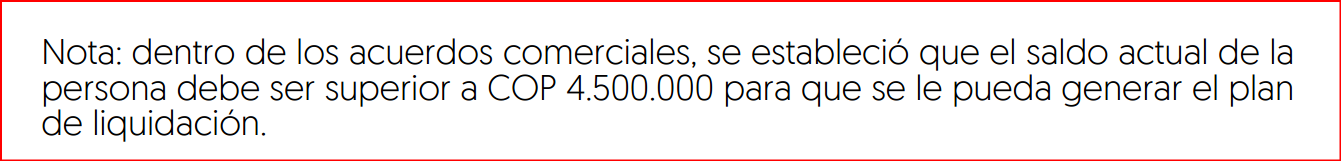
# PREGUNTAS

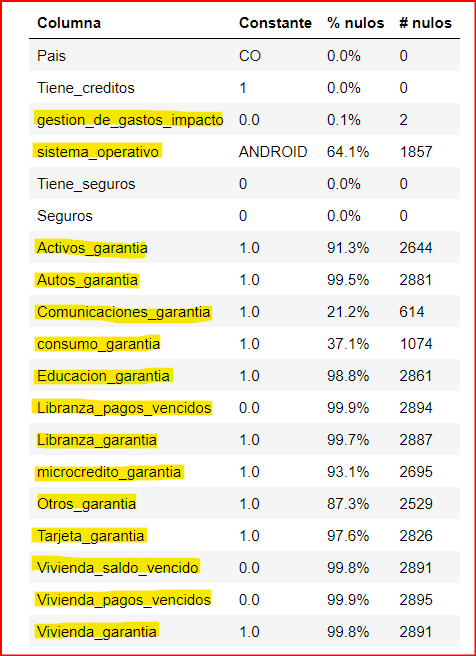
## Regla de negocio acuerdos comerciales



Aparece esta nota en la parte 1. Asumo que la regla aplica en el momento de predecir observaciones nuevas, pero me causa curiosidad que este en el apartado de entrenamiento. Encontré que la cantidad de registros que cumplen esta condición, en el conjunto de datos de entrenamiento, son 450 usuarios. ¿Esta regla debe aplicarse efectivamente en tiempo de predicción? Por favor aclarar la regla.

Si. Para eso se pone la nota aclaratoria. Recomiendo analizar si los datos descartados permiten encontrar relaciones estructurales de los datos o generar nuevas variables, a pesar de que esos usuarios no cumplan las características del aliado.

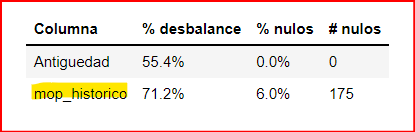
## Variables con valores constantes



Encontré una lista de variables que solo tienen un valor. Algunas no tienen valores nulos y por tanto las descarto, ya que no generan valor en el modelado. Sin embargo, aquellas subrayadas en amarillo necesito confirmar si tienen un valor por defecto que genere balance o contraparte. Por favor me indicas cuales y que valor por defecto, con el fin de completar o descartar las que no cumplan.

Si son nulos, el valor por defecto es cero. Si el valor no es cero, indica que existe algún tipo de garantía.

## Variables categóricas no balanceadas



En el MOP histórico los valores indicados como enteros están bien, sin embargo, el valor “N” me queda la duda si corresponde al MOP-00. ¿Por favor indicar el significado de este valor? Es buena aclarar que con este valor están el 78.8% de los datos en la muestra. ¿Cuál es el valor por defecto para registros nulos o se adicionaría una categoría “Sin MOP histórico?

Esa variable indica la morosidad. Te adjunto significados:

|  |  |
| --- | --- |
| Al día | N |
| Mora 30 días | 1 |
| Mora 60 días | 2 |
| Mora 90 días | 3 |
| Mora 120 días | 4 |
| Mora > 120 días | 5 y 6 |

## Variables con sesgo alto

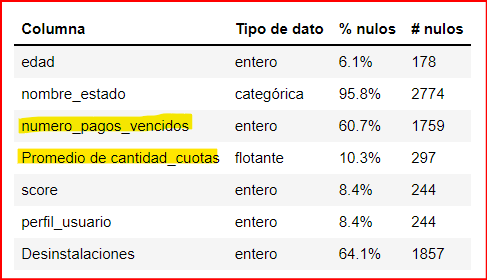
Tabla

Descripción generada automáticamente

Las variables sombreadas tienen valores negativos, exactamente el valor -1. ¿Estos se interpretan como cero o tienen un significado diferente para el negocio? La decisión que considero es volverlos cero.

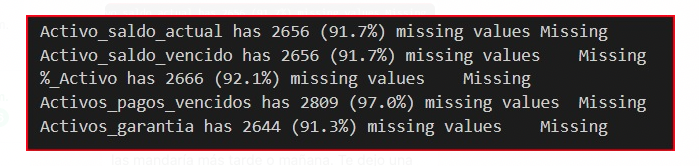
Si. Convertir a cero o como se considere

## Variables con valores nulos



La variable **numero\_pagos\_vencidos** tiene valores negativos (-1 y -10) y un alto porcentaje de valores nulos (60.7%). ¿Se puede considerar que tanto los negativos como los nulos se conviertan a cero? Por favor me explicas si los valores negativos tienen significado de negocio para tenerlo en cuenta. Pasa lo mismo con la variable Promedio de cantidad\_cuotas, la cual tiene valores nulos y negativos, realizar las mismas aclaraciones.

Los valores negativos no tienen significado. Imputar a conveniencia



Estoy terminando de analizar el último bloque de variables, donde todas tienen una estructura similar... Si me surgen preguntas al respecto, te las mandaría más tarde o mañana. Te dejo una muestra, porque hay varios bloques similares con Autos, Comunicaciones, Consumo, Educacion, Libranza, Microcredito, Otros, Tarjeta credito y Vivienda (Parece que hace alusión a productos que tiene el cliente)

OK