

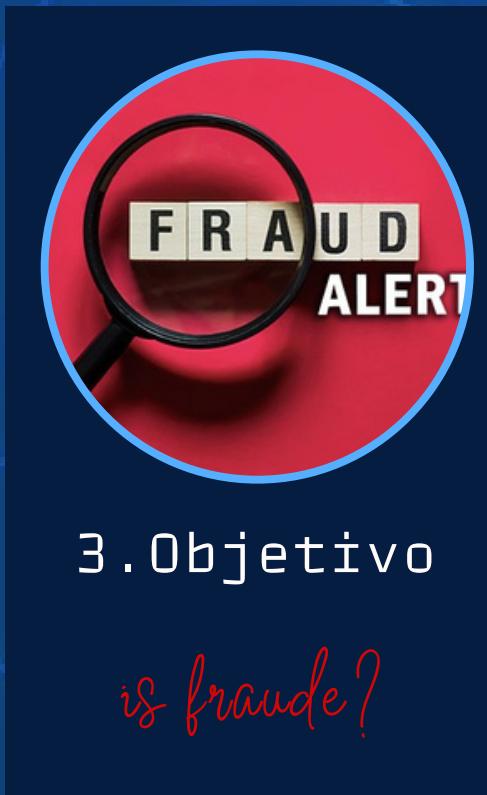
DETECCIÓN IN CREDIT CARDS

is fraude?

Proyecto Final Data Science - Gimena Cruz

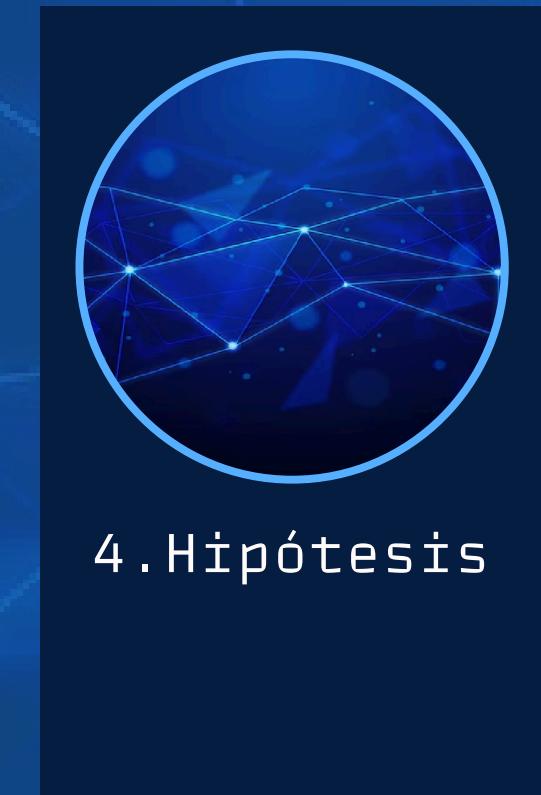


INDICE

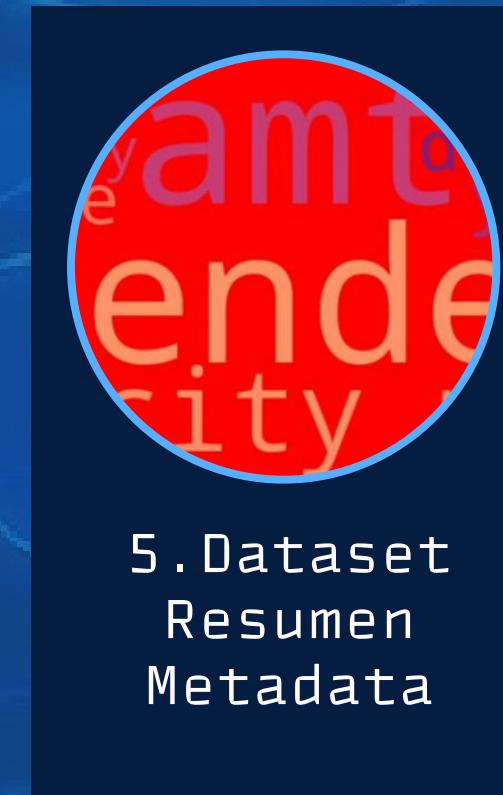


3. Objetivo

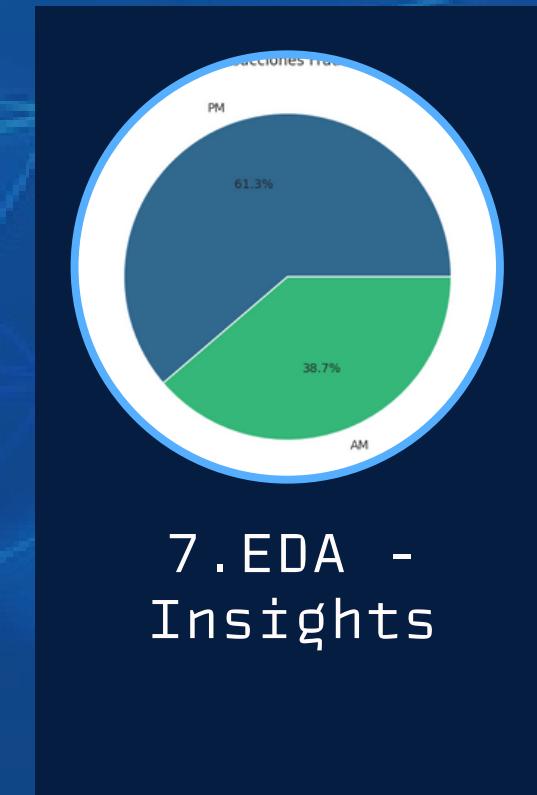
is fraude?



