

## 1. Carga y exploración básica

- Carga el dataset en un DataFrame.
- Muestra los primeros 5 registros.
- Describe brevemente las columnas y sus tipos.
- ¿Hay valores nulos? ¿Cómo los manejarías?

	fecha	id_proveedor	proveedor	prod_id	producto	proveedor_unificado	valor_ventas	cantidad_tx
0	2025-05-26	3	TIGO	222	9.5GB+MIN.ILIM.+20D	TIGO	5670000	252
1	2025-05-27	3	TIGO	109	10 MIN LDI + 2 DIAS	TIGO	151200	28
2	2025-05-28	3	TIGO	575	5 GB+MIN.ILIM+15D	TIGO	4284000	306
3	2025-05-29	3	TIGO	101	5 MIN LDI + 1DIA	TIGO	58900	19
4	2025-05-30	3	TIGO	485	13 GB+MINILI+RDS+30D	TIGO	15307500	471

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 84 entries, 0 to 83
Data columns (total 8 columns):
#   Column                Non-Null Count  Dtype
---  -
0   fecha                 84 non-null    object
1   id_proveedor          84 non-null    int64
2   proveedor             84 non-null    object
3   prod_id               84 non-null    int64
4   producto              84 non-null    object
5   proveedor_unificado   84 non-null    object
6   valor_ventas          84 non-null    int64
7   cantidad_tx           84 non-null    int64
dtypes: int64(4), object(4)
memory usage: 5.4+ KB
```

```
fecha                 0
id_proveedor          0
proveedor             0
prod_id               0
producto              0
proveedor_unificado   0
valor_ventas          0
cantidad_tx           0
dtype: int64
```

Si tuviera datos nulos en este caso específico, eliminaría el registro, ya que esto podría alterar los análisis y los cálculos a la hora de sacar porcentajes, promedios, etc.

```
ventas = ventas.dropna()
```

## 2. Limpieza de datos

- Corrige los tipos de datos si es necesario (por ejemplo, fechas).
- Elimina columnas irrelevantes si aplica.
- Crea una columna nueva calculada, por ejemplo: Ganancia = Ingreso - Costo.

