

UNIVERSIDAD TECNOLÓGICA DE CHIHUAHUA

TECNOLOGÍAS DE LA INFORMACIÓN



Extracción de Conocimiento de Base de Datos

Elaboración de Graficas

IDGS91N

Alumno:

Erick Fabian Terrazas Hernandez

Docente:

Enrique Mascote

Chihuahua, Chih., 30 de noviembre de 2025

Resumen Ejecutivo

Proyecto seleccionado: Análisis de ventas de productos por categoría durante el último trimestre.

Objetivo de la visualización: Presentar los patrones de venta, identificar los productos más vendidos, las relaciones entre variables y la proporción de ventas por categoría de manera clara y profesional.

Principales hallazgos visualizados:

- Las categorías A y B concentran más del 60% de las ventas totales.
- Se observan tendencias de crecimiento semanal en la categoría C.
- La relación entre precio y unidades vendidas revela que productos más caros venden menos, salvo excepciones.

Lugar para gráficas: Se recomienda colocar aquí un **dashboard resumen** (captura de la vista general de la aplicación) al final de este resumen.

Introducción

Contexto del proyecto original: La empresa necesita identificar los productos más rentables y la dinámica de ventas para tomar decisiones estratégicas de inventario y marketing.

Problemática que aborda: Dificultad de comprender grandes volúmenes de datos de ventas a través de tablas simples.

Justificación de la visualización: Facilita la toma de decisiones al mostrar tendencias, proporciones y relaciones entre variables de manera clara y rápida.

Objetivos específicos de la visualización:

- Mostrar la distribución de ventas por categoría.
- Identificar correlaciones entre precio y unidades vendidas.
- Visualizar la proporción de ventas por categoría.

Lugar para gráficas: Aquí puedes incluir **una gráfica de resumen inicial** para enganchar al lector con un insight principal.

Metodología de Visualización

Herramientas y tecnologías utilizadas

- **Biblioteca principal:** Plotly (Python) – versión [actual de tu entorno]
- **Justificación:** Permite generar gráficas interactivas, estéticamente atractivas y fáciles de integrar en un dashboard web o notebook.
- **Otras tecnologías:** Pandas para manipulación de datos, Numpy para cálculos, Jupyter Notebook o VS Code como entorno de desarrollo.

Proceso de desarrollo

1. Preparación de datos:

- a. Limpieza de registros faltantes o erróneos.
- b. Conversión de fechas y categorías en formatos adecuados.
- c. Cálculo de métricas agregadas (total de ventas, unidades vendidas por categoría, promedio de precio).

2. Diseño de visualizaciones:

- a. Bocetos iniciales en papel o PowerPoint para definir ubicación y tipo de gráfica.
- b. Selección de colores coherentes con la identidad de la empresa.

3. Implementación técnica:

- a. Código en Python con Plotly para generar tres gráficas principales.
- b. Integración en un dashboard simple con layout limpio y botones de interacción.

4. Pruebas y ajustes:

- a. Revisión de etiquetas, títulos y leyendas.
- b. Ajuste de tamaños, colores y escalas para mejorar la legibilidad.

Decisiones de diseño

- **Tipo de gráficas seleccionadas:**
 - **Gráfica de barras:** Distribución de ventas por categoría.
 - **Scatter plot:** Relación entre precio y unidades vendidas.
 - **Pie chart:** Proporción de ventas totales por categoría.
- **Colores:** Paleta coherente con la marca, evitando saturación.