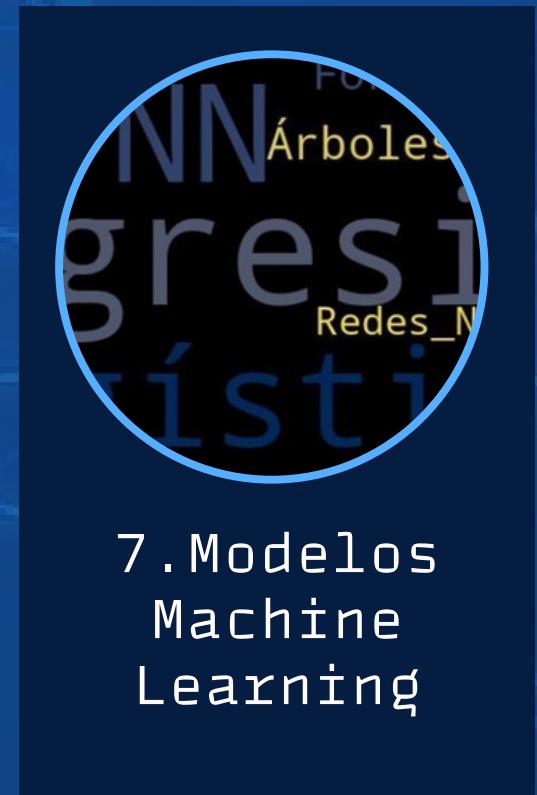
4. Hipótesis



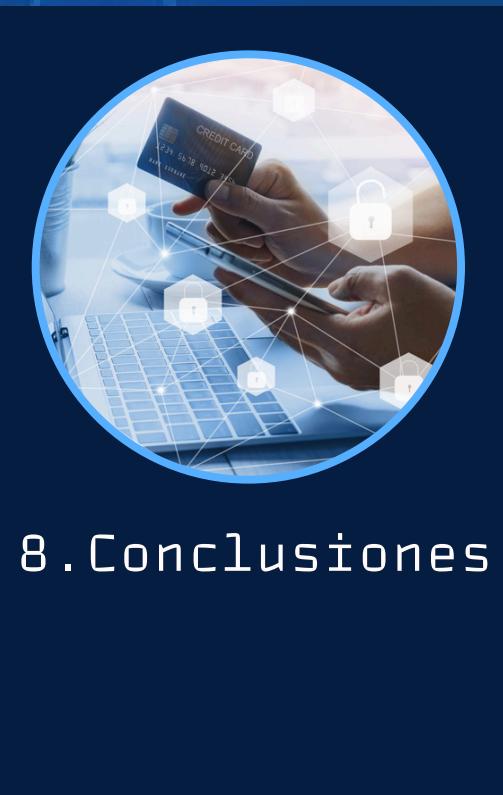
5. Dataset
Resumen
Metadata



7. EDA -
Insights



7. Modelos
Machine
Learning



8. Conclusiones

OBJETIVO



**FRAUD
ALERT**

Análisis e identificación de patrones + modelos predictivos de ML para la detección temprana de fraude orientado a instituciones financieras y clientes (titulares, comerciantes)



HIPOTESIS

“Las transacciones fraudulentas se realizan por única vez con la misma credit card.”

DATASETS



ORIGINAL

transacciones_long

1.296.675 x 21



REDUCIDO

transacciones_short

15.506 x 22

99.4%

NO

0.6%

SI

is fraude?

