

Universidad Nacional Rosario Castellanos

Campus: Casco de Santo Tomás

Ciencia de Datos para Negocios

Análisis de Nuevos Tipos de Fraudes Bancarios

UCA: Herramientas para Grandes Volúmenes de Datos

Grupo: 502

Alumno: Daniel Cureño Martínez

Fecha: 08 Octubre 2025

Objetivo

Simular y analizar la ocurrencia de nuevos tipos de fraudes bancarios mediante el uso de **Python y NumPy**, generando datos aleatorios que representen la frecuencia de detección de cada tipo de fraude.

Se busca:

- Calcular la **media** y **desviación estándar** de ocurrencias.
- Estimar el **total de pérdidas** considerando una pérdida promedio de \$2000 por evento.
- Ordenar los resultados de forma ascendente.
- Visualizar los resultados en una gráfica con líneas de referencia estadísticas.

Importación de librerías

En esta sección se importan las librerías requeridas: **NumPy** para la generación y análisis de datos, y **Matplotlib** para la visualización gráfica.

```
In [1]: import numpy as np
import matplotlib.pyplot as plt
```

Configuración de variables y simulación de datos

Se define un conjunto de 8 nuevos tipos de fraude, incluyendo la modalidad “**monta deudas**”, la cual representa una de las estafas digitales emergentes más comunes en Latinoamérica.

Cada tipo de fraude tendrá una ocurrencia aleatoria entre 0 y 10,000, representando casos simulados detectados por sistemas de monitoreo bancario.

```
In [2]: np.random.seed(42)
perdida_promedio = 2000
max_ocurrencias = 10000
n = 8

tipos_de_fraude = [
    "Fraude de monta deudas",
    "Fraude por comercio electrónico",
    "Fraude por compra remota",
    "Fraude de devolución",
    "Fraude por aplicaciones falsas",
    "Fraude por intercambio de SIM",
    "Fraude por compromiso de cuenta",
    "Fraude por facturación duplicada"
]

ocurrencias = np.random.randint(0, max_ocurrencias + 1, size=n)
ocurrencias
```

```
Out[2]: array([7270, 860, 5390, 5191, 5734, 6265, 466, 4426])
```

Cálculo estadístico

A partir de los datos simulados, se calcula la **media** y la **desviación estándar** de las ocurrencias.

También se ordenan los resultados de manera ascendente para facilitar la interpretación visual.

```
In [3]: media = np.mean(ocurrencias)
desviacion_estandar = np.std(ocurrencias)

orden_idx = np.argsort(ocurrencias)
ocurrencias_ordenadas = ocurrencias[orden_idx]
tipos_ordenados = [tipos_de_fraude[i] for i in orden_idx]

print("Ocurrencias (originales):")
for t, o in zip(tipos_de_fraude, ocurrencias):
    print(f" - {t:<35}: {o:6d}")

print("\nOcurrencias (ordenadas):")
for t, o in zip(tipos_ordenados, ocurrencias_ordenadas):
    print(f" - {t:<35}: {o:6d}")

print(f"\nMedia: {media:.2f}")
print(f"Desviación estándar: {desviacion_estandar:.2f}")
```

Ocurrencias (originales):

- Fraude de monta deudas	:	7270
- Fraude por comercio electrónico	:	860
- Fraude por compra remota	:	5390
- Fraude de devolución	:	5191
- Fraude por aplicaciones falsas	:	5734
- Fraude por intercambio de SIM	:	6265
- Fraude por compromiso de cuenta	:	466
- Fraude por facturación duplicada	:	4426

Ocurrencias (ordenadas):

- Fraude por compromiso de cuenta	:	466
- Fraude por comercio electrónico	:	860
- Fraude por facturación duplicada	:	4426
- Fraude de devolución	:	5191
- Fraude por compra remota	:	5390
- Fraude por aplicaciones falsas	:	5734
- Fraude por intercambio de SIM	:	6265
- Fraude de monta deudas	:	7270

Media: 4450.25

Desviación estándar: 2320.80

Estimación del total de pérdidas

Cada tipo de fraude representa una pérdida promedio de **\$2,000**.

Se estima el impacto económico total multiplicando las ocurrencias por la pérdida promedio.

```
In [5]: perdidas_por_tipo = ocurrencias * perdida_promedio
total_perdidas = perdidas_por_tipo.sum()

print("Pérdidas estimadas por tipo de fraude:\n")
for t, p in zip(tipos_de_fraude, perdidas_por_tipo):
    print(f" - {t:<35}: ${p:,.2f}")

print(f"\n Total estimado de pérdidas: ${total_perdidas:,.2f}")
```

Pérdidas estimadas por tipo de fraude:

- Fraude de monta deudas	:	\$14,540,000.00
- Fraude por comercio electrónico	:	\$1,720,000.00
- Fraude por compra remota	:	\$10,780,000.00
- Fraude de devolución	:	\$10,382,000.00
- Fraude por aplicaciones falsas	:	\$11,468,000.00
- Fraude por intercambio de SIM	:	\$12,530,000.00
- Fraude por compromiso de cuenta	:	\$932,000.00
- Fraude por facturación duplicada	:	\$8,852,000.00

