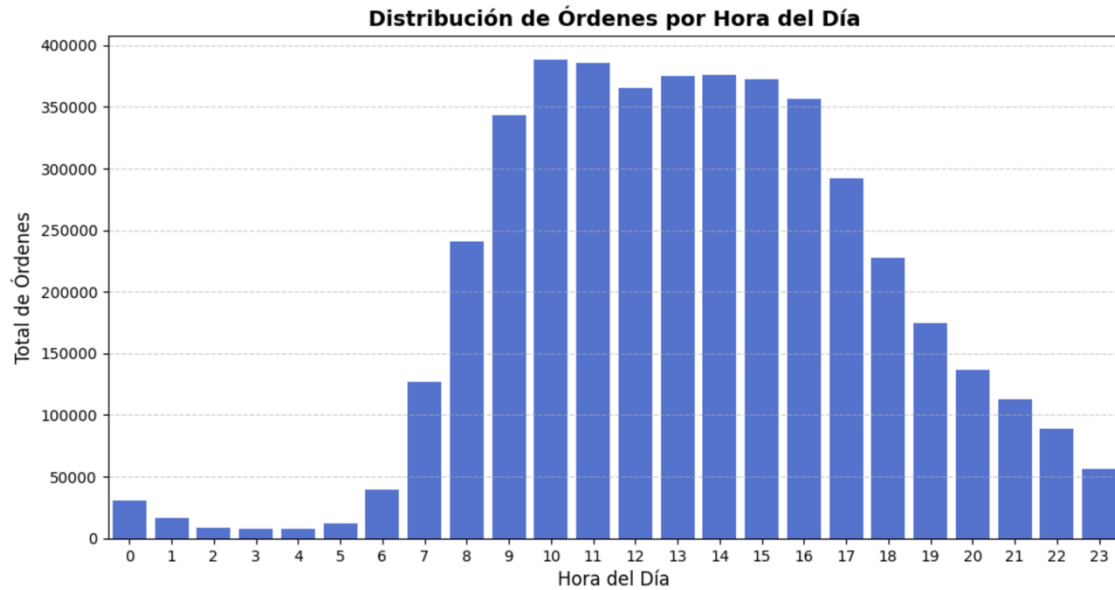
4. Principales Insights del Análisis

1. Comportamiento de compra según día de la semana.



1. Domingo es el día con más órdenes, lo que indica que muchas personas hacen compras en ese día, probablemente para abastecerse para la semana.
2. Menor actividad entre miércoles y jueves, sugiriendo que estos días tienen menos tráfico de compras.
3. Los fines de semana tienen más órdenes en comparación con los días laborales, lo que puede estar relacionado con la disponibilidad de tiempo de los clientes para hacer pedidos.

2. Comportamiento de compra según hora del día



1. Horas Pico (9 AM - 4 PM)

- Se observa un fuerte incremento en el número de órdenes a partir de las 7 AM.
- El punto máximo ocurre entre las 9 AM y 4 PM, con el mayor volumen de pedidos.
- Este horario puede representar un patrón de compra vinculado a las rutinas matutinas, como realizar pedidos antes del almuerzo.