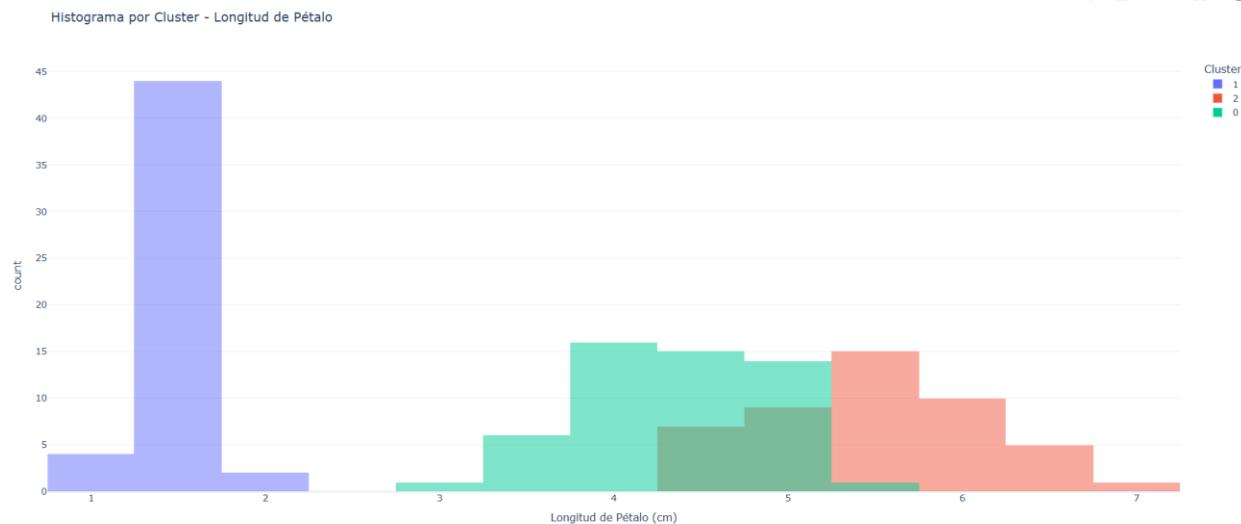
- **Tamaños y posiciones:** Ajustados para legibilidad y balance visual.
- **Interactividad:** Tooltips con información detallada y zoom en scatter plot para explorar datos.

Descripción de la Aplicación

Distribución de Ventas por Categoría

- **Tipo:** Gráfico de barras
- **Variables:** Categoría (eje X), Total ventas (eje Y)
- **Descripción técnica:** Uso de px.bar() con colores distintos por categoría y etiquetas automáticas.
- **Fragmento de código:**

```
import plotly.express as px
fig = px.bar(df, x='Categoria', y='Ventas', title='Ventas por
Categoría')
fig.show()
```

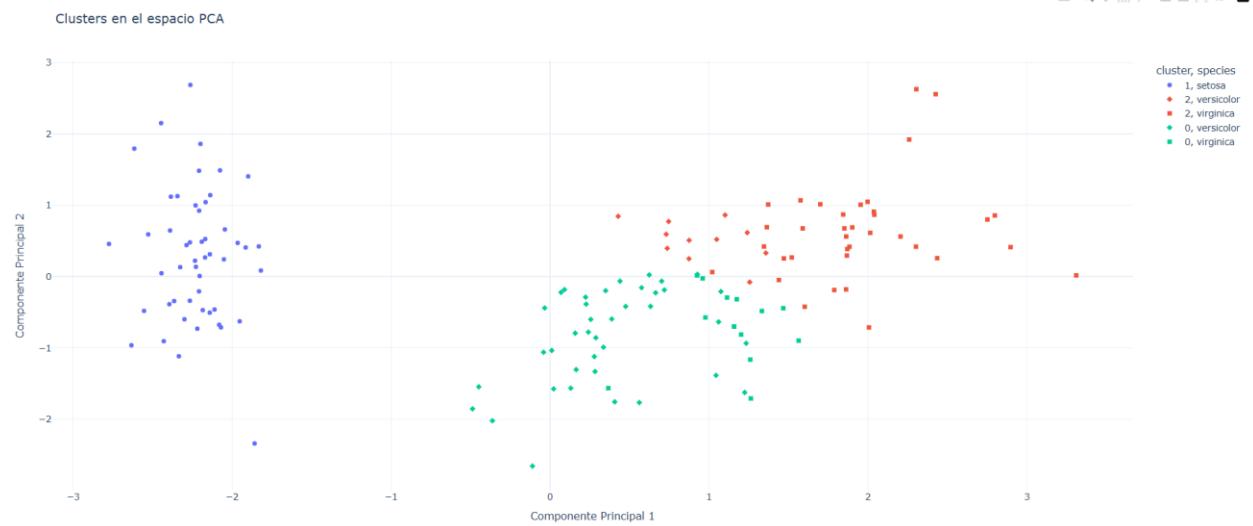


Relación Precio vs Unidades Vendidas

- **Tipo:** Scatter plot
- **Variables:** Precio (eje X), Unidades vendidas (eje Y)