Total estimado de pérdidas: \$71,204,000.00

Visualización de resultados

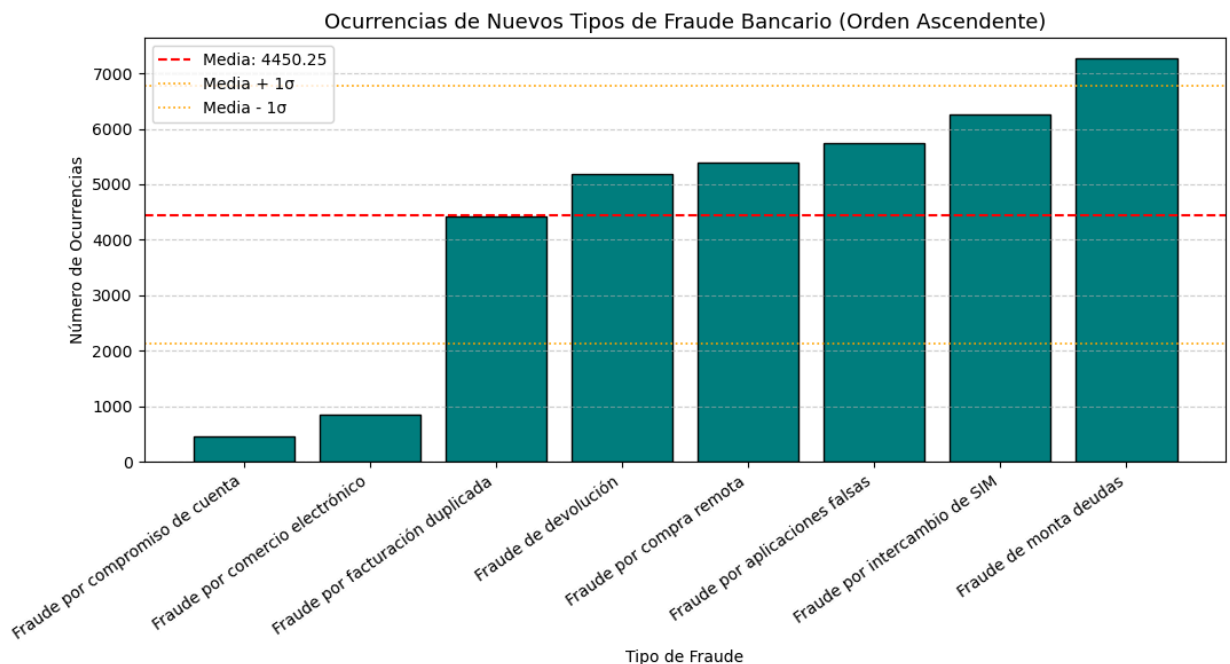
A continuación, se grafica el número de ocurrencias ordenadas de menor a mayor.

Se incluyen líneas de referencia para la **media** y la **desviación estándar**, permitiendo identificar fraudes con ocurrencias significativamente diferentes al promedio.

```
In [6]: plt.figure(figsize=(11, 6))
plt.bar(tipos_ordenados, ocurrencias_ordenadas, color='teal', edgecolor='black')
plt.title("Ocurrencias de Nuevos Tipos de Fraude Bancario (Orden Ascendente)",
plt.xlabel("Tipo de Fraude")
plt.ylabel("Número de Ocurrencias")
plt.xticks(rotation=35, ha='right')
plt.grid(axis='y', linestyle='--', alpha=0.6)

plt.axhline(media, color='red', linestyle='--', linewidth=1.5, label=f"Media: {media}")
plt.axhline(media + desviacion_estandar, color='orange', linestyle=':', linewidth=1.5, label=f"Media + 1σ")
plt.axhline(max(media - desviacion_estandar, 0), color='orange', linestyle=':', linewidth=1.5, label=f"Media - 1σ")

plt.legend()
plt.tight_layout()
plt.show()
```



Conclusión

El análisis evidencia que los distintos tipos de fraude bancario presentan **variabilidad considerable** en su frecuencia.

La media y la desviación estándar permiten identificar cuáles fraudes se encuentran **por encima o por debajo del comportamiento promedio**.

Entre ellos, el **fraude de monta deudas** se destaca como una amenaza emergente dentro del ecosistema financiero digital, reforzando la necesidad de:

- fortalecer los sistemas de **detección temprana**,
- mejorar los **modelos de aprendizaje automático** para detectar patrones inusuales, y
- concientizar a los usuarios sobre las **modalidades de estafa digital actuales**.

Este ejercicio demuestra cómo las herramientas de análisis de datos permiten simular, visualizar e interpretar tendencias en fraudes bancarios, siendo un paso esencial para el desarrollo de modelos predictivos en entornos de Big Data.

In []: