

El presente Diccionario de Datos describe las características de las solicitudes de cuentas bancarias sintéticas utilizadas en el análisis. Cada registro representa una aplicación con diversas características tanto numéricas como categóricas, diseñadas para capturar información relevante sobre el solicitante y su solicitud.

Comprobemos las 32 variables que conforman el conjunto de datos:

- `fraud_bool` (binaria, variable target)\*\*: Indica si la solicitud es fraudulenta o no.
- `income` (numérica): Ingreso anual del solicitante en forma de decil (0.1 a 0.9).
- `name_email_similarity` (numérica): Similitud entre el correo electrónico y el nombre del solicitante (0 a 1).
- `prev_address_months_count` (numérica): Meses en la dirección anterior del solicitante (-1 a 380, donde -1 indica valor faltante).
- `current_address_months_count` (numérica): Meses en la dirección actual del solicitante (-1 a 429, donde -1 indica valor faltante).
- `customer_age` (numérica): Edad del solicitante, redondeada a la década (10 a 90 años).
- `days_since_request` (numérica): Días transcurridos desde que se realizó la solicitud (0 a 79 días).
- `intended_balcon_amount` (numérica): Monto inicial transferido para la solicitud (-16 a 114, los valores negativos indican datos faltantes).
- `payment_type` (categórica): Tipo de plan de pago de crédito. 5 valores posibles anónimos.
- `zip_count_4w` (numérica): Número de solicitudes dentro del mismo código postal en las últimas 4 semanas (1 a 6830).
- `velocity_6h` (numérica): Velocidad de solicitudes totales en las últimas 6 horas (-175 a 16818).
- `velocity_24h` (numérica): Velocidad de solicitudes totales en las últimas 24 horas (1297 a 9586).
- `velocity_4w` (numérica): Velocidad de solicitudes totales en las últimas 4 semanas (2825 a 7020).
- `bank_branch_count_8w` (numérica): Número de solicitudes totales en la sucursal bancaria seleccionada en las últimas 8 semanas (0 a 2404).
- `date_of_birth_distinct_emails_4w` (numérica): Número de correos electrónicos de solicitantes con la misma fecha de nacimiento en las últimas 4 semanas (0 a 39).
- `employment_status` (categórica): Estado de empleo del solicitante. 7 valores posibles anónimos.

- `credit_risk_score` (numérica): Puntuación interna de riesgo de la solicitud (-191 a 389).
- `email_is_free` (binaria): Si el dominio del correo electrónico de la solicitud es gratuito o de pago.
- `housing_status` (categórica): Estado residencial actual del solicitante. 7 valores posibles anónimos.
- `phone_home_valid` (binaria): Validez del teléfono fijo proporcionado.
- `phone_mobile_valid` (binaria): Validez del teléfono móvil proporcionado.
- `bank_months_count` (numérica): Antigüedad de la cuenta bancaria anterior en meses (-1 a 32, donde -1 indica valor faltante).
- `has_other_cards` (binaria): Si el solicitante tiene otras tarjetas de la misma entidad bancaria.
- `proposed_credit_limit` (numérica): Límite de crédito propuesto para el solicitante (200 a 2000).
- `foreign_request` (binaria): Si el país de origen de la solicitud es diferente al del banco.
- `source` (categórica): Fuente en línea de la solicitud, navegador (INTERNET) o aplicación (TELEAPP).
- `session_length_in_minutes` (numérica): Duración de la sesión del usuario en el sitio web bancario en minutos (-1 a 107, donde -1 indica valor faltante).
- `device_os` (categórica): Sistema operativo del dispositivo que realizó la solicitud.
- `keep_alive_session` (binario): Opción del usuario sobre el cierre de sesión.
- `device_distinct_emails` (numérica): Número de correos electrónicos distintos en el sitio web bancario desde el dispositivo utilizado en las últimas 8 semanas (-1 a 2, donde -1 indica valor faltante).
- `device_fraud_count` (numérica): Número de solicitudes fraudulentas realizadas con el dispositivo en cuestión. Varía entre [0, 1].
- `month` (numérica): Mes en el que se realizó la solicitud. Varía entre [0, 7].