METADATA RESUMEN



ANALISIS DE DATOS_INSIGHTS

TRANSACCIONES FRAUDE CC FRECUENCIA AM/PM

7506

Transacciones
Fraude

01-01-19
21-06-20

Período

762

Tarjetas
únicas

02

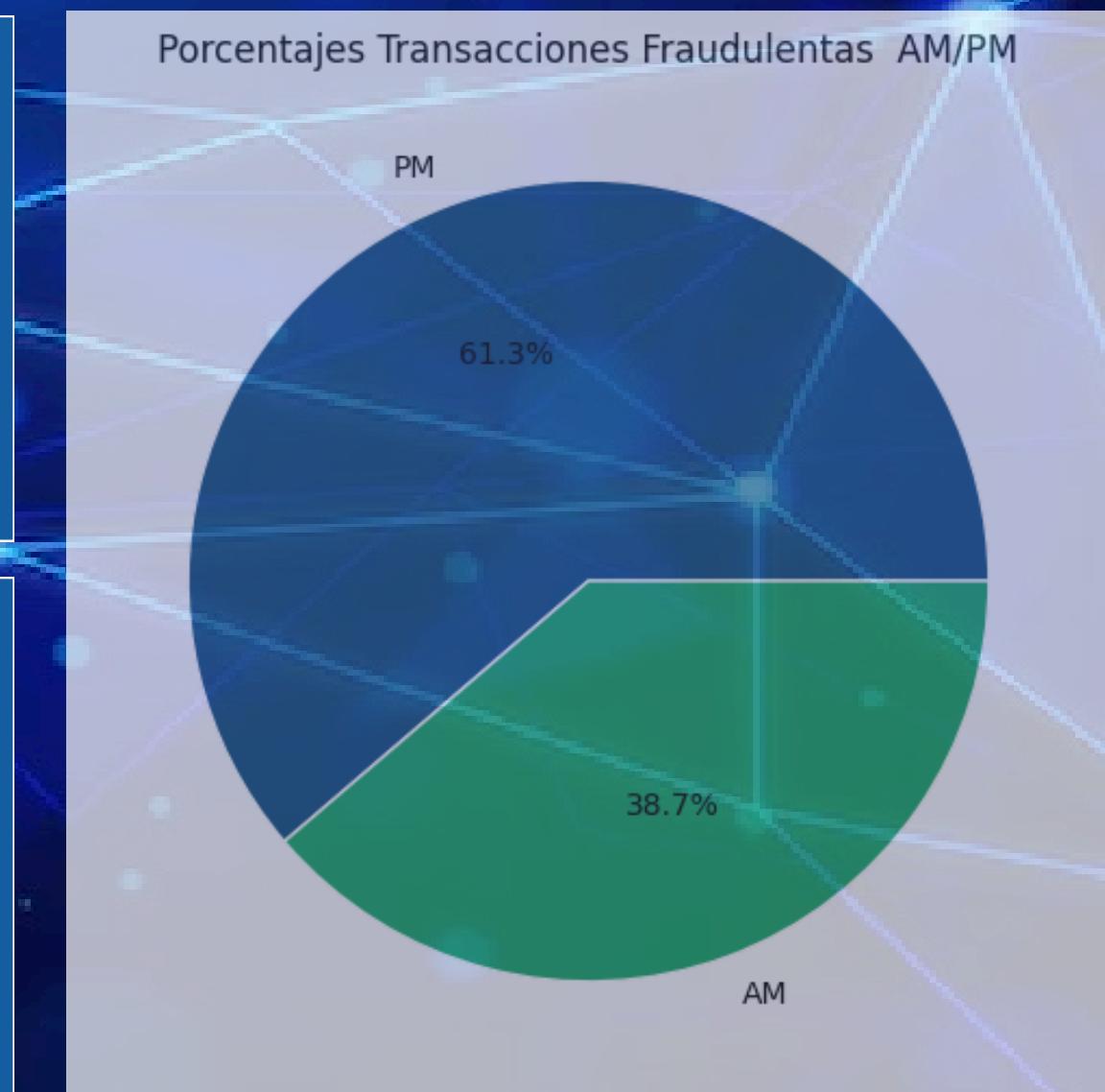
Mínimo uso
cc

10

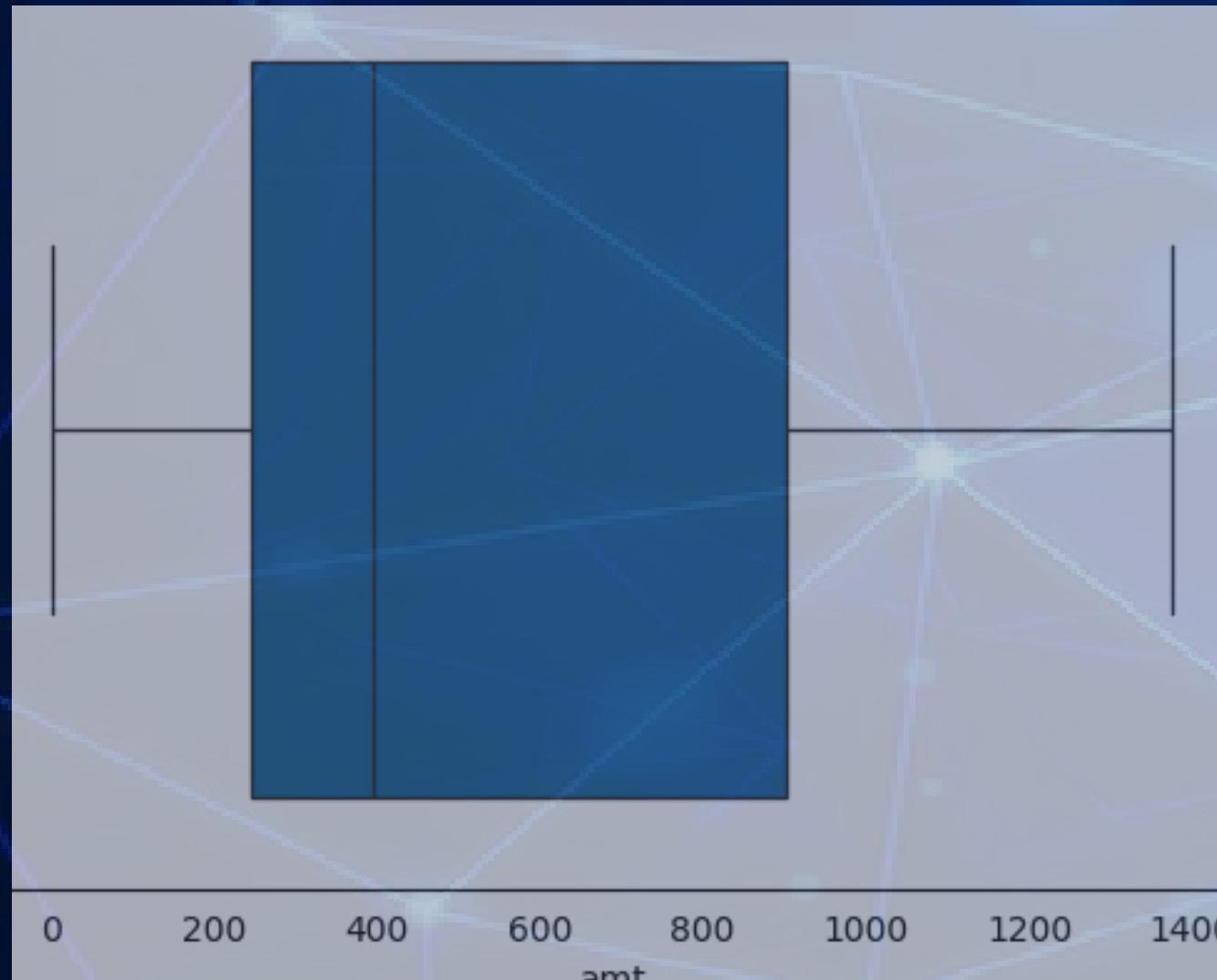
