
Análisis del abandono de clientes en BlueBank



Análisis del abandono de clientes en BlueBank

• Índice de contenidos:

Introducción	<ul style="list-style-type: none">• Breve introducción al proyecto
Tecnología	<ul style="list-style-type: none">• Tecnologías empleadas en el desarrollo• Apoyo con capturas de pantalla
Flujo análisis	<ul style="list-style-type: none">• Fases del proyecto• Procesos realizados
Conclusiones	<ul style="list-style-type: none">• Hallazgos y estrategias propuestas• Explicación con dashboards en PowerBI
Cierre	<ul style="list-style-type: none">• Links de interés





CONTEXTO

BlueBank es una entidad financiera que busca comprender los factores que influyen en el **abandono de clientes** con el objetivo de mejorar la retención.

OBJETIVOS DEL ANÁLISIS

- Identificar patrones en el comportamiento de los clientes.
- Analizar factores clave que afectan la retención.
- Proponer estrategias basadas en datos para reducir el abandono.

ENFOQUE

- Exploración y análisis de datos históricos de clientes.
- Segmentación de clientes según su historial de abandono.
- Generación de insights accionables para el banco.

HERRAMIENTAS UTILIZADAS



Introducción

Tecnología

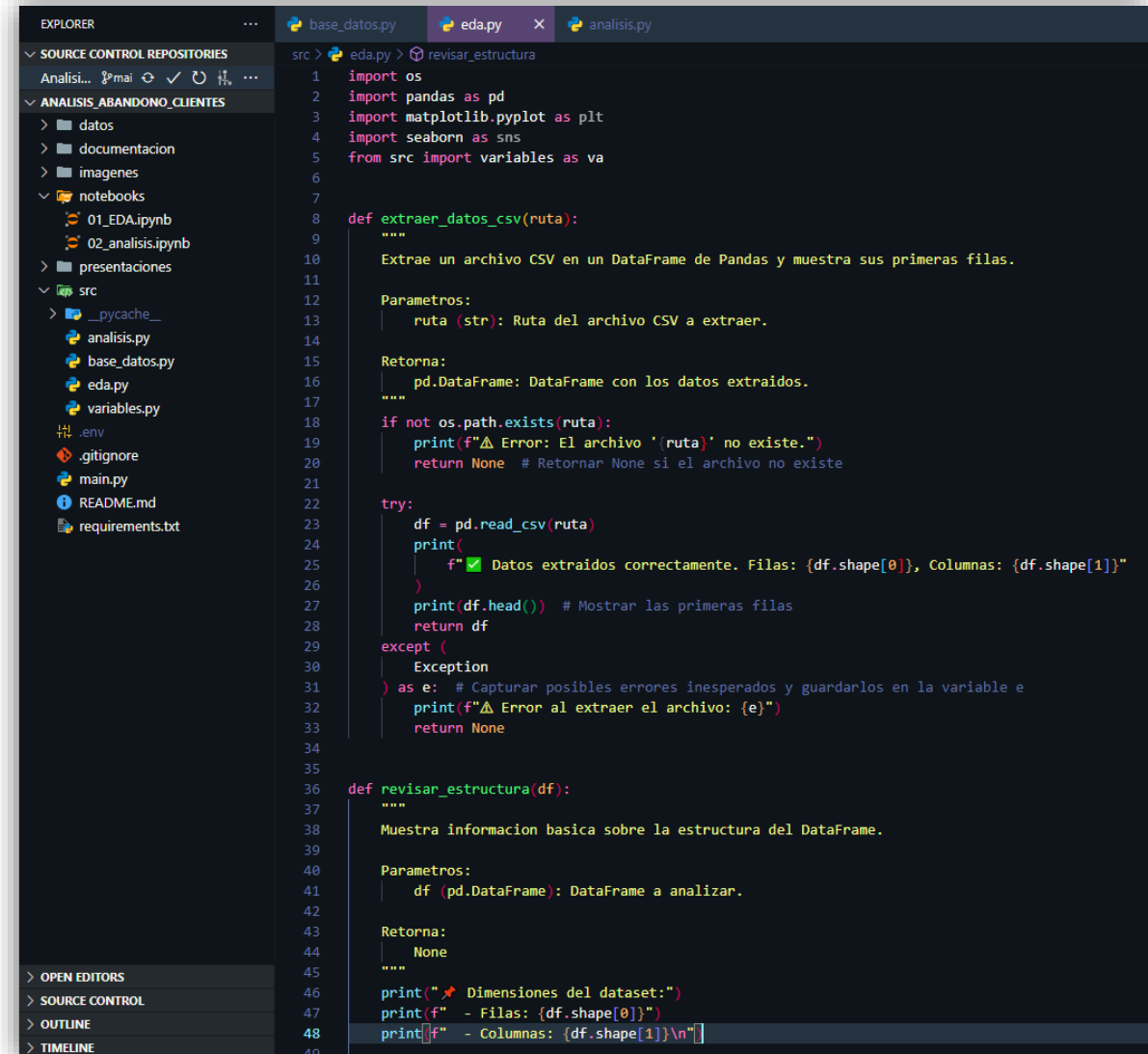
Flujo
análisis

Conclusiones

Cierre

Lenguaje principal: PYTHON

- **Análisis, limpieza, transformación y carga** de los datos.
- **Estructura modular:** Código organizado en distintos módulos para facilitar la reutilización y mantenibilidad.
- **Librerías clave:** pandas, matplotlib, seaborn y mysql-connector-python.
- **Automatización de procesos.**
- **Gestión de la base de datos** para conectar, crear tablas e insertar datos en MySQL.



```
EXPLORER
  SOURCE CONTROL REPOSITORIES
  Analisis...
  ANALISIS_ABANDONO_CLIENTES
    datos
    documentacion
    imagenes
    notebooks
      01_EDA.ipynb
      02_analisis.ipynb
    presentaciones
    src
      __pycache__
      analisis.py
      base_datos.py
      eda.py
      variables.py
    .env
    .gitignore
    main.py
    README.md
    requirements.txt

base_datos.py
eda.py
analysis.py

src > eda.py > revisar_estructura
1 import os
2 import pandas as pd
3 import matplotlib.pyplot as plt
4 import seaborn as sns
5 from src import variables as va
6
7
8 def extraer_datos_csv(ruta):
9     """
10     Extrae un archivo CSV en un DataFrame de Pandas y muestra sus primeras filas.
11
12     Parametros:
13         ruta (str): Ruta del archivo CSV a extraer.
14
15     Retorna:
16         pd.DataFrame: DataFrame con los datos extraidos.
17     """
18     if not os.path.exists(ruta):
19         print(f"⚠ Error: El archivo '{ruta}' no existe.")
20         return None # Retornar None si el archivo no existe
21
22     try:
23         df = pd.read_csv(ruta)
24         print(
25             f"✅ Datos extraidos correctamente. Filas: {df.shape[0]}, Columnas: {df.shape[1]}"
26         )
27         print(df.head()) # Mostrar las primeras filas
28         return df
29     except (
30         Exception
31     ) as e: # Capturar posibles errores inesperados y guardarlos en la variable e
32         print(f"⚠ Error al extraer el archivo: {e}")
33         return None
34
35
36 def revisar_estructura(df):
37     """
38     Muestra informacion basica sobre la estructura del DataFrame.
39
40     Parametros:
41         df (pd.DataFrame): DataFrame a analizar.
42
43     Retorna:
44         None
45     """
46     print("🔥 Dimensiones del dataset:")
47     print(f"  - Filas: {df.shape[0]}")
48     print(f"  - Columnas: {df.shape[1]}\n")
49
```