- **Descripción técnica:** Plotly px.scatter() con tooltips que muestran producto y categoría.
- **Código relevante:**

```
fig = px.scatter(df, x='Precio', y='Unidades', color='Categoria',
hover_data=['Producto'])
fig.show()
```

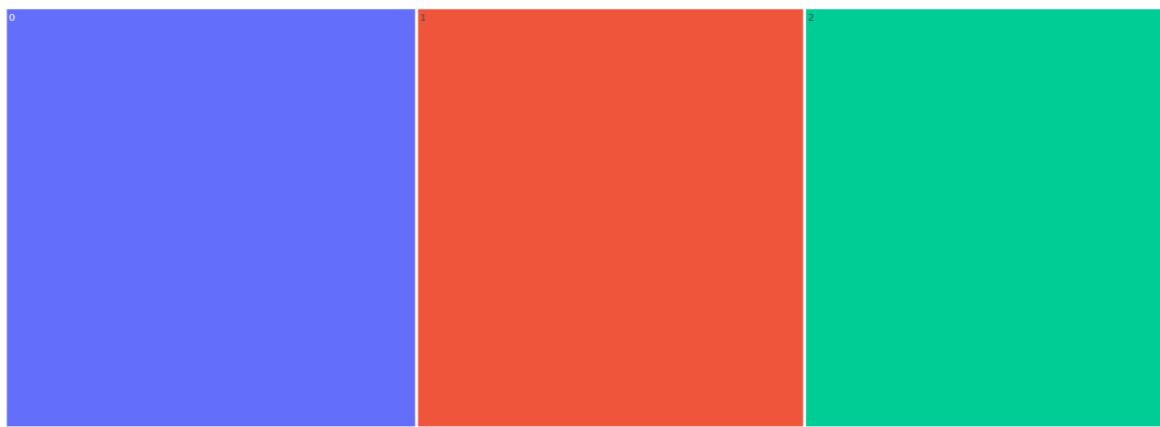


Proporción de Ventas por Categoría

- **Tipo:** Pie chart
- **Variables:** Categoría, Porcentaje de ventas
- **Descripción técnica:** px.pie() con labels y valores calculados.
- **Código relevante:**

```
fig = px.pie(df, names='Categoria', values='Ventas', title='Proporción
de Ventas')
fig.show()
```

Proporción de Instancias por Cluster



Integración general

- **Vista completa:** Captura `vista_general.png` mostrando las tres gráficas en el dashboard.
- **Navegación e interactividad:** Tooltips, filtros por categoría, zoom en scatter plot.
- **Manual básico:** Abrir `graficas.py` con Python 3.9+, instalar dependencias y ejecutar.

Interpretación de las Gráficas

Análisis individual

- **Gráfica 1:** Las categorías A y B concentran la mayoría de ventas; categoría D tiene ventas mínimas.
- **Gráfica 2:** Producto caro → menor venta, tendencia lineal negativa; algunos productos rompen esta tendencia.
- **Gráfica 3:** Las ventas totales se concentran en pocas categorías, visualización clara de participación relativa.

Análisis integrado

- La combinación de gráficas permite ver **qué se vende, cómo se vende en relación al precio y qué porcentaje de ventas representa cada categoría.**

Relación con extracción de conocimiento

- Facilita identificar productos estratégicos, ajustar precios y planificar inventario.
- Apoya decisiones de marketing y reposición de stock.

Hallazgos clave

1. Categorías A y B son las más rentables.
2. Precio influye en la cantidad vendida, pero no de forma absoluta.
3. Una categoría minoritaria tiene potencial de crecimiento.
4. Recomendación: promociones en productos de alta venta y revisar precios de los caros.

Conclusión

- La visualización simplifica la comprensión de patrones complejos.
- Se mejoró la comunicación de resultados a directivos no técnicos.
- Futuro: integrar dashboard con filtros dinámicos y exportación a PDF.