Apuntes Taller Final – Modulo 3: Coding para procesamiento de Datos

**Exploración df comportamiento tarjetas**

Hallazgos:

* Hay 1048575 observaciones de las cuales 209230 tienen la columna dias vencidos en nulo. Equivale al 20%
* De un total de 45,985 clientes con comportamiento del manejo de tarjeta, 25,005 clientes tienen al menos un campo nulo en dias vencidos.

Opino que debemos considerar todas las observaciones para definir el tipo de pagador, toda vez que podemos asegurar un mal comportamiento con información cierta. Y la información faltante de dias vencidos transcurridos podría ser MNAR (Missing not at random), es decir que no han estado en vencido en ese mes ni con pago en el día de vencimiento que sería el 0.

**Definiendo la etiqueta mal pagador**

Existen cuatro categorías de pagadores:

a) Los que quieren y pueden pagar. – buen pagador

b) Los que quieren pagar, pero por diferentes circunstancias no pueden hacerlo. – mal pagador o buen pagador.

c) Los que pudiendo pagar deciden no hacerlo. – mal pagador.

d) Los que no pueden pagar y no quieren hacerlo. – mal pagador.

Para definir un mal pagador propongo crear un score de comportamiento:

* Mayor que 25 y menor igual que 30, se puntúa el comportamiento con 0.25.
* Mayor que 30, se puntúa con 1

Con el punto uno buscaría etiquetar como mal pagador a los que pudiendo pagar, deciden no hacerlo de manera oportuna sino esperar hasta los últimos días y mantenerlo como un comportamiento habitual.

Definición mal pagador según su score y antigüedad

0 a <=6 meses, score mayor que 0.

>6 a <=12 meses, score mayor que a 0.5

>12 a <=24 meses, score mayor igual a 1.5.

>24 a <=36 meses, score mayor igual a 2.5.

>36 a <=48 meses, score mayor igual a 3.5.

>48 a <=60 meses, score mayor igual a 4.5.

>60 meses, score mayor igual a 5.5

**Busqueda ChatGPT**

En las economías latinoamericanas, las estadísticas de buenos y malos pagadores de tarjetas de crédito varían, pero se observan algunas tendencias comunes. En general, los niveles de morosidad en las tarjetas de crédito para consumo son significativos. Por ejemplo, en Ecuador, aproximadamente el 8.33% de los clientes con tarjetas de crédito de consumo tienen deudas vencidas【7†source】.

Además, se ha identificado que la morosidad total de los clientes con tarjetas de crédito para consumo representa el 14.2% de todos los clientes en esta modalidad de crédito. Este comportamiento refleja un riesgo significativo y debilidad en la capacidad de pago de los clientes, particularmente en los niveles de ingresos más bajos【7†source】.

En cuanto al impacto de los impagos, si un cliente se atrasa en el pago de una deuda más de 30 días, entra en mora, y si el retraso supera los 90 días, se clasifica como parte de la cartera castigada por los bancos. Estos reportes negativos pueden permanecer hasta por 4 años en los burós de crédito【8†source】.

Por otro lado, la implementación de sistemas de información crediticia que incluyen tanto datos positivos como negativos ha demostrado mejorar la inclusión financiera y reducir los costos del crédito. Países que han adoptado estos sistemas, como España, han visto una reducción en la morosidad y una mayor competencia entre las entidades financieras【9†source】.

Estas estadísticas y tendencias indican que, aunque existe una proporción considerable de malos pagadores en la región, también hay iniciativas para mejorar la gestión de crédito y reducir la morosidad.

**Join de la información adicional**

Después de la definición de mal pagador se examina si hay valores nulos para obtener las dos columnas que pertenecen al ID\_CLIENTE y TIPO\_CLIENTE.

Se procede a analizar el dataset que tiene información del cliente, no existen valores duplicados, pero si un 31% de valores faltantes para el campo de profesión.

Se observan que en la información hay valores repetidos en los ids, se analiza que no hay un patrón de los valores cambiantes. Se descarta por MCAR ya que la información de un cliente debe ser única y se enviará al departamento de Gobierno este subconjunto de datos que se rastreo cuál es la información verídica.

Se trabaja con un dataframe de valores únicos y se hará join con el conjunto de datos de la definición de mal pagador.

Tenemos 91% de datos sin categorizar para el tipo de cliente.