**Exploratory Data Analysis (EDA) Report**

**Objetivo del análisis**

Este análisis tiene como objetivo entender qué factores influyen en que las personas acepten suscribir depósitos a plazo fijo, basándonos en la información de clientes recopilada de las campañas de marketing telefónico realizadas por un banco en Portugal.

**1. Transformación y limpieza de los datos**

* Se corrigieron errores de formato en el archivo CSV original, eliminando comas innecesarias y limpiando las comillas dobles.
* Se homogeneizaron formatos numéricos, reemplazando comas por puntos en columnas como cons.price.idx y euribor3m.
* Se unieron dos conjuntos de datos (bank-additional.csv y customer-details.xlsx) usando el campo ID.

**2. Análisis descriptivo**

**Distribución de variables clave:**

* **Edad**: la mayoría de los clientes tienen entre 30 y 50 años.
* **Ingreso**: distribuido de forma sesgada a la derecha. Se detectaron algunos valores extremos.
* **Duración de llamada**: se observan diferencias notables entre aquellos clientes que se suscribieron y los que no.
* **Campañas**: la mayoría de los contactos se realizaron 1 o 2 veces por cliente.

**Correlaciones relevantes:**

* Se detectaron correlaciones fuertes entre variables económicas como euribor3m, emp.var.rate y nr.employed.
* Las correlaciones identificadas entre las variables económicas, como euribor3m, emp.var.rate y nr.employed, reflejan patrones comunes que suelen estar influenciados por factores macroeconómicos. Por ello, es importante analizarlas con cuidado para evitar posibles sesgos o interpretaciones inexactas en modelos predictivos.

**3. Visualizaciones**

Se generaron gráficos de barras, histogramas y mapas de calor para ilustrar:

* La distribución de campañas exitosas vs. fallidas por segmentos (job, education, housing).
* Diferencias en duración de llamadas y número de contactos según el resultado (y).
* Boxplots para variables numéricas críticas.

Se utilizaron librerías de Python de visualización como Matplotlib y Seaborn

**4. Principales insights**

* **Perfiles con alta conversión**:
  + Profesionales con educación superior (tertiary), sin hipotecas, y contacto por celular presentan mayor tasa de éxito.
  + Los perfiles con tasas de conversión más altas fueron identificados agrupando por job, education y marital.
* **Impacto del channel**:
  + Las campañas realizadas mediante teléfono móvil (contact = cellular) tuvieron una tasa de éxito superior comparado con el canal telefónico fijo.
* **Duración y número de contactos**:
  + Las llamadas más largas están asociadas a mayor probabilidad de conversión.
  + Un exceso de contactos (campaign > 3) no mejora significativamente la tasa de conversión y puede ser contraproducente.