

# Dashboard estadístico

## Creadores

- Santarsiero Mateo
- Mendoza Pedro
- Tarricone Thiago

## Objetivo

Visualizar un dashboard estadístico con gráficos para el análisis de ventas de productos en un comercio, se calculan tanto medias, desvíos y correlaciones. Además se debe poder agregar ventas y clientes.

## Entidades Principales

---

- **categories** : Almacena las categorías de los productos (ej. 'Bebida', 'Snack').
- **products** : Contiene la información de los productos individuales, como nombre, precio y stock. Se relaciona con **categories**.
- **zones** : Define las distintas zonas geográficas (ej. 'Centro', 'La Perla').
- **clients** : Guarda los datos de los clientes, incluyendo su edad y la zona a la que pertenecen. Se relaciona con **zones**.
- **payment\_methods** : Lista los métodos de pago aceptados (ej. 'Mercado Pago').
- **sales** : Es la tabla transaccional principal. Registra cada venta, vinculando un cliente (`client_id`), un producto (`product_id`) y un método de pago (`payment_id`). Calcula el `total` automáticamente.

## Estructura de cada tabla

---

### Categories

Columna	Tipo de Dato	Clave/Restricción	Descripción
category_id	int	PK , AUTO_INCREMENT	Identificador único de la categoría.
category_name	varchar(100)	NOT NULL	Nombre de la categoría.

### Products

Columna	Tipo de Dato	Clave/Restricción	Descripción
product_id	int	PK , AUTO_INCREMENT	Identificador único del producto.
name	varchar(150)	NOT NULL	Nombre del producto.
category_id	int	FK (-> categories.category_id)	Enlace a la tabla categories .
unit_price	decimal(10,2)		Precio unitario del producto.
stock	int	DEFAULT 0	Cantidad disponible en inventario.
description	varchar(255)		Descripción del producto.

### Clients

Columna	Tipo de Dato	Clave/Restricción	Descripción
client_id	int	PK , AUTO_INCREMENT	Identificador único del cliente.
first_name	varchar(100)		Nombre del cliente.
last_name	varchar(100)		Apellido del cliente.
email	varchar(150)	UNIQUE	Email único del cliente.
zone_id	int	FK (-> zones.zone_id)	Enlace a la tabla zones .

Columna	Tipo de Dato	Clave/Restricción	Descripción
age	int		Edad del cliente.

---

## Sales

Columna	Tipo de Dato	Clave/Restricción	Descripción
sale_id	int	PK , AUTO_INCREMENT	Identificador único de la venta.
sale_date	datetime	NOT NULL	Fecha y hora de la transacción.
client_id	int	FK (-> clients.client_id)	Enlace al cliente que compró.
product_id	int	FK (-> products.product_id)	Enlace al producto vendido.
quantity	int	CHECK > 0	Cantidad de unidades vendidas.
unit_price	decimal(10,2)		Precio unitario (al momento de la venta).
payment_id	int	FK (-> payment_methods.payment_id)	Enlace al método de pago usado.
total	decimal(10,2)	GENERATED	Columna calculada ( quantity * unit_price ).

---

## Payment\_methods

Columna	Tipo de Dato	Clave/Restricción	Descripción
payment_id	int	PK , AUTO_INCREMENT	Identificador único del método de pago.
payment_name	varchar(50)		Nombre del método.

---

## Zones

Columna	Tipo de Dato	Clave/Restricción	Descripción
zone_id	int	PK , AUTO_INCREMENT	Identificador único de la zona.
zone_name	varchar(100)		Nombre de la zona.

## Principales Consultas SQL

En los casos de consulta de modificación / agregado , los valores son agregados

### 1. Clientes y sus respectivas zonas:

```
SELECT c.client_id, c.first_name, c.last_name, c.email, c.age, z.zone_id, z.zone_name FROM clients AS c
INNER JOIN zones AS z ON c.zone_id = z.zone_id
```

### 2. Productos con categorías

```
SELECT p.product_id, p.name, c.category_name, p.unit_price, p.stock, p.description FROM products AS p
inner JOIN categories AS c ON p.category_id = c.category_id WHERE p.stock > 0
```

### 3. Ventas con cliente y método de pago

```
SELECT s.sale_id, s.sale_date, c.first_name, c.last_name, c.client_id, p.name, s.quantity, s.unit_price,
pay.payment_name, s.total FROM sales s
INNER JOIN clients c ON s.client_id = c.client_id
INNER JOIN products p ON s.product_id = p.product_id
INNER JOIN payment_methods pay ON s.payment_id = pay.payment_id
```

### 4. Insertado de venta

```
INSERT INTO sales (client_id, product_id, quantity, unit_price, payment_id, sale_date) VALUES (?, ?, ?, ?, ?, ?)
```

## 5. Insertado de cliente

```
INSERT INTO clients (first_name, last_name, email, age, zone_id) VALUES (?, ?, ?, ?, ?)
```

## Variables analizadas y estadísticos realizados

Tener en cuenta que los cálculos son realizados en el back-end (PHP) y, en este documento, serán representados en el formato de JSON (**JavaScript Object Notation**). Por otro lado, los datos utilizados son los de la base de datos MySQL adjunta a este documento, en caso de requerir un input.

