

UNIVERSIDAD INTERNACIONAL
DE LAS AMÉRICAS
ESCUELA DE INGENIERÍA INFORMÁTICA

PROYECTO

CURSO

Base de Datos 2

Nombre de (los) estudiante (s)

Brayan Chavarria Arias

Joshua Navarro Corrales

PROFESOR

Olman Nuñez Peralta

SAN JOSÉ, COSTA RICA

- **REQ01: Identificar al menos una fuente pública**

Kaggle fue el medio utilizado para obtener datos publicos para el EDA y el modelo multidimensional. Además Kaggle al ser una pagina con fuentes de datos publicos como de paga hay múltiples fuentes interesantes. Nuestra base de datos llamada Conjunto de datos de predicción de abandono de clientes bancarios.

(Badole, 2024)

- **REQ02: Plantear tres preguntas con al menos tres variables cada pregunta, justificando su relevancia y su objetivo.**

¿Cuál es el perfil demográfico de los clientes que tienden a abandonar el banco?

Género: Masculino, Femenino.

Edad: Rango de edades.

Fuera Del Banco: Dentro o fuera del banco en su totalidad.

Justificación: Identificar el perfil demográfico de los clientes que abandonan el banco puede ayudar a comprender si ciertos grupos tienen una mayor tendencia a hacerlo, lo que permitiría diseñar estrategias de retención específicas.

Objetivo: Descubrir patrones demográficos entre los clientes que abandonan el banco para desarrollar campañas de retención personalizadas y reducir la tasa de abandono

¿Cómo influye el balance de la cuenta en la permanencia o salida de los clientes?

- **Variables:**

1. **Saldo** Cantidad de dinero disponible en la cuenta.
2. **Edad:** Rango de edades.
3. **Fuera Del Banco:** Dentro o fuera del banco en su totalidad.

- **Justificación:** El balance de la cuenta puede ser un indicador importante de la estabilidad financiera y la relación del cliente con el banco. Comprender cómo varía el balance entre los clientes que permanecen y los que abandonan puede ayudar a identificar riesgos financieros.

- **Objetivo:** Analizar si existe una correlación entre el Saldo de la cuenta y la probabilidad de abandono, influyen en esta relación.

¿Qué impacto tiene la posesión de una tarjeta de crédito en la retención de clientes?

- **Variables:**
 1. **Tarjeta de crédito:** Si el cliente tiene o no tarjeta de crédito.
 2. **Género:** Masculino, Femenino.
 3. **Fuera Del Banco:** Dentro o fuera del banco en su totalidad.
- **Justificación:** La posesión de una tarjeta de crédito puede estar relacionada con la lealtad del cliente y su satisfacción con los servicios bancarios. Analizar esta relación puede proporcionar información valiosa sobre cómo mejorar la retención de clientes.
- **Objetivo:** Evaluar si tener una tarjeta de crédito está asociado con una mayor o menor probabilidad de abandono.

• REQ03: Entendimiento de los datos: cuál es su origen, como se recolectaron, si son susceptibles de fallos, entre otros (para todas las fuentes de datos)

- **Origen:** Base de datos interna del banco, almacena información sobre los clientes, sus transacciones, productos contratados (como tarjetas de crédito), y otros detalles relevantes como el tiempo que llevan en el banco, sus saldos, etc.
- **Recolección:** La mayoría de los datos son recolectados automáticamente a través de sistemas de transacciones, donde cada acción del cliente (como un depósito, retiro o uso de tarjeta) se registra en tiempo real.
- **Susceptibilidad a Fallos:** Datos incompletos, Errores de Registro, Actualización de Datos son algunos de los posibles fallos dentro de la base de datos.

•**REQ04: Preparación de los datos: clasificación de las variables cuantitativas (discretas o continuas) y cualitativas (ordinales o nominales)**

Variables:

Las variables cuantitativas son aquellas que se pueden medir numéricamente.

- **Continuas:** Pueden tomar cualquier valor dentro de un rango continuo.
 - **Saldo:** Es una variable continua porque el saldo de una cuenta puede ser cualquier valor dentro de un rango (por ejemplo, \$100.25, \$100.50, etc.).

- **Discretas:** Pueden tomar valores específicos y separados.
 - **Edad:** Es una variable discreta porque la edad generalmente se mide en años enteros (20, 21, 22, etc.).

Las variables cualitativas son aquellas que describen categorías o cualidades.

- **Nominales:** No tienen un orden intrínseco.
 - **Género:** Es una variable nominal porque las categorías (Masculino, Femenino) no tienen un orden natural.
 - **Tarjeta de crédito (Sí o No):** Es nominal porque las categorías no tienen un orden.

- **REQ05: Entendimiento de cada una de las variables: análisis estadístico descriptivo sobre estas (promedio, mediana, desviación estándar)**

```
# Mostrar descripción completa  
print(datos.describe())
```

	Cal_Crediticia	Tiempo_Cliente	Saldo
## count	10000.000000	10000.000000	10000.000000
## mean	650.528800	5.012800	76485.889288
## std	96.653299	2.892174	62397.405202
## min	350.000000	0.000000	0.000000
## 25%	584.000000	3.000000	0.000000
## 50%	652.000000	5.000000	97198.540000
## 75%	718.000000	7.000000	127644.240000
## max	850.000000	10.000000	250898.090000

Cal_Crediticia:

- **Count (10000):** Hay 10,000 registros de calificaciones crediticias en la base de datos.
- **Mean (650.53):** La calificación crediticia promedio de los clientes es de 650.53.
- **Std (96.65):** La desviación estándar es 96.65, lo que indica una variabilidad moderada en las calificaciones crediticias de los clientes.
- **Min (350):** La calificación crediticia mínima es 350, que puede corresponder a un cliente con un historial crediticio más bajo.
- **25% (584):** El 25% de los clientes tiene una calificación crediticia menor o igual a 584, lo que indica que este grupo de clientes tiene un perfil de crédito más bajo.
- **50% (652):** La mediana de la calificación crediticia es 652, lo que significa que el 50% de los clientes tiene una calificación crediticia menor o igual a este valor.
- **75% (718):** El 75% de los clientes tiene una calificación crediticia menor o igual a 718, lo que indica que solo el 25% de los clientes tiene calificaciones más altas.
- **Max (850):** La calificación crediticia máxima es 850, que podría indicar un cliente con un excelente historial crediticio.