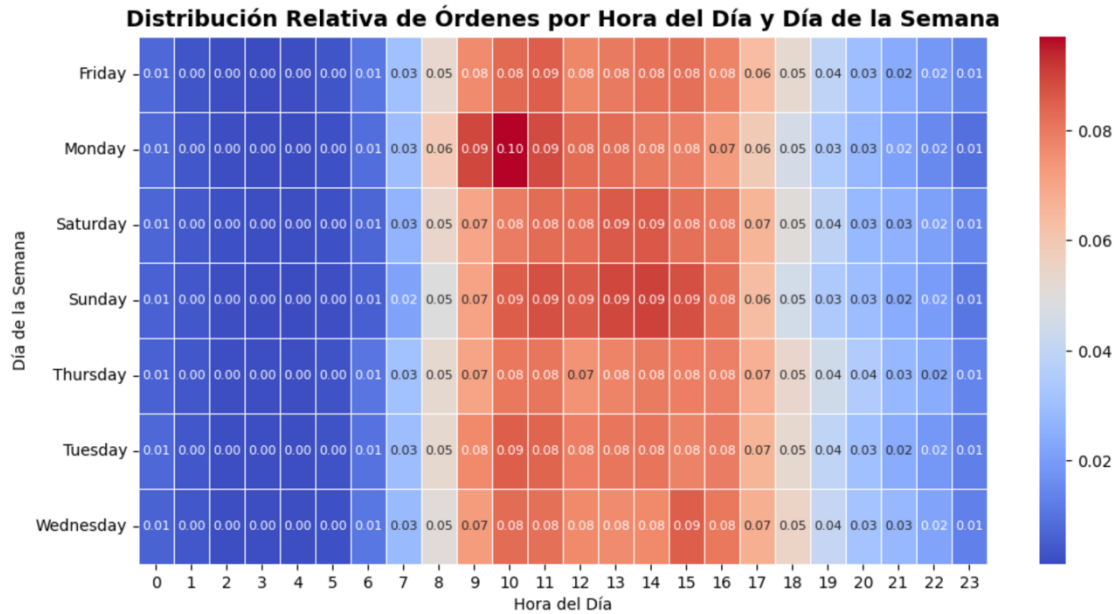
2. Disminución en la Tarde-Noche

- Después de las 4 PM, la cantidad de órdenes empieza a disminuir gradualmente.
- La caída más pronunciada se da a partir de las 6 PM, lo que sugiere que la mayoría de los pedidos ya han sido realizados en el día.

3. Madrugada con Menor Actividad

- Entre la 1 AM y las 6 AM hay una actividad mínima de pedidos.
- Esto indica que la mayoría de los clientes no realizan compras en este horario.

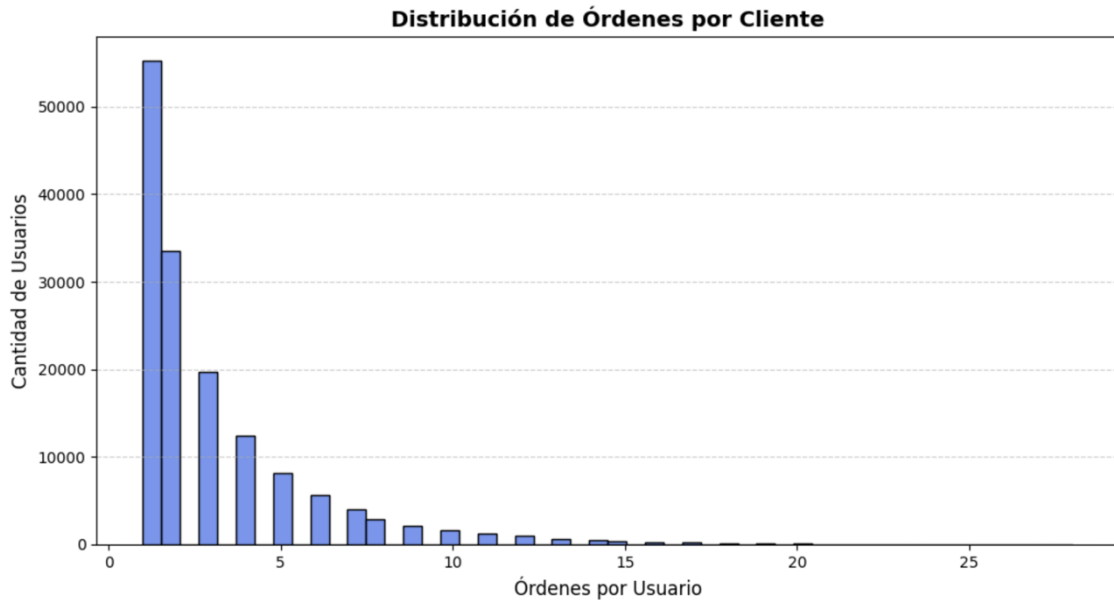
3. Comportamiento según hora del día y día de la semana



1. Mayor concentración de compras en la mañana (8 AM - 3 PM) con un pico alrededor de las 10 AM y 11 AM.
2. El patrón es similar en todos los días, pero los fines de semana muestran una ligera caída en compras a primera hora.
3. En la madrugada (0 AM - 6 AM), la actividad es prácticamente nula.
4. Por la tarde y noche (5 PM - 10 PM), la actividad disminuye, pero se mantiene relativamente estable.