Introducción

Tecnología

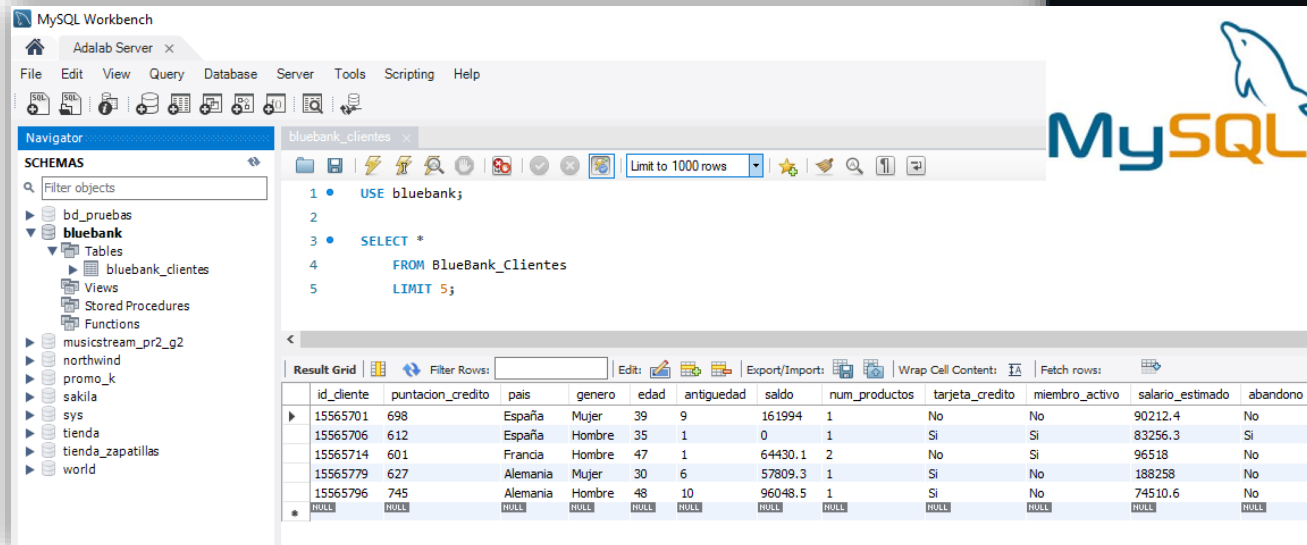
Flujo
análisis

Conclusiones

Cierre

Base de datos: MySQL

- Creación base de datos, definición de tabla *BlueBank_Clientes* e inserción de datos desde Python.
- Consultas para exploración de datos en MySQL Workbench.
- Carga directa desde MySQL a Power BI.