```
<class 'pandas.core.series.Series'>
RangeIndex: 84 entries, 0 to 83
Series name: fecha
Non-Null Count  Dtype
-----
84 non-null    datetime64[ns]
dtypes: datetime64[ns](1)
memory usage: 804.0 bytes
```

Todas las columnas están relacionadas así que no serían necesario eliminar ninguna.

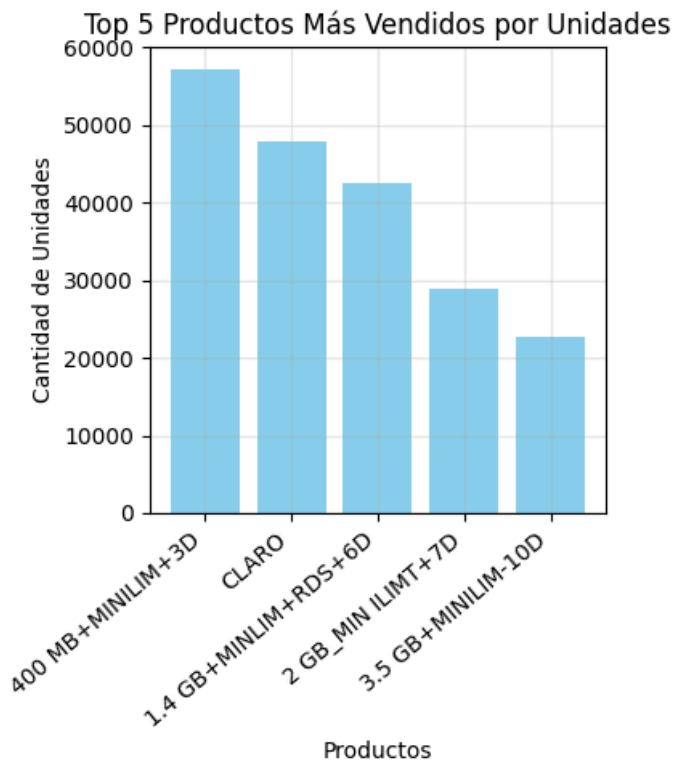
Si fuera necesario eliminar una columna: `ventas = ventas.drop(columns=['Columna_a_eliminar'])`

	fecha	id_proveedor	proveedor	prod_id	producto	proveedor_unificado	valor_ventas	cantidad_tx	valor_unitar
0	2025-05-26	3	TIGO	222	9.5GB+MIN.I.LIM.+20D	TIGO	5670000.0	252	22500

### 3. Análisis exploratorio

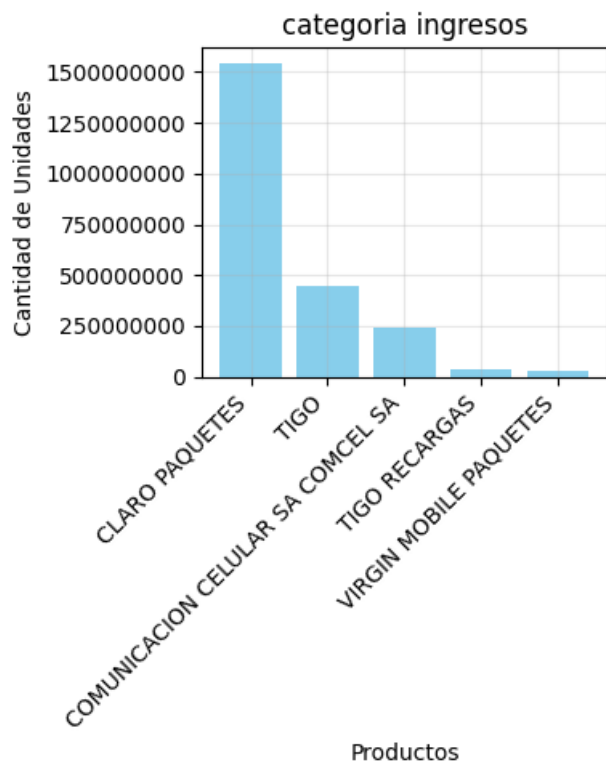
- ¿Cuáles son los 5 productos más vendidos por unidades?

	producto	cantidad_tx
0	400 MB+MINILIM+3D	57218
1	CLARO	47968
2	1.4 GB+MINILIM+RDS+6D	42523
3	2 GB_MIN ILIMT+7D	28985
4	3.5 GB+MINILIM-10D	22623



- ¿Cuál es la categoría que más ingresos genera?

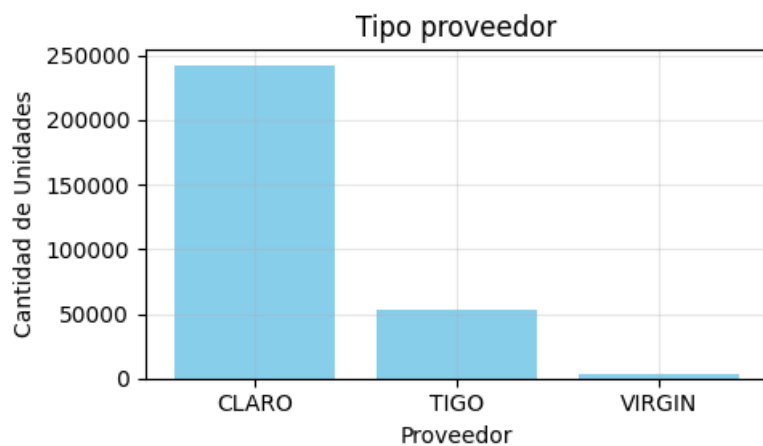
	proveedor	valor_ventas
0	CLARO PAQUETES	1.542106e+09