### Promedios

Las siguientes variables son analizadas en los promedios: `Día`, `Nombre de producto`, `Nombre de Cliente`, todas estas son utilizadas para realizar medias teniendo como denominador su respectivo total y clasificarlas.

El principal uso de este tipo de estadísticos es el de poder resumir las numerosas ventas en distintos factores, ¿en qué días hubo más ventas? ¿Quién es nuestro mayor cliente o cual es nuestro producto más exitoso?

Ese tipo de preguntas se pueden, es un aproximado, responder mediante los estadísticos antes nombrados.

Resultado:

```
{
  "status": "ok",
  "data": {
    "Average by day": {
      "18 11 25": 1709.090909090909,
      "14 11 25": 17100,
      "11 08 25": 57000
    },
    "Average by product": {
      "Papas Lays": 3090.909090909091,
      "Coca cola": 33250
    },
    "Average by client": {
      "Mateo Santarsiero": 19562.5,
      "Testing Trigger": 1500
    }
  }
}
```

### Desviación

Por otro lado, gracias a la `desviación` calculada con las ventas se puede responder, entre otra consultas, una muy fundamental para un comercio: ¿Qué tan importante o impactante son los altibajos en las ventas de mi negocio?

Puesto que esta dispersión estadística se calcula en un periodo específico (`origin_date` y `end_date`), los resultados pueden analizarse en adición a su contexto para concluir, por ejemplo, en nuevas estrategias para reforzar el negocio en las fechas que posean mayor desviación.

Input:

```
{
  "origin_date": "2025-08-11",
  "end_date": "2025-11-19"
}
```

Resultado:

```
{
  "status": "success",
  "data": 24643.563826996247
}
```

## Correlaciones de Pearson

Finalmente, el análisis se profundiza mediante el cálculo de Correlaciones . A diferencia de los promedios o las desviaciones que miran una variable aislada, este cálculo busca patrones ocultos conectando dos variables simultáneamente: Precio Unitario vs Cantidad , Cantidad vs Día de la Semana , y Total vs Método de Pago .

El objetivo de procesar estos pares de datos es responder a interrogantes sobre el comportamiento del consumidor. Al cruzar el Precio Unitario con la Cantidad , por ejemplo, intentamos resolver: **¿El costo de un producto incentiva o disuade a nuestros clientes de llevar más unidades?** Si la correlación es negativa, confirmaría que a precios más altos, el volumen de compra por ticket disminuye.

Del mismo modo, al correlacionar Cantidad con Día de la semana , la administración puede visualizar tendencias temporales: **¿Existen días específicos donde los clientes tienden a realizar compras masivas?** Esto permite ajustar la logística y el stock preventivo para esos días críticos.

Por último, el vínculo estadístico entre el Total monetario y el ID de Pago ofrece una perspectiva financiera fundamental: **¿Los montos elevados de venta están atados a un medio de pago específico?** Entender esta relación permite al negocio optimizar sus pasarelas de pago o lanzar promociones específicas para el método que genere mayores ingresos, basando estas decisiones no en la intuición, sino en la fuerza de la relación matemática calculada entre las ventas históricas.

```
{  
  "status": "success",  
  "sales_count": 15,  
  "correlations": {  
    "price_vs_quantity": 0.5403834774873848,  
    "quantity_vs_day_of_week": -0.15871790922363302,  
    "total_vs_payment_method": -0.33037643057802313  
  }  
}
```