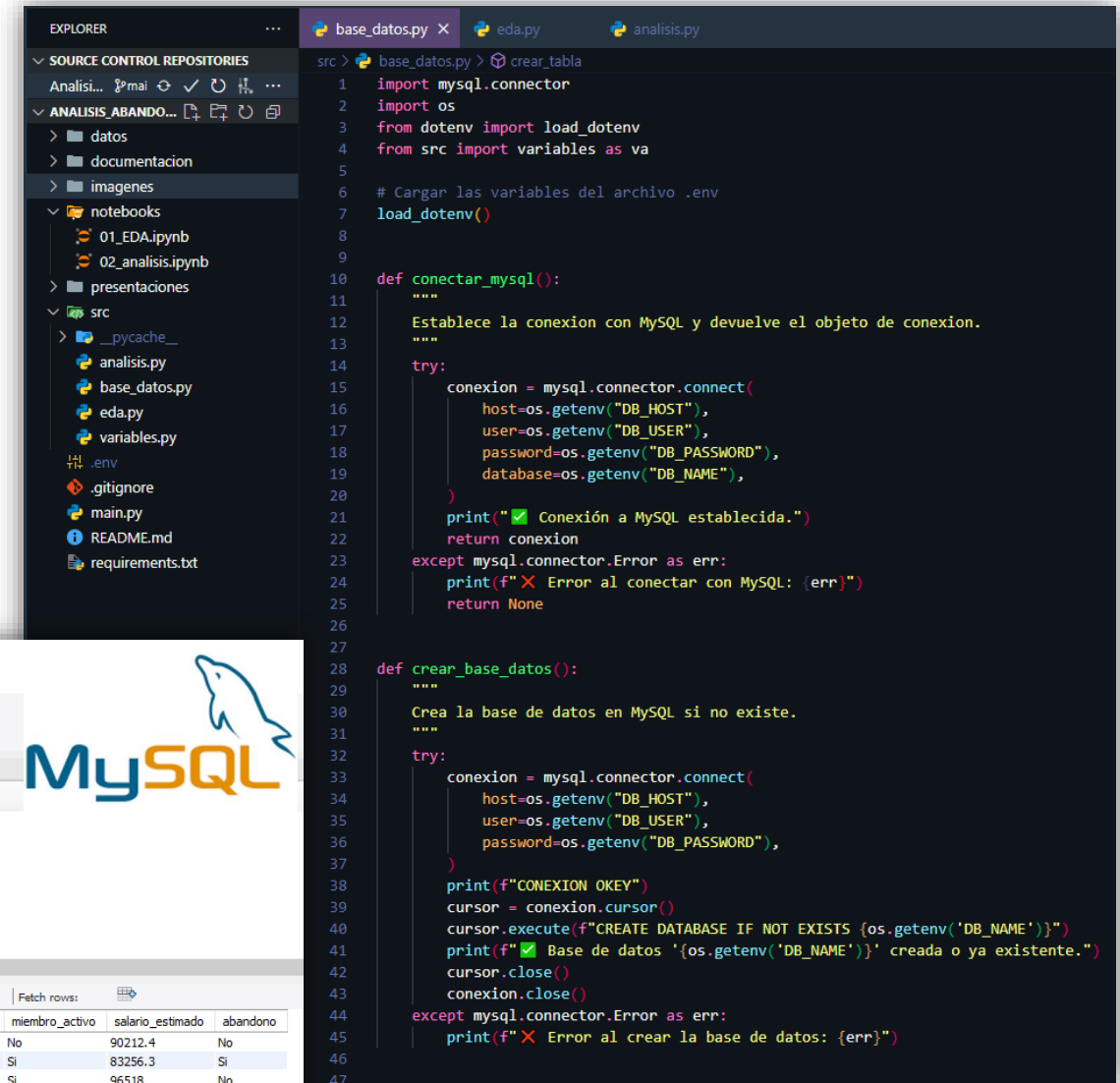


MySQL Workbench interface showing a query and its results. The query is:

```
1 USE bluebank;
2
3 SELECT *
4 FROM BlueBank_Clientes
5 LIMIT 5;
```

The result grid shows the following data:

id_cliente	puntacion_credito	pais	genero	edad	antiguedad	saldo	num_productos	tarjeta_credito	miembro_activo	salario_estimado	abandono
15565701	698	España	Mujer	39	9	161994	1	No	No	90212.4	No
15565706	612	España	Hombre	35	1	0	1	Si	Si	83256.3	Si
15565714	601	Francia	Hombre	47	1	64430.1	2	No	Si	96518	No
15565779	627	Alemania	Mujer	30	6	57809.3	1	Si	No	188258	No
15565796	745	Alemania	Hombre	48	10	96048.5	1	Si	No	74510.6	No



```
EXPLORER
...
base_datos.py x eda.py analisis.py

SOURCE CONTROL REPOSITORIES
Analisi...
ANALISIS_ABANDO...
  > datos
  > documentacion
  > imagenes
  > notebooks
    01_EDA.ipynb
    02_analisis.ipynb
  > presentaciones
  > src
    > _pycache_
    analisis.py
    base_datos.py
    eda.py
    variables.py
  .env
  .gitignore
  main.py
  README.md
  requirements.txt

src > base_datos.py > crear_tabla
1 import mysql.connector
2 import os
3 from dotenv import load_dotenv
4 from src import variables as va
5
6 # Cargar las variables del archivo .env
7 load_dotenv()
8
9
10 def conectar_mysql():
11     """
12     Establece la conexión con MySQL y devuelve el objeto de conexión.
13     """
14     try:
15         conexion = mysql.connector.connect(
16             host=os.getenv("DB_HOST"),
17             user=os.getenv("DB_USER"),
18             password=os.getenv("DB_PASSWORD"),
19             database=os.getenv("DB_NAME"),
20         )
21         print("✅ Conexión a MySQL establecida.")
22         return conexion
23     except mysql.connector.Error as err:
24         print(f"❌ Error al conectar con MySQL: {err}")
25         return None
26
27
28 def crear_base_datos():
29     """
30     Crea la base de datos en MySQL si no existe.
31     """
32     try:
33         conexion = mysql.connector.connect(
34             host=os.getenv("DB_HOST"),
35             user=os.getenv("DB_USER"),
36             password=os.getenv("DB_PASSWORD"),
37         )
38         print(f"CONEXION OKEY")
39         cursor = conexion.cursor()
40         cursor.execute(f"CREATE DATABASE IF NOT EXISTS {os.getenv('DB_NAME')}")
41         print(f"✅ Base de datos '{os.getenv('DB_NAME')}' creada o ya existente.")
42         cursor.close()
43         conexion.close()
44     except mysql.connector.Error as err:
45         print(f"❌ Error al crear la base de datos: {err}")
46
47
```