En este insight, se ha normalizado los datos de la columna **"total_orders"** para analizar la distribución relativa de compras a lo largo del día dentro de cada día de la semana.

4. Distribución de las órdenes hechas por los clientes



1. Distribución Asimétrica

- La mayoría de los usuarios realizan entre 1 y 5 órdenes.
- A medida que aumenta el número de órdenes por usuario, la frecuencia de usuarios disminuye drásticamente.

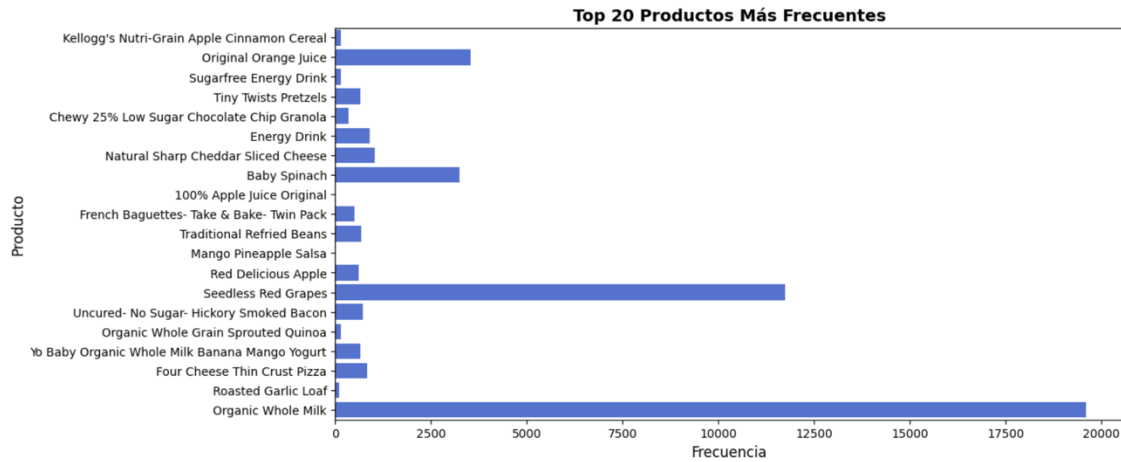
2. Usuarios de una Sola Orden

- Se observa un pico alto en 1 orden, indicando que muchos usuarios solo compraron una vez.
- Este fenómeno podría deberse a clientes ocasionales o que no encontraron valor en repetir la compra.

3. Clientes Leales y Recurrentes

- Aunque en menor cantidad, algunos usuarios han realizado más de 10 órdenes, lo que representa a los clientes fieles de la tienda.