## Análisis de gráficos

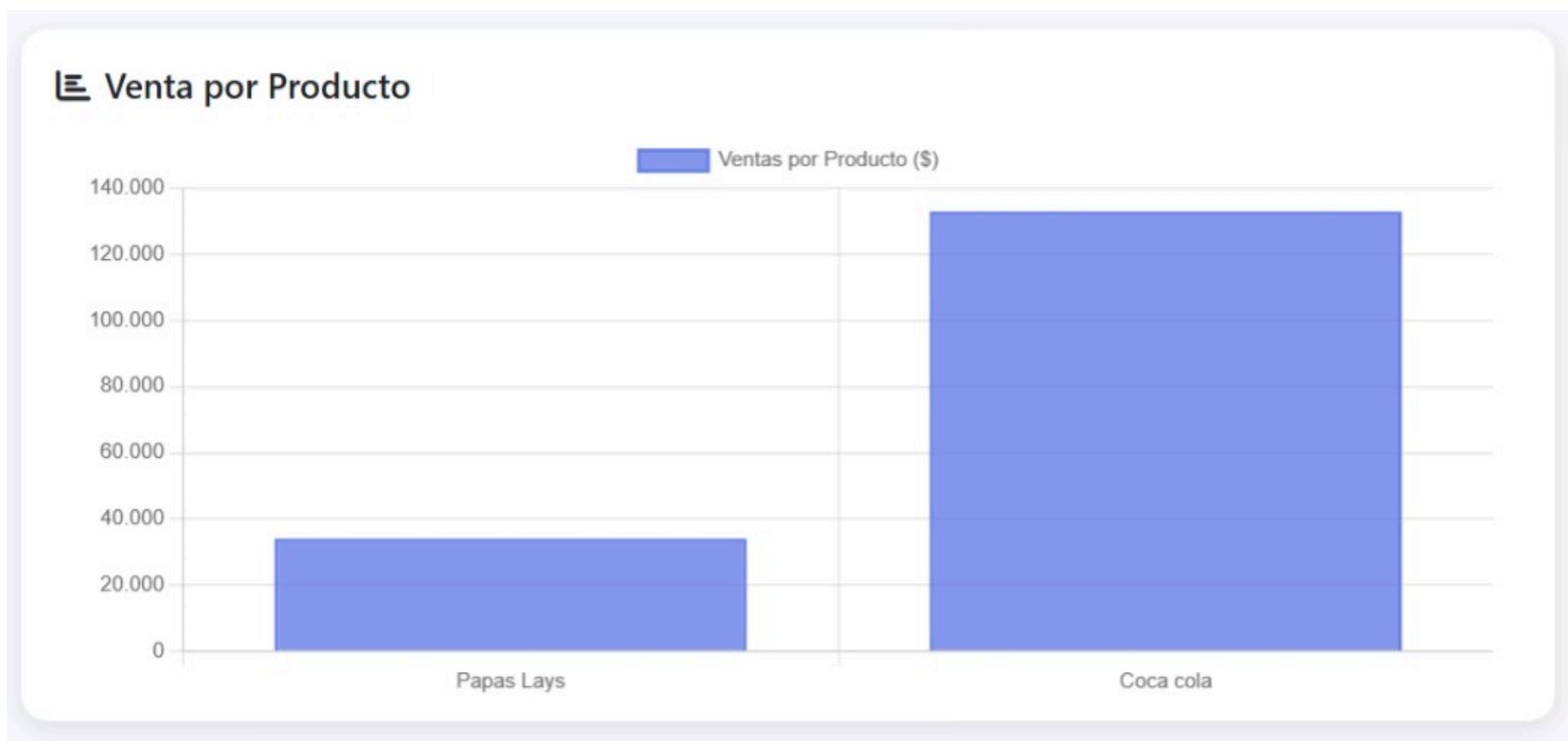
### 1. Resumen



El tablero refleja un negocio con un volumen de transacciones bajo pero de alto valor promedio, impulsado por ventas mayoristas o en lote.

- Volumen de Negocio:** Se registraron ventas totales por **\$167,000.00** a través de 15 transacciones.
- Ticket Promedio:** El valor medio por venta es sorprendentemente alto (**\$11,133.33**). Esto se explica al revisar la base de datos: existe una venta masiva (ID 1) de 25 unidades de Coca-Cola que distorsiona el promedio hacia arriba.
- Producto Estrella:** "Coca cola" es indiscutiblemente el producto más vendido.

### 2. Análisis de Ventas y Preferencias



#### Venta por Producto (Ingresos)

El gráfico de barras muestra una disparidad masiva entre los ingresos generados por "Coca cola" frente a "Papas Lays".

- **Datos de Respaldo:** En la base de datos, la "Coca cola" tiene un precio unitario de 3800.00, mientras que las "Papas Lays" cuestan 1500.00.
- **Anomalía Detectada:** Aunque las papas son más baratas, la base de datos muestra un registro (Venta ID 2) donde se vendieron 5 unidades de "Papas Lays" a precio de "Coca cola" (3800.00) en lugar de su precio de lista (1500.00). Esto infló artificialmente la barra de ingresos de Papas Lays en el gráfico.

## Métodos de Pago



El gráfico de dona indica una preferencia clara por **Mercado Pago** (color azul) sobre Cuenta DNI (color rosa).