Visualización: Power BI

- Creación de **dashboards** con **filtros interactivos**.
- Aplicación de **lenguaje DAX** para **cálculos personalizados**.
- Uso de **gráficos visuales efectivos**.

Introducción

Tecnología

Flujo
análisis

Conclusiones

Cierre



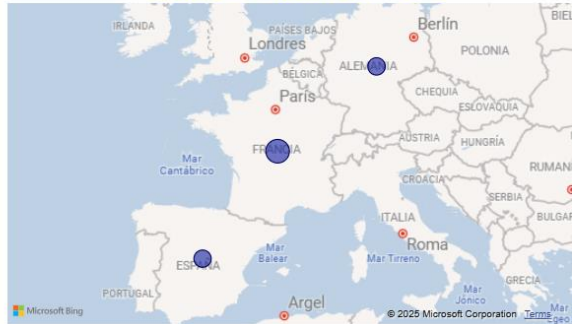
Análisis clientes BlueBank



Perfil general



Número de clientes por país

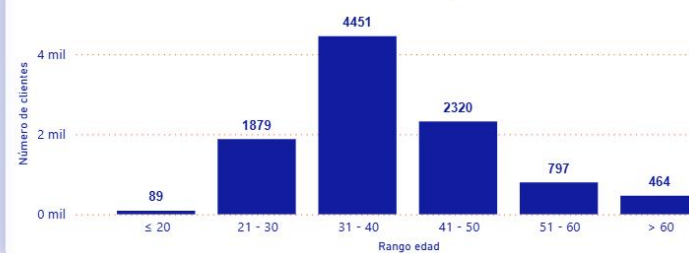


% Activos
51,51

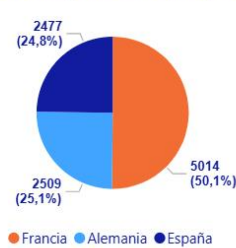
Total clientes
10000

% Abandono
20,37

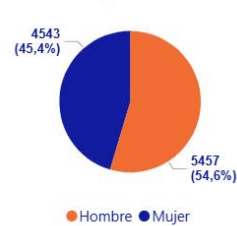
Distribución de clientes por edad



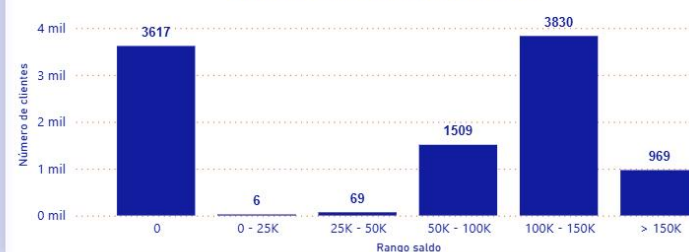
Porcentaje clientes por país

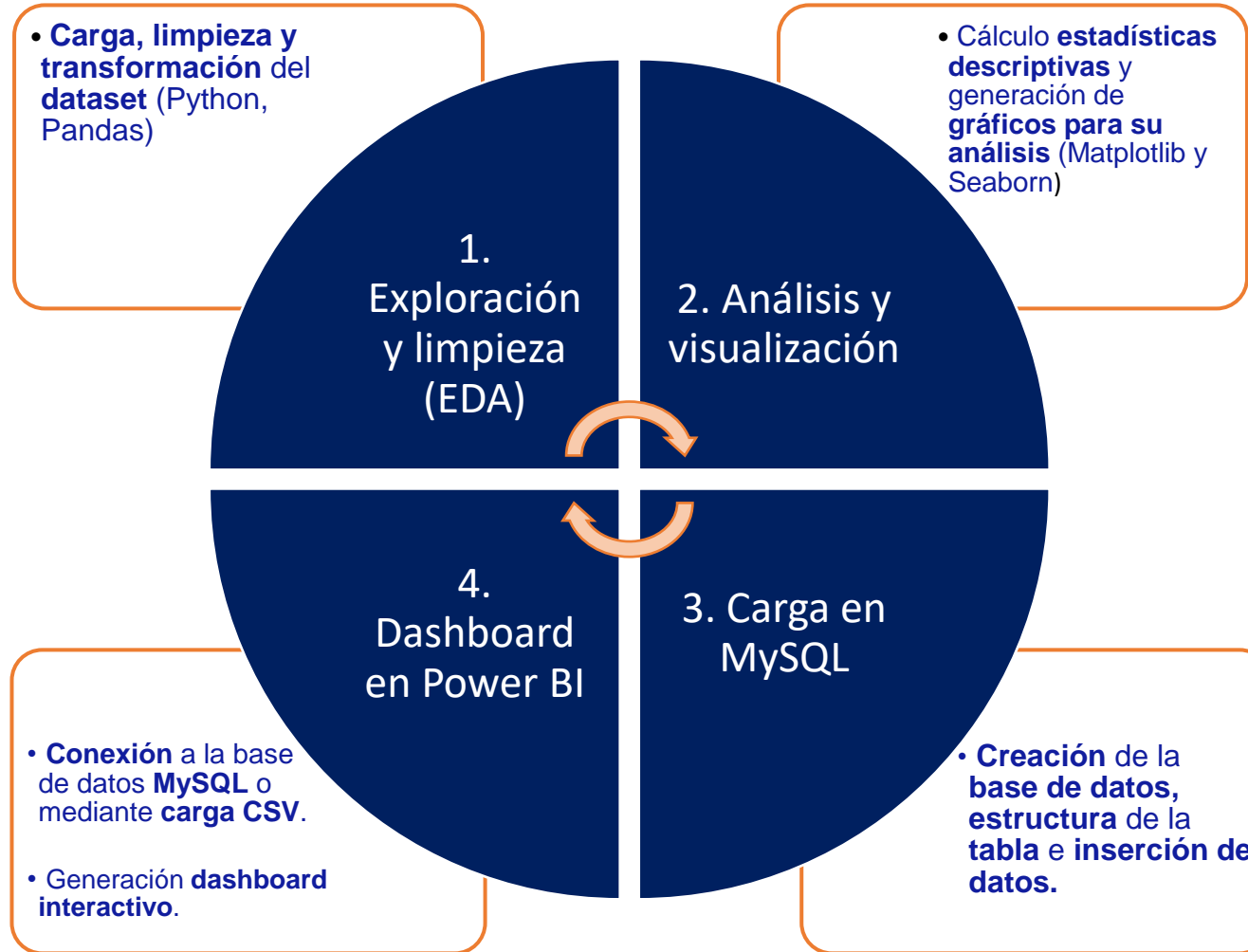


Porcentaje clientes por género



Distribución de clientes por saldo

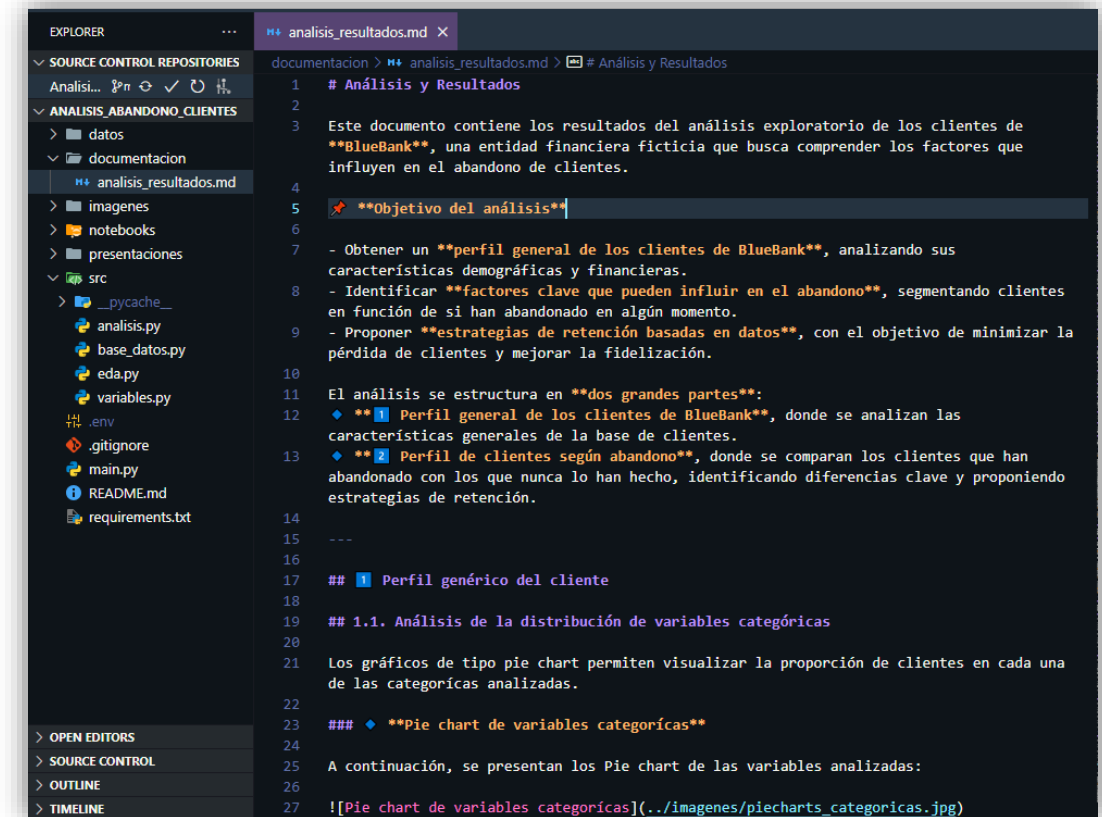






Conclusiones

- **Análisis del cliente general.**
- **Análisis del cliente que abandonó el banco, pero ha vuelto.**
- **Factores más relevantes en el abandono.**
- **Conclusiones y estrategias de retención.**