promedio de uso
cc

19

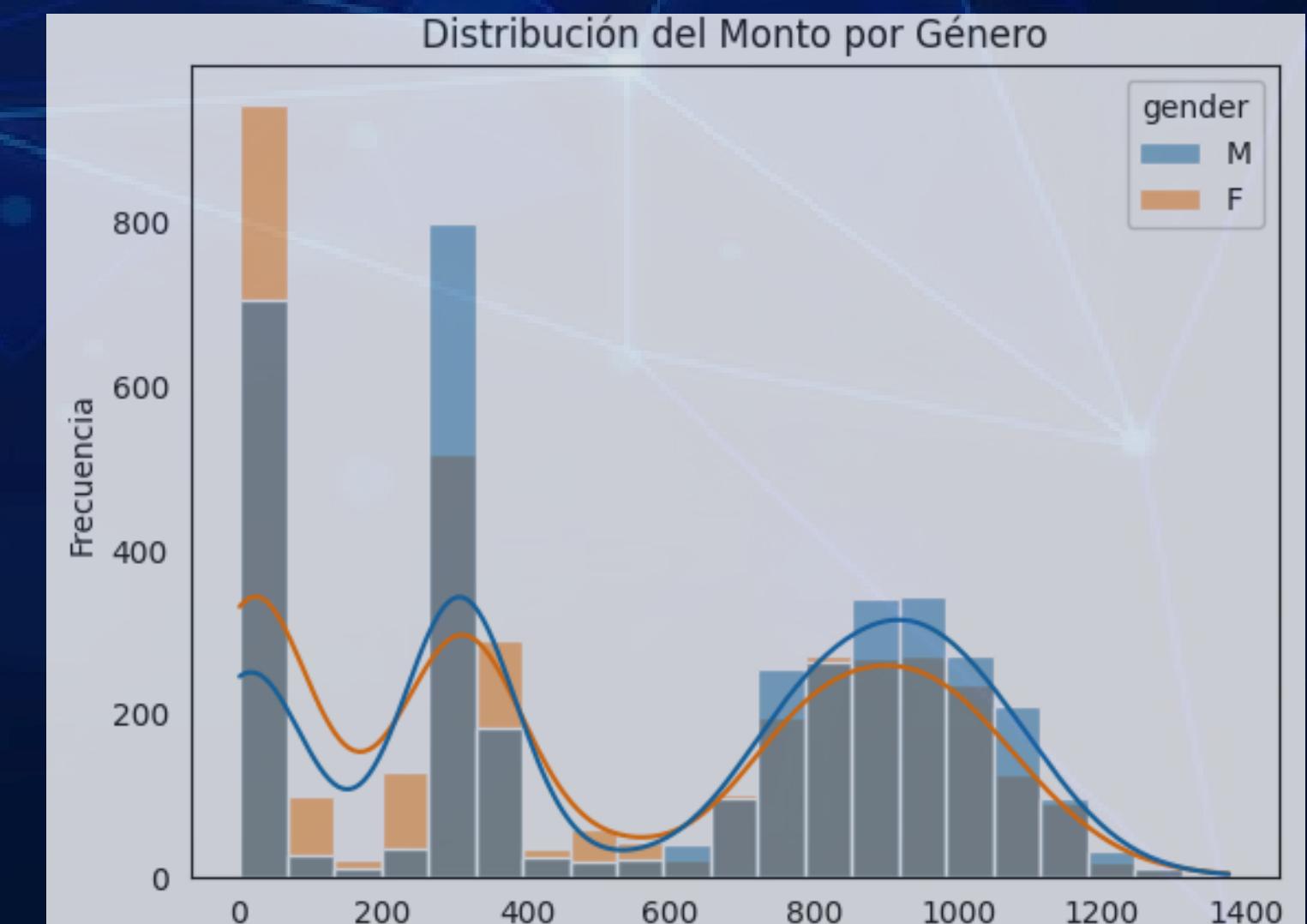
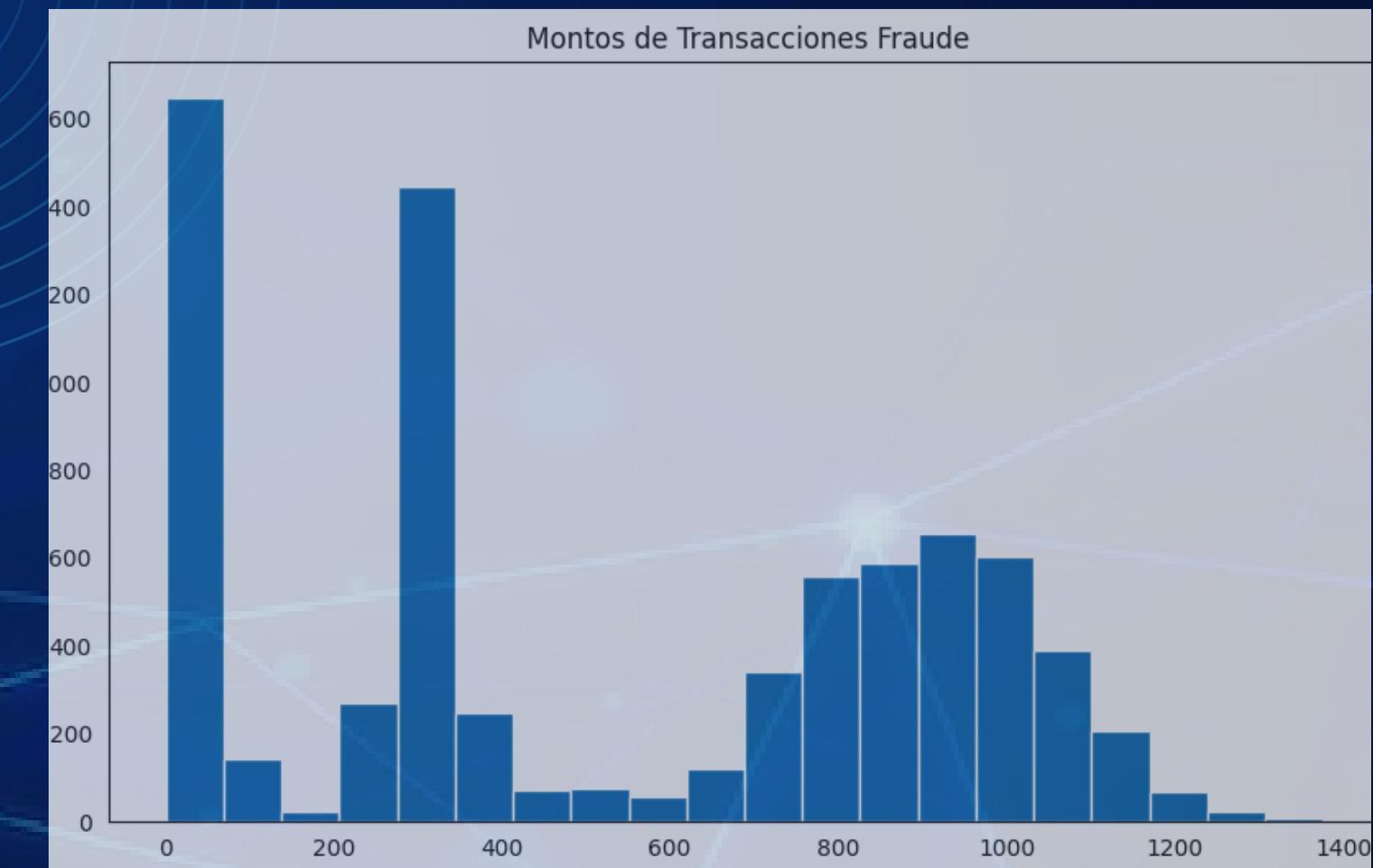
Máximo uso
cc