5. Top 20 productos más frecuentes



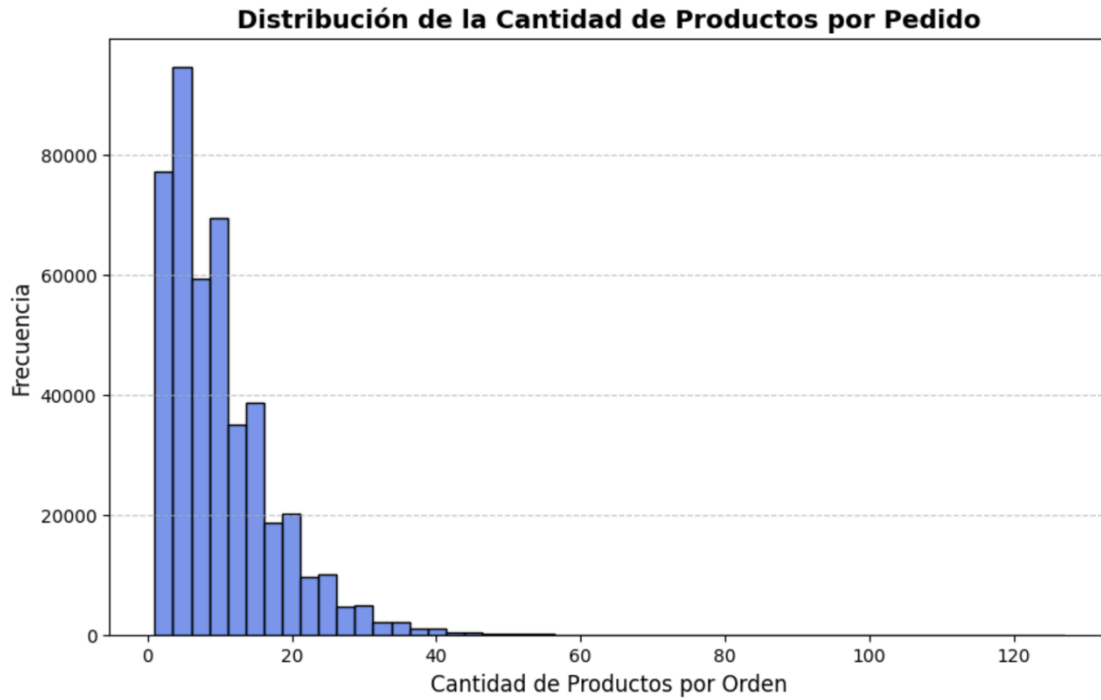
1. Productos más comprados

- El producto "Organic Whole Milk" destaca como el más comprado, con una frecuencia significativamente mayor que otros productos.
- "Seedless Red Grapes" y "Original Orange Juice" también tienen una presencia fuerte en las órdenes.

2. Preferencias de los clientes

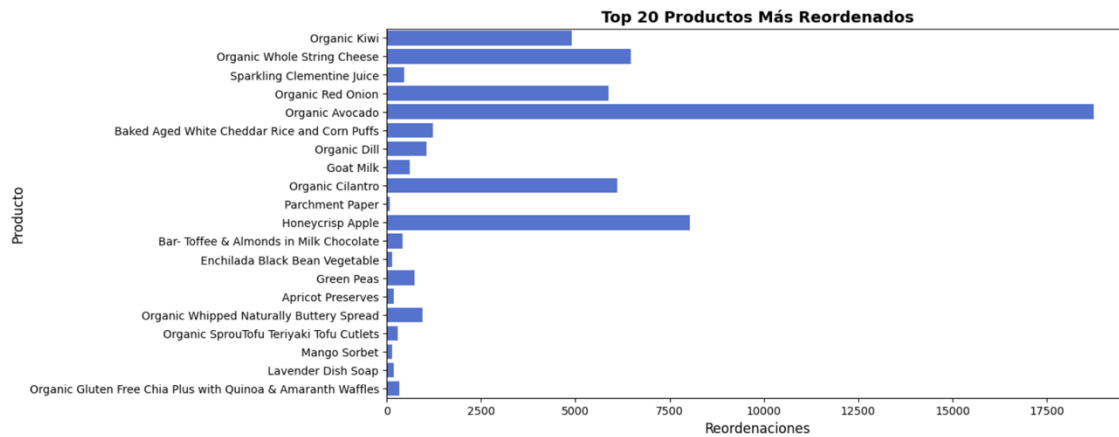
- Se observa una preferencia por productos frescos y orgánicos como leche, jugos naturales, frutas y productos sin azúcar.
- También hay productos de panadería y refrigerados, como French Baguettes y Four Cheese Thin Crust Pizza.

6. ¿Cuántos artículos se compran generalmente en un pedido?



