Exploración estadística del perfil de clientes de una entidad bancaria

En el mercado financiero, más precisamente en las entidades bancarias, debido a la mayor velocidad de procesamiento de las computadoras y el aumento continuo del tamaño de los dispositivos de almacenamiento de información, permiten que las entidades agreguen constantemente información de sus clientes en sus bases de datos, lo que las convierten en un poderoso reservorio de información útil para reducir la incertidumbre a la hora de tomar decisiones.

La base de datos que se analizará corresponde a clientes de una entidad bancaria, desde marzo a julio de 2023. Las variables registradas están relacionadas con la gestión y análisis de productos bancarios y el comportamiento financiero de los clientes. Este *dataset* puede ser utilizado para diversos análisis en el ámbito bancario, tales como: permitir identificar patrones de uso de productos y servicios, agrupar clientes según sus características demográficas y financieras, predecir la probabilidad de venta de productos financieros, mejorar la efectividad de campañas de *marketing* basadas en el comportamiento del cliente, evaluar la estabilidad financiera y el riesgo crediticio de los clientes, entre otras. Dentro de la base de datos se tienen los siguientes campos:

- **Variables Generales**

codmes: Mes de evaluación.

short_date: Día de impresión.

key_value: Número de documento encriptado del cliente.

target: Indicador de si la impresión generó una venta (0: no, 1: sí).

dia: Día de la impresión.

prox_15: Cercanía a la quincena del mes.

day_week: Orden del día de la semana.

- **Información Demográfica y de Segmentación**

rango_edad: Rango de edad utilizado para segmentar clientes.

flagcy_sex: Género del cónyuge (M-1, F-2).

flag_civil: Estado civil (C-3, V-2, S-1).

gency_m: Género del cónyuge masculino.

flag_costa: Indicador de si el cliente vive en una zona costera.

flag_frontera: Indicador de si el cliente vive en una zona fronteriza.

- **Información de Productos Bancarios**

productos: Total de productos del banco que posee el cliente.

saldo_prom_parcial: Saldo promedio de productos pasivos sin incluir plazo fijo.

saldo_prom_tot_millo_premia: Saldo promedio de la cuenta millonaria premia.

saldo_fdp_tot_txs: Saldo final de periodo de todas las transacciones.

saldo_fdp_tot_alcancia: Saldo final de periodo en productos de alcancía.

saldo_prom_tot_activo: Saldo promedio total de productos activos.

saldo_prom_tot_tc: Saldo promedio total de tarjetas de crédito.

saldo_prom_tot_extracash_tc: Saldo promedio total de extracash en tarjeta de crédito.

saldo_prom_tot_libre_dispon: Saldo promedio total de libre disponibilidad.

saldo_fdp_tot_activo: Saldo final de periodo total de productos activos.

saldo_fdp_tot_adel_sueldo: Saldo final de periodo de adelanto de sueldo.

saldo_fdp_tot_ahor_cred: Saldo final de periodo de ahorro-crédito.

saldo_fdp_tot_alcancia: Saldo final de periodo en productos de alcancía.

saldo_fdp_tot_convenio: Saldo final de periodo en productos de convenio.

saldo_fdp_tot_cts: Saldo final de periodo en CTS (Compensación por Tiempo de Servicios).

saldo_fdp_tot_extracash_tc: Saldo final de periodo de extracash en tarjeta de crédito.

saldo_fdp_tot_hipotecario: Saldo final de periodo de productos hipotecarios.

saldo_fdp_tot_hipotecario_cov19: Saldo final de periodo de productos hipotecarios reprogramados debido a COVID-19.

saldo_fdp_tot_inversion: Saldo final de periodo de productos de inversión.

saldo_fdp_tot_libre_dispon: Saldo final de periodo de libre disponibilidad.

saldo_fdp_tot_millo_premia: Saldo final de periodo de cuenta millonaria premia.

saldo_fdp_tot_millo_super_tasa: Saldo final de periodo de cuenta millonaria super tasa.

saldo_fdp_tot_millonaria: Saldo final de periodo de cuentas millonarias.

saldo_fdp_tot_pasivo: Saldo final de periodo de productos pasivos.

saldo_fdp_tot_planilla: Saldo final de periodo de productos de planilla.

saldo_fdp_tot_plazo_fijo: Saldo final de periodo de plazo fijo.

- **Información de Antigüedad y Cantidad de Productos**

ant_1_prod_month: Antigüedad en meses del primer producto en el banco hace un mes.

ant_1_prod_day: Antigüedad en días del primer producto en el banco hace un mes.

ant_1_prodact_day: Antigüedad en días del primer producto activo retail en el banco hace un mes.

ct_prodact: Cantidad de productos activos retail en el banco hace un mes.

ant_1_prodpas_day: Antigüedad en días del primer producto pasivo retail en el banco hace un mes.

ct_prodpas: Cantidad de productos pasivos retail en el banco hace un mes.

ant_1_prodpas_month: Antigüedad en meses del primer producto pasivo retail en el banco hace un mes.

ant_1_prodact_month: Antigüedad en meses del primer producto activo retail en el banco hace un mes.

- **Información Adicional**

flag_bonificacion: Indicador de bonificación extra en el mes.

tenencia: Cantidad de productos sin extracash, pago automático y seguros.

prm_sowe1ptcallsf06m: Promedio del share of wallet en la entidad principal por tarjeta de crédito reportado por el sistema financiero en los últimos seis meses.

prm_usoep1tcallsf06m_ingb: Promedio del uso de tarjeta de crédito en la entidad principal reportado por el sistema financiero en los últimos seis meses (entre el ingreso bruto).

prm_sowe1ptcallsf06m_ingb: Promedio del share of wallet en la entidad principal por tarjeta de crédito reportado por el sistema financiero en los últimos seis meses (entre el ingreso bruto).

prm_usotcrrstsf06m_ingb: Promedio del porcentaje de uso de la tarjeta de crédito de otras entidades bancarias en los últimos seis meses (entre el ingreso bruto).

prm_usotcribksf06m_ingb: Promedio del % de uso de la tarjeta de crédito del banco en los últimos 6 meses (entre el ingreso bruto).

El objetivo es llevar a cabo un análisis predictivo y/o exploratorio exhaustivo con el propósito de discernir y clasificar aquellos clientes con una inclinación alta hacia la adquisición de plazos fijos. Este análisis se fundamentará en el escrutinio meticuloso de las variables contextuales, demográficas y transaccionales dispuestas en el conjunto de datos proporcionado por la entidad bancaria, lo que permitirá a dicha entidad potenciar las ventas de plazos fijos.