MONTO VS GÉNERO

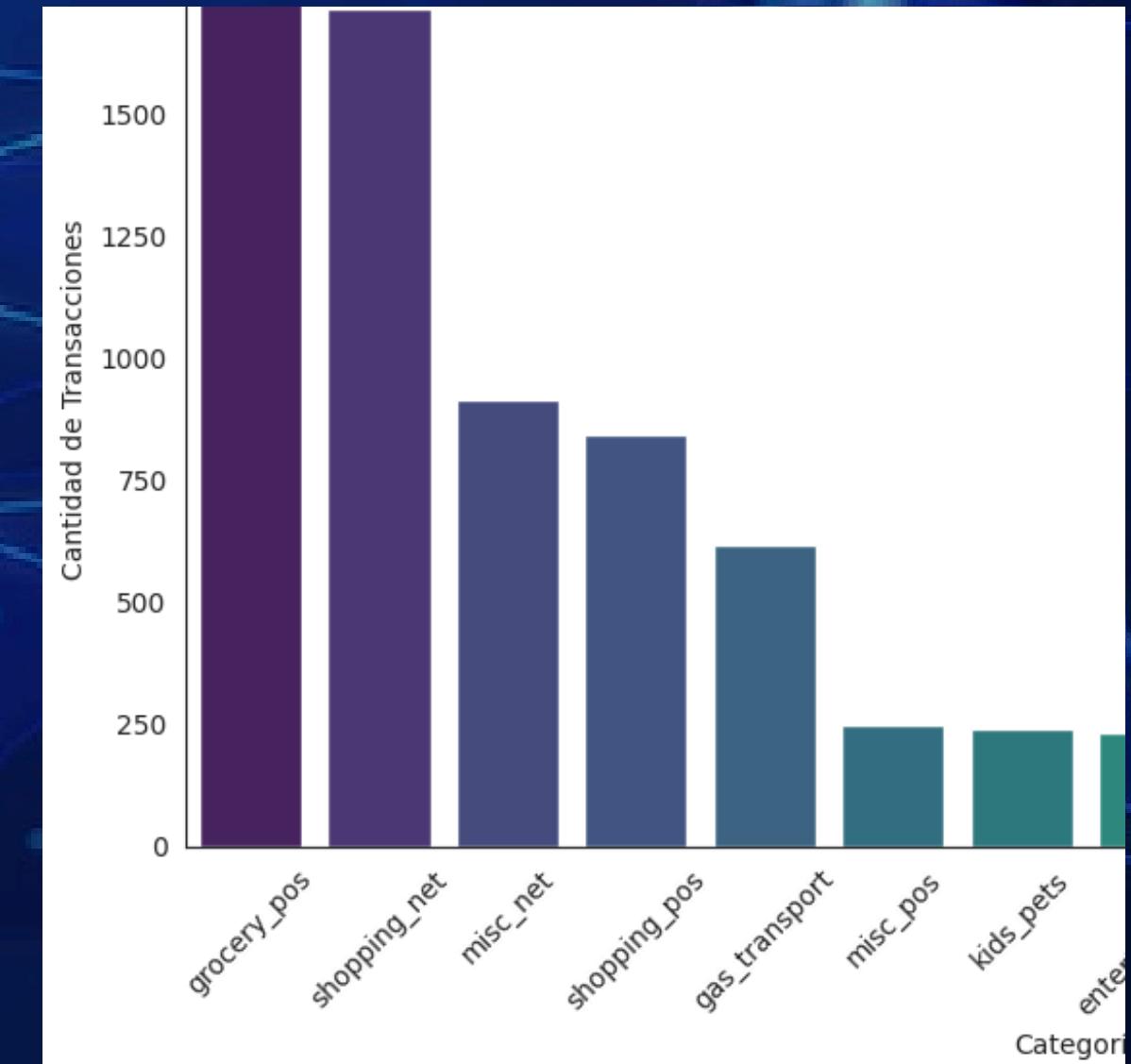


El monto más repetido es 400. Un 25% de transacciones con montos menores a 100 dólares. Un 50% son montos entre 250 y 900 dólares y otro 25% con montos entre 1000 y 1400 dólares.-

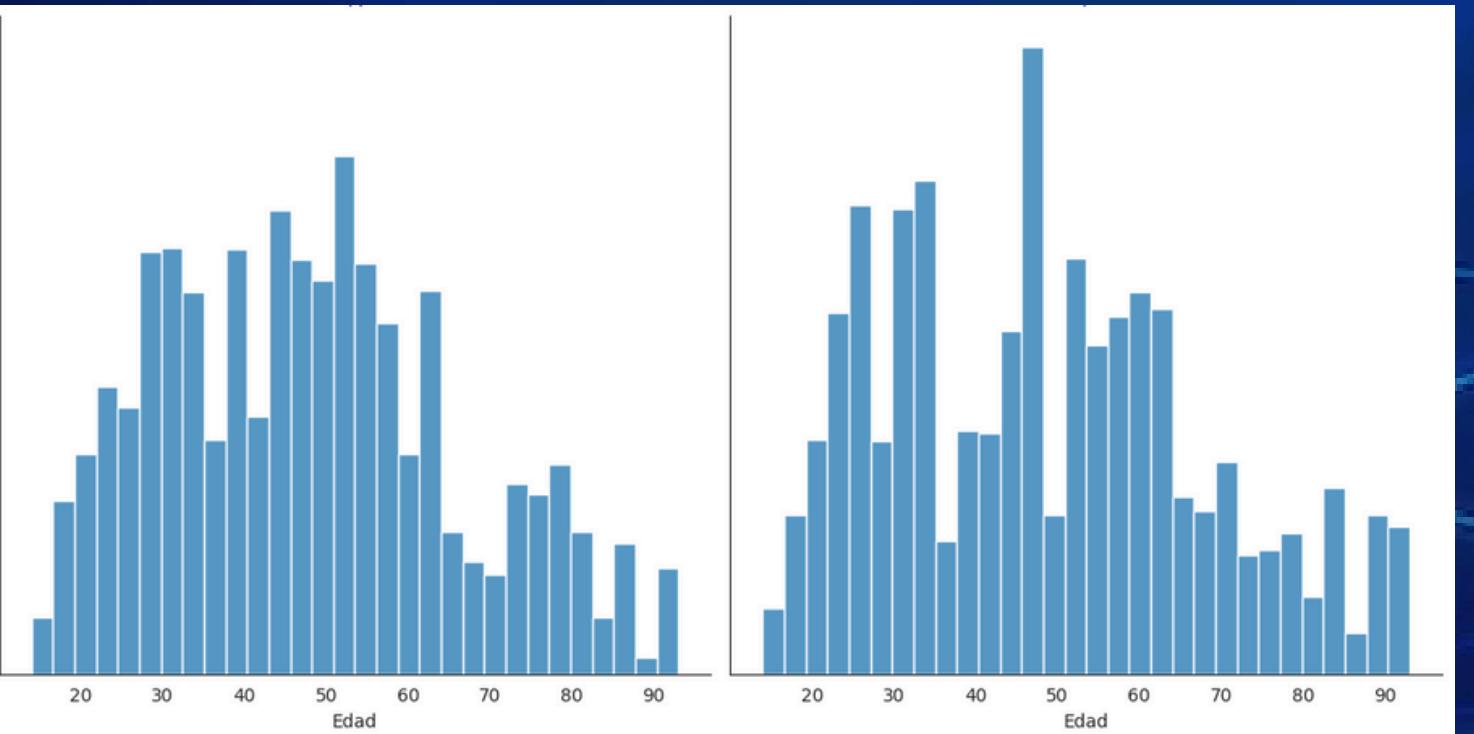
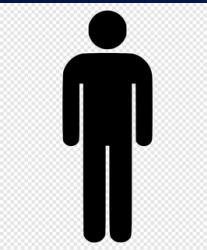