- **Análisis completo con gráficos y métricas en *análisis_resultados.md***
- **Next Steps: automatización de consultas SQL definidas, almacenamiento de resultados en archivo txt, visualización de las consultas en la terminal y creación de la base de datos en la nube.**
- **Continuación en Power BI, archivo ubicado en el repositorio del proyecto.**



```
1 # Análisis y Resultados
2
3 Este documento contiene los resultados del análisis exploratorio de los clientes de
4 **BlueBank**, una entidad financiera ficticia que busca comprender los factores que
5 influyen en el abandono de clientes.
6
7 **Objetivo del análisis**
8
9 - Obtener un **perfil general de los clientes de BlueBank**, analizando sus
10 características demográficas y financieras.
11 - Identificar **factores clave que pueden influir en el abandono**, segmentando clientes
12 en función de si han abandonado en algún momento.
13 - Proponer **estrategias de retención basadas en datos**, con el objetivo de minimizar la
14 pérdida de clientes y mejorar la fidelización.
15
16 El análisis se estructura en **dos grandes partes**:
17
18 1. **Perfil general de los clientes de BlueBank**, donde se analizan las
19 características generales de la base de clientes.
20
21 2. **Perfil de clientes según abandono**, donde se comparan los clientes que han
22 abandonado con los que nunca lo han hecho, identificando diferencias clave y proponiendo
23 estrategias de retención.
24
25 ---
26
27 **1. Perfil genérico del cliente**
28
29 **1.1. Análisis de la distribución de variables categóricas**
30
31 Los gráficos de tipo pie chart permiten visualizar la proporción de clientes en cada una
32 de las categorías analizadas.
33
34 **Pie chart de variables categóricas**
35
36 A continuación, se presentan los Pie chart de las variables analizadas:
37
38 **Pie chart de variables categóricas**(../imagenes/piecharts_categoricas.jpg)
```




Links de interés

Repositorio del
proyecto

Repositorio
consultas SQL

GitHuB

LinkedIN

Correo: m.penacasares@gmail.com