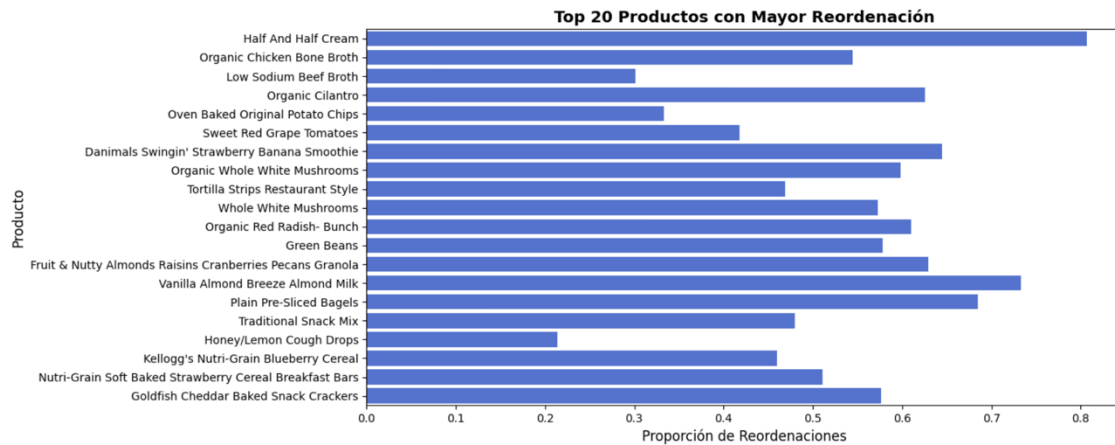
1. La mayoría de los pedidos contienen entre 5 y 15 productos.
 - Se observa una distribución sesgada a la derecha, con la mayoría de las órdenes en un rango de pocos productos.
2. El punto más alto indica que la mayoría de los pedidos tienen menos de 10 productos.
 - Esto sugiere que los clientes suelen realizar compras pequeñas y frecuentes, en lugar de grandes pedidos ocasionales.

7. Top 20 artículos que se vuelven a pedir con más frecuencia



1. Ciertos productos son reordenados con mucha más frecuencia que otros.
 - Los productos en los primeros lugares como el Organic Avocado, Honeycrisp Apple y Organic Whole String Cheese destacan significativamente.
2. Tendencia hacia productos orgánicos y saludables.
 - Muchos de los productos más reordenados son frutas, lácteos y alimentos orgánicos.
 - Esto sugiere que los clientes de esta tienda tienen una preferencia marcada por productos frescos y saludables.
3. Algunos productos tienen una frecuencia muy baja de reordenación.
 - Productos como Enchilada Black Bean Vegetable y Parchment Paper tienen menos reordenaciones.
 - Esto podría indicar que son compras ocasionales en lugar de productos esenciales.

8. Proporción de pedidos que se vuelven a pedir para cada producto



1. Diferencia entre frecuencia de compra y tasa de reordenación

- Aunque algunos productos se compran con mucha frecuencia, no necesariamente tienen la tasa de reordenación más alta.
- En este caso, los productos con más alta proporción de reordenación no siempre son los más comprados.

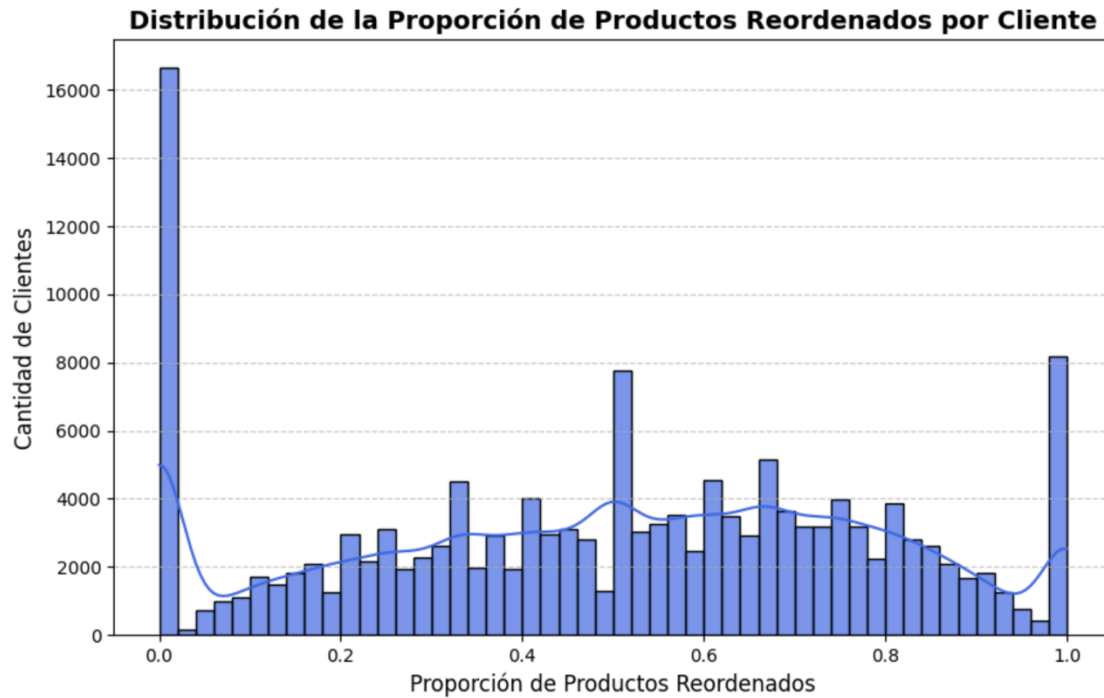
2. Patrón de productos con alta reordenación