ANALISIS DE DATOS_INSIGHTS

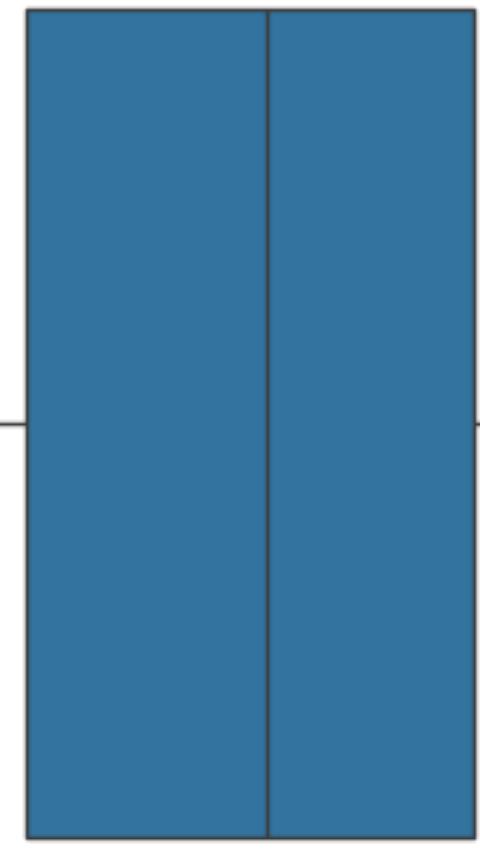
PROMEDIOS DE USO CC PROMEDIOS CATEGORY



GENDER AGE



48



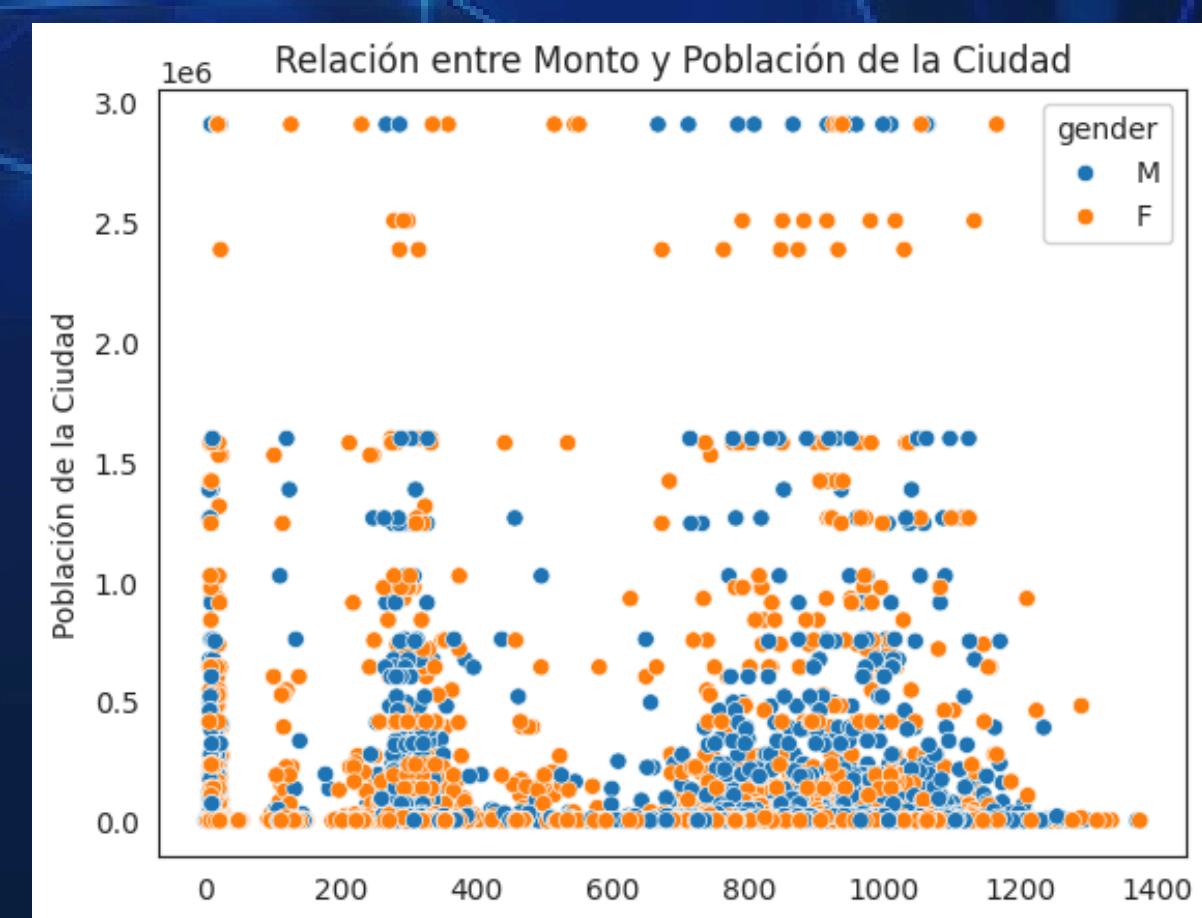
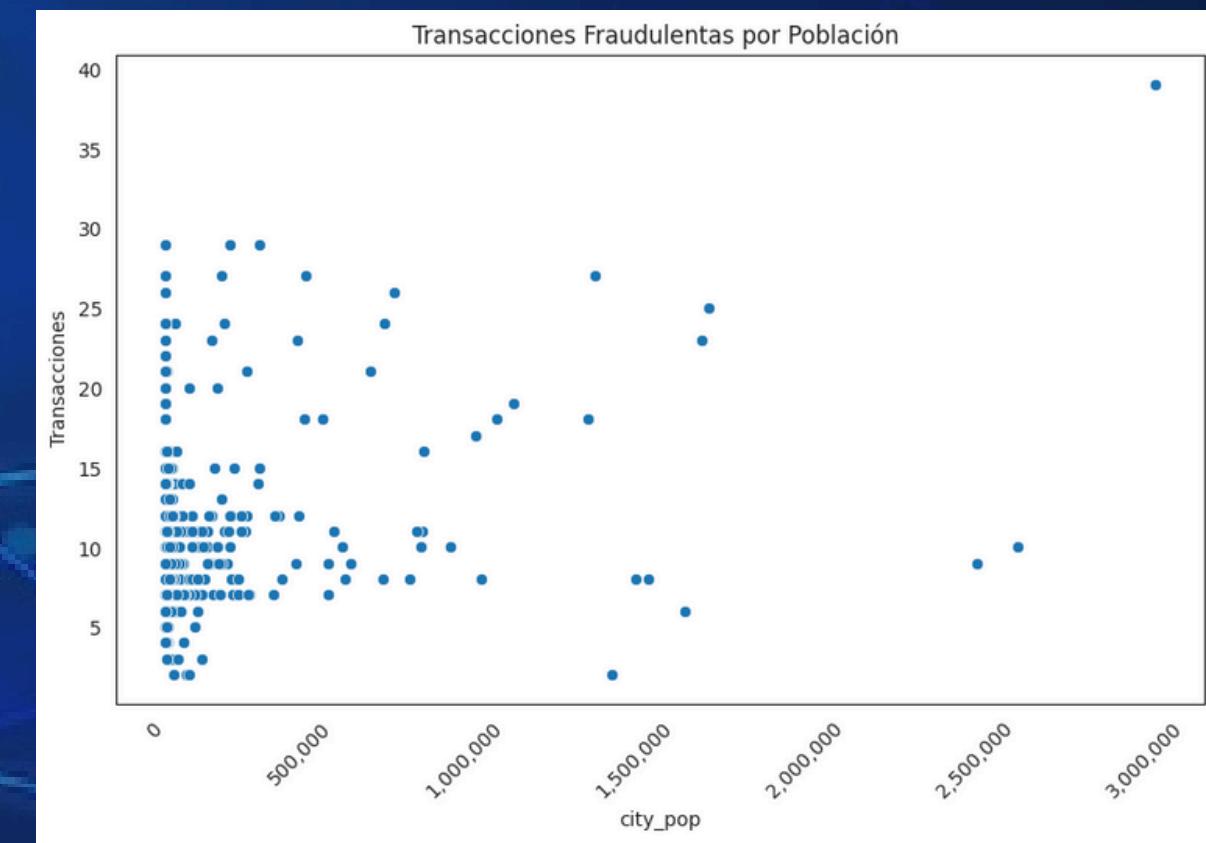
25% tienen menos de 32 años, otro 25% con edades > 60 años.
50% de titulares con edades entre 30 y 60 años.-

¿Fraude?

GENDER CITY POP

Relación monto con cantidad de población por ciudad y por género.

Concentración en poblaciones de pocos habitantes con algunos montos variables y otros en rango de 300 - 400 y 800-110 dólares.



INN 459356 NRO. VÁLIDO DE VISA
EMITIDO POR SCOTIABANK EN COLOMBIA

MODELO MACHINE LEARNING

▼ LogisticRegression  

LogisticRegression()

¡Gracias por su atención!

CONCLUSIONES

¡Gracias por su atención!

Recall: 0.97

Precision: 0.98

Train Accuracy: 0.98

Test Accuracy: 0.98

¡GRACIAS POR SU ATENCIÓN!