

## Proyecto Final: Análisis Exploratorio de Pérdida de Clientes (CHURN)

En el sector bancario, la pérdida de clientes (churn) representa un desafío significativo, ya que implica pérdidas financieras y disminución de la satisfacción del cliente. Este proyecto se centra en realizar un análisis exploratorio de datos exhaustivo para entender en profundidad el comportamiento de los clientes y los factores asociados a su decisión de abandonar la entidad. El objetivo es extraer conocimiento a partir del dataset para identificar patrones, tendencias y relaciones entre las variables, lo que servirá de base para futuras estrategias de retención.

### Descripción de las Variables

- **RowNumber:**  
Identificador único para cada registro. Aunque no aporta información relevante para el análisis, ayuda a organizar y referenciar los datos.
- **CustomerId:**  
Permite rastrear y diferenciar a cada cliente en el conjunto de datos.
- **Surname:**  
Proporciona información sobre el apellido del cliente, útil para análisis descriptivos o segmentaciones familiares.
- **CreditScore:**  
Valor numérico que evalúa la solvencia crediticia del cliente basado en su historial y comportamiento financiero.
- **Geography:**  
Indica la ubicación geográfica de los clientes, lo que posibilita realizar análisis basados en factores regionales o nacionales.
- **Gender:**  
Categoriza a los clientes según su género, permitiendo identificar posibles diferencias en el comportamiento o en la propensión al churn.
- **Age:**  
Representa la edad del cliente en años, útil para detectar patrones relacionados con distintos grupos etarios.
- **Tenure:**  
Mide el tiempo (en años o meses) que el cliente ha estado vinculado al banco, lo que puede influir en su fidelidad o en la decisión de abandonar.
- **Balance:**  
Monto de dinero presente en la cuenta bancaria del cliente en un momento dado, indicador de su situación financiera.
- **NumOfProducts:**  
Número de productos o servicios bancarios que utiliza el cliente (cuentas, préstamos, tarjetas de crédito, etc.), lo que puede reflejar su nivel de compromiso con el banco.
- **HasCrCard:**  
Variable binaria que indica si el cliente posee una tarjeta de crédito (1) o no (0).
- **IsActiveMember:**  
Variable binaria que señala si el cliente es un miembro activo del banco (1) o no (0), lo cual puede estar relacionado con su fidelidad.

- **EstimatedSalary:**  
Aproximación del nivel de ingresos del cliente, relevante para analizar la relación entre el poder adquisitivo y la propensión a abandonar el banco.
- **Exited:**  
Indicador que señala si el cliente ha abandonado (1) o se ha quedado (0) en el banco, siendo la variable de interés para analizar el churn.

## Entregables del Proyecto

El estudiante debe realizar un análisis exploratorio de datos (EDA) robusto y detallado que incluya:

- **Análisis Descriptivo:**
  - Revisión del tipo de datos y cantidad de registros.
  - Estadísticas descriptivas (media, mediana, desviación estándar, percentiles, etc.).
  - Identificación y tratamiento de valores faltantes o duplicados.
- **Visualizaciones:**
  - Histogramas y boxplots para analizar la distribución de variables numéricas (CreditScore, Age, Balance, EstimatedSalary, etc.).
  - Gráficos de barras para variables categóricas (Geography, Gender, HasCrCard, IsActiveMember).
  - Series temporales o análisis de tendencias (si se identifican datos relevantes en variables como Tenure o comportamientos a lo largo del tiempo).
  - Matrices de correlación que faciliten la identificación de relaciones entre variables.
- **Visualizaciones Interactivas:**
  - Desarrollo de gráficos interactivos que permitan explorar dinámicamente el dataset y profundizar en la identificación de patrones.
- **Conclusiones:**
  - Una sección final donde se resuman los hallazgos clave del análisis, destacando las tendencias y relaciones relevantes que podrían influir en la decisión de los clientes de abandonar el banco.

## Nota Adicional:

Aunque el enfoque principal es el análisis exploratorio, se valorará positivamente el desarrollo de un modelo de churn. Se espera que el trabajo haga uso de las librerías vistas en el curso, tales como NumPy, Pandas, Polars, SciPy, Plotly, seaborn, entre otras.

**El trabajo deberá ser entregado en un Notebook. Puede ser montado en el aula virtual o subirlo a un repositorio en GitHub.**