- **Contexto:** Mercado Pago (ID 1) se utilizó en las primeras ventas grandes (incluida la venta de \$95,000), mientras que Cuenta DNI (ID 2) aparece recién en las transacciones finales (ID 9 a 14) y solo para montos pequeños (\$1,500). Esto sugiere que los clientes que gastan más prefieren Mercado Pago.

## Temporalidad (Promedio por Día)



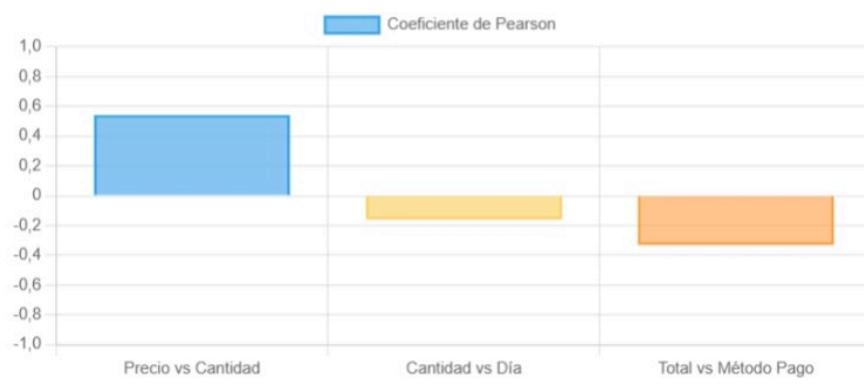
El gráfico "Promedio de Ventas por día" muestra un pico gigante el día **11/08/25**.

- **Causa:** Esto coincide con las dos primeras inserciones en la base de datos (Ventas ID 1 y 2) realizadas el 2025-08-11, que suman \$114,000 en un solo día. Los días subsiguientes (14 y 18 de noviembre) tienen muchas más transacciones, pero de montos mucho menores (unitarios), lo que baja drásticamente el promedio diario visualizado.

## 3. Interpretación de las Correlaciones de Pearson

## Correlaciones de Pearson

Coefficientes de Correlación de Pearson



### Correlación Precio vs Cantidad

$r = 0,540$

Correlación moderada positiva. A mayor precio mayor cantidad la otra variable.

### Correlación Cantidad vs Día

$r = -0,159$

El día de la semana no influye significativamente en la otra variable.

### Correlación Total vs Pago

$r = -0,330$

Correlación débil negativa. El método de pago se relaciona con la otra variable.

## 1. Precio vs. Cantidad ( $r = 0,540$ )

- Interpretación:** Existe una **correlación positiva moderada**. Esto es contra intuitivo en economía (normalmente, a mayor precio, menor demanda).
- El producto más caro (Coca cola, \$3800) fue el que se vendió en la mayor cantidad física (25 unidades en una sola venta).
- Conclusión de Negocio:** Tus clientes no son sensibles al precio alto; de hecho, el producto premium es el que compran en mayor volumen.

## 2. Cantidad vs. Día ( $r = -0,159$ )

- Interpretación:** Correlación negativa muy débil, prácticamente inexistente.
- Conclusión de Negocio:** No existe un patrón claro que indique que en un día específico de la semana la gente compre más unidades. Las compras masivas y las unitarias ocurren sin depender del día de la semana.

## 3. Total vs. Pago ( $r = -0,330$ )

- Interpretación:** Correlación negativa débil.
- Análisis de la variable:**
  - Pago ID 1 = Mercado Pago.
  - Pago ID 2 = Cuenta DNI.
  - Total = Dinero gastado.
- Explicación:** Dado que el ID de Mercado Pago es "1" (número bajo) y está asociado a los Totales más altos (\$95,000), y Cuenta DNI es "2" (número alto) y está asociado a los Totales más bajos (\$1,500), la matemática detecta que "a medida que sube el número de ID del pago, baja el total de la venta".
- Conclusión de Negocio:** Existe una tendencia ligera a que las compras grandes se paguen con Mercado Pago y las compras chicas ("kiosco") se paguen con Cuenta DNI.

## Tareas desarrolladas

Tarea	Descripción	Miembro	Tiempo
Planeación	Planteamiento de la idea del proyecto y los requerimientos a cumplir	Mateo, Pedro, Thiago	15 minutos
Base de datos	Diseño y creación de la base de datos MySQL	Thiago	25 minutos
Backend	Utilizado PHP, desarrollo de tanto la conexión a la base de datos y los endpoints planteados en los requerimientos	Mateo	8 horas
Frontend	Usado HTML, Javascript y CSS, desarrollo del dashboard, incluyendo gráficos y tablas consumiendo las API's del backend	Pedro	5 horas
Corrección de errores	Fixes a bugs tanto en el back-end como en el front-end. Siendo visuales y lógicos.	Mateo, Pedro	21 horas
Documentación	Desarrollo de este documento	Mateo	2 horas