- Se observa una tendencia hacia productos fáciles de consumir regularmente, como:
 - Snacks (Goldfish Crackers, Traditional Snack Mix)
 - Lácteos (Half And Half Cream, Vanilla Almond Breeze Almond Milk, Danimals Smoothie)
 - Productos Orgánicos (Organic Cilantro, Organic Whole White Mushrooms)

3. Algunos productos tienen más del 50% de reordenación

- Lo que indica que la mayoría de las personas que los compran, vuelven a hacerlo.

9. Proporción de productos pedidos que se vuelven a pedir para cada cliente



1. Distribución Bimodal

- Se observa un pico alto en 0.0, lo que indica que hay muchos clientes que nunca reordenan productos.
- También hay un pico en 1.0, lo que significa que ciertos clientes reordenan el 100% de sus compras, lo que sugiere patrones de compra altamente repetitivos.

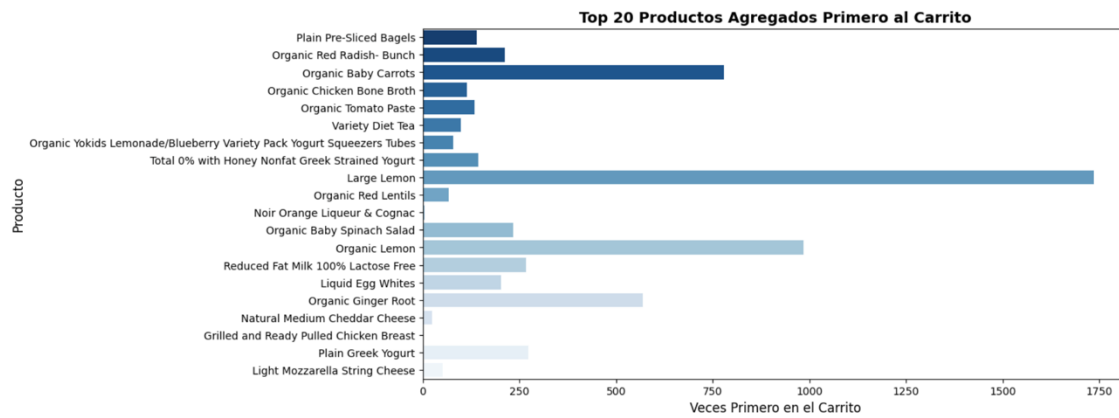
2. Patrones Intermedios

- Entre 0.2 y 0.8 hay una distribución uniforme con algunas concentraciones en ciertos valores.
- Esto indica que la mayoría de los clientes reordenan algunos productos pero también compran nuevos.

3. Segmentación de Clientes

- Compradores exploradores (0.0 - 0.2): Compran productos nuevos en la mayoría de sus órdenes.
- Compradores balanceados (0.3 - 0.7): Tienen una mezcla de productos nuevos y reordenaciones.
- Compradores leales (0.8 - 1.0): Recompran los mismos productos de manera consistente.

10. Top 20 artículos que la gente pone primero en el carrito



1. Large Lemon es el producto más agregado primero al carrito con una diferencia significativa respecto a otros productos.
2. Organic Baby Carrots y Organic Baby Spinach también tienen una alta frecuencia, lo que indica que los consumidores tienden a agregar estos productos como sus primeras elecciones.
3. La distribución sugiere que los productos más agregados en primer lugar suelen ser frescos y básicos, como frutas, vegetales y productos esenciales en la alimentación diaria.
4. Algunos productos tienen una baja frecuencia, lo que indica que no suelen ser la primera elección de los compradores.