La categoria que mas ingresos genera es CLARO PAQUETES con \$1,542,106,250 en ventas

- ¿Hay diferencias significativas de ventas por región o canal?

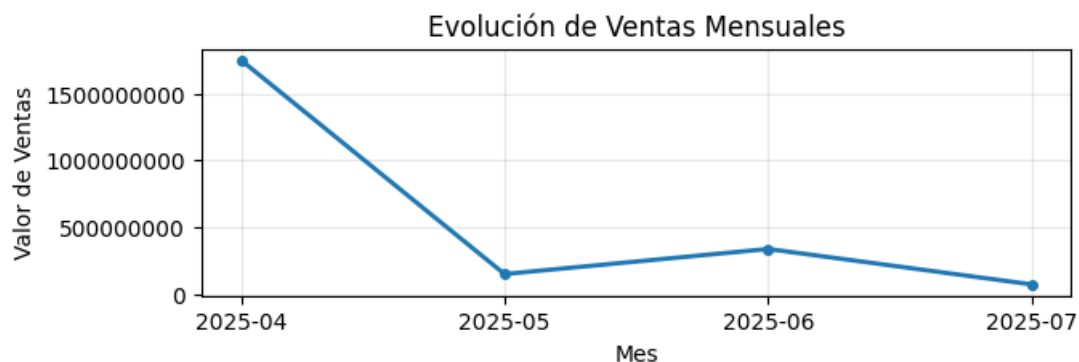
	proveedor_unificado	cantidad_tx	porcentaje_cantidad_tx
0	CLARO	242152	81.01
1	TIGO	53631	17.94
2	VIRGIN	3135	1.05



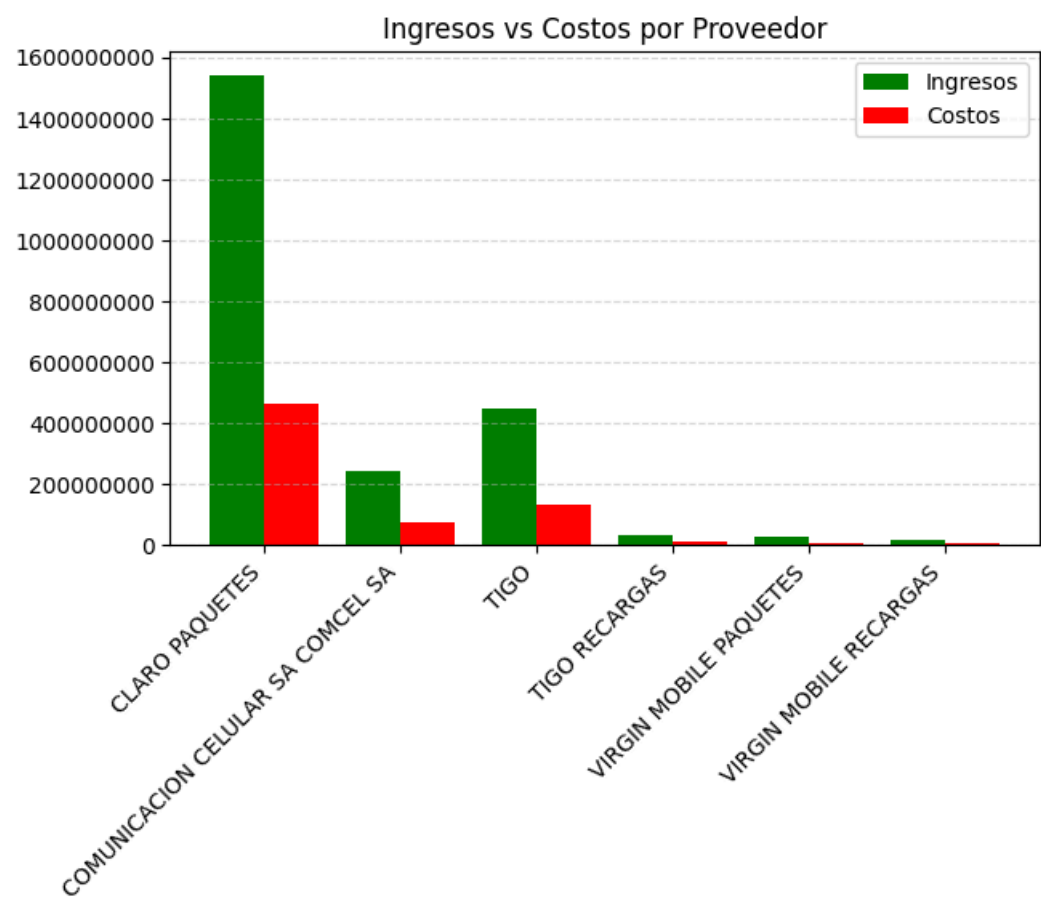
CLARO domina el mercado con (81%), mientras que TIGO (18%) y VIRGIN (1%) tienen ventas considerablemente menores.

## 4. Visualización

- Crea una visualización que muestre la evolución de las ventas mensuales.

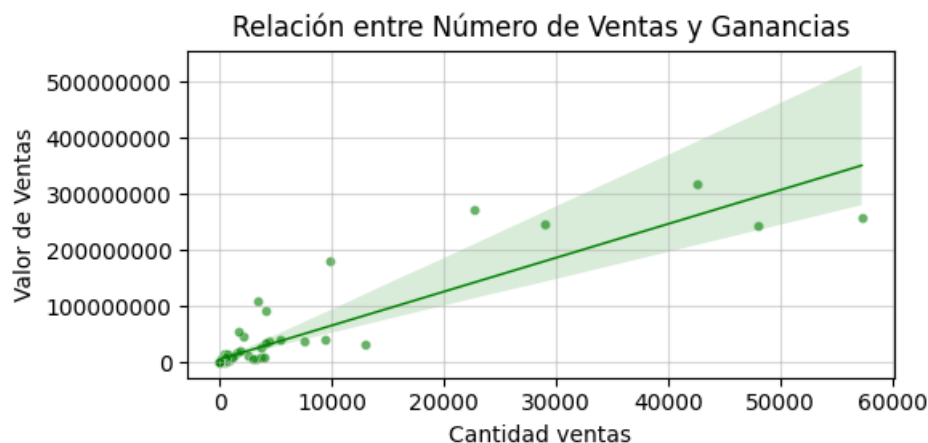


- Muestra un gráfico comparativo entre ingresos y costos por categoría.



Debido a que no se tiene la columna de costos tomé un supuesto del 30% por costo de venta para calcularla.

- Usa un gráfico de dispersión (scatter) para mostrar la relación entre el número de ventas y las ganancias.



## 5. Insight y comunicación

- En máximo 5 líneas, escribe dos hallazgos interesantes del análisis.
- Propón una acción que el negocio podría tomar basada en esos hallazgos.

### Hallazgo 1:

CLARO domina completamente el mercado de telecomunicaciones con 242,152 unidades vendidas, representando más del 80% del total de transacciones, mientras que TIGO (53,631) y VIRGIN (3,135) tienen participaciones significativamente menores.

### Hallazgo 2:

Existe una correlación positiva fuerte entre la cantidad de ventas y el valor de ventas, donde el producto "400 MB+MINILIM+3D" lidera tanto en unidades vendidas (57,218) como en generación de ingresos, sugiriendo que los productos de mayor volumen también generan mayores ingresos.

### Acción recomendada:

El negocio debería enfocar sus esfuerzos comerciales para productos similares al "400 MB+MINILIM+3D". Teniendo en cuenta que el volumen es el factor principal del negocio, tomar un enfoque en aumentar la frecuencia de recompra y recurrencia de clientes. Debido a la fuerte caída en ventas después del primer mes se debería diseñar estrategias para convertir ventas puntuales en ventas recurrentes.