Tiempo_Cliente:

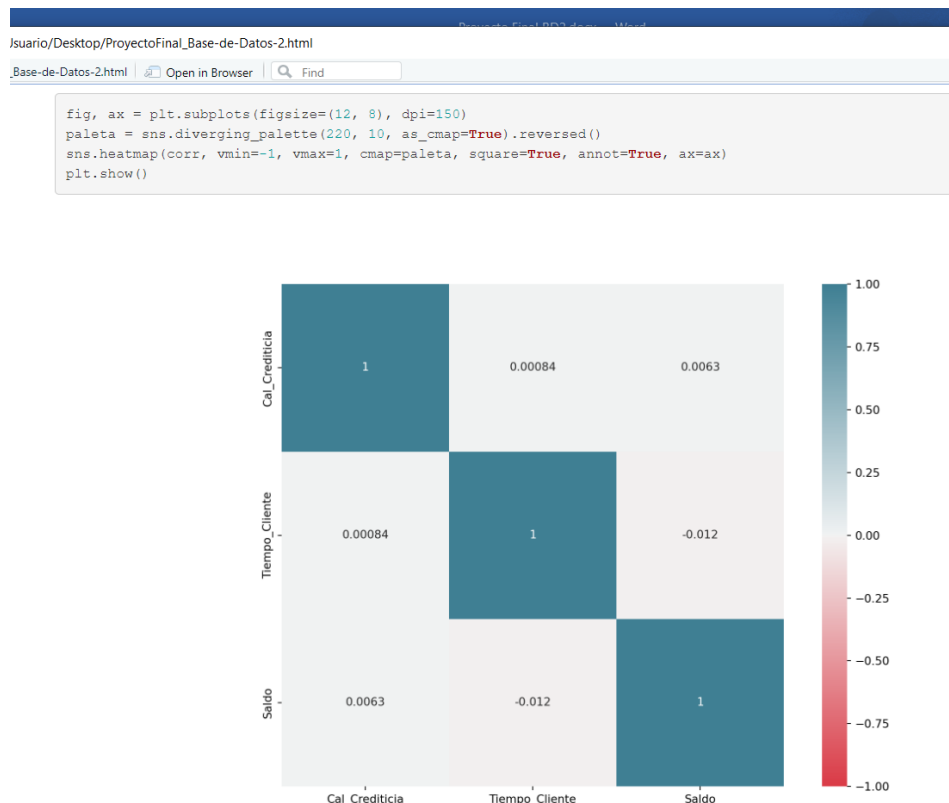
- **Count (10000):** Hay 10,000 registros del tiempo que los clientes han estado con el banco.
- **Mean (5.01):** En promedio, los clientes han estado con el banco por aproximadamente 5 años.
- **Std (2.89):** La desviación estándar es 2.89 años, lo que sugiere una variabilidad considerable en la duración de la relación de los clientes con el banco.
- **Min (0):** El tiempo mínimo que un cliente ha estado con el banco es 0 años, lo que podría representar a clientes nuevos o aquellos que acaban de cerrar su cuenta.
- **25% (3):** El 25% de los clientes ha estado con el banco por 3 años o menos.
- **50% (5):** La mediana del tiempo que los clientes han estado con el banco es de 5 años.
- **75% (7):** El 75% de los clientes ha estado con el banco por 7 años o menos, lo que indica que solo el 25% ha estado con el banco por más tiempo.
- **Max (10):** El tiempo máximo que un cliente ha estado con el banco es de 10 años, representando a los clientes más antiguos.

Saldo:

- **Count (10000):** Hay 10,000 registros de saldo de clientes.
- **Mean (76485.89):** El saldo promedio en las cuentas de los clientes es de aproximadamente 76,486 unidades monetarias.
- **Std (62397.41):** La desviación estándar del saldo es bastante alta (62,397.41), lo que indica una gran variabilidad en los saldos de los clientes.
- **Min (0):** El saldo mínimo es 0, lo que significa que algunos clientes no tienen dinero en sus cuentas.
- **25% (0):** El 25% de los clientes tiene un saldo de 0, lo que indica que una cuarta parte de los clientes no tiene dinero en sus cuentas.
- **50% (97,198.54):** La mediana del saldo es 97,198.54, lo que sugiere que el 50% de los clientes tiene menos de esta cantidad en sus cuentas.

- **75% (127,644.24):** El 75% de los clientes tiene un saldo menor o igual a 127,644.24 unidades monetarias, lo que sugiere que el 25% restante tiene saldos significativamente mayores.
- **Max (250,898.09):** El saldo máximo es 250,898.09, lo que representa a los clientes con los saldos más altos en el banco.

• **REQ06: Estudio de las relaciones entre variables tanto cualitativas (usando variables dummies) como cuantitativas: Análisis de correlaciones.**



cal_crediticia vs saldo (0.0063):

- Este valor indica una correlación muy débil y positiva entre la calificación crediticia y el saldo. Esto sugiere que no hay una relación significativa entre estas dos variables.

cal_crediticia vs tiempo_cliente (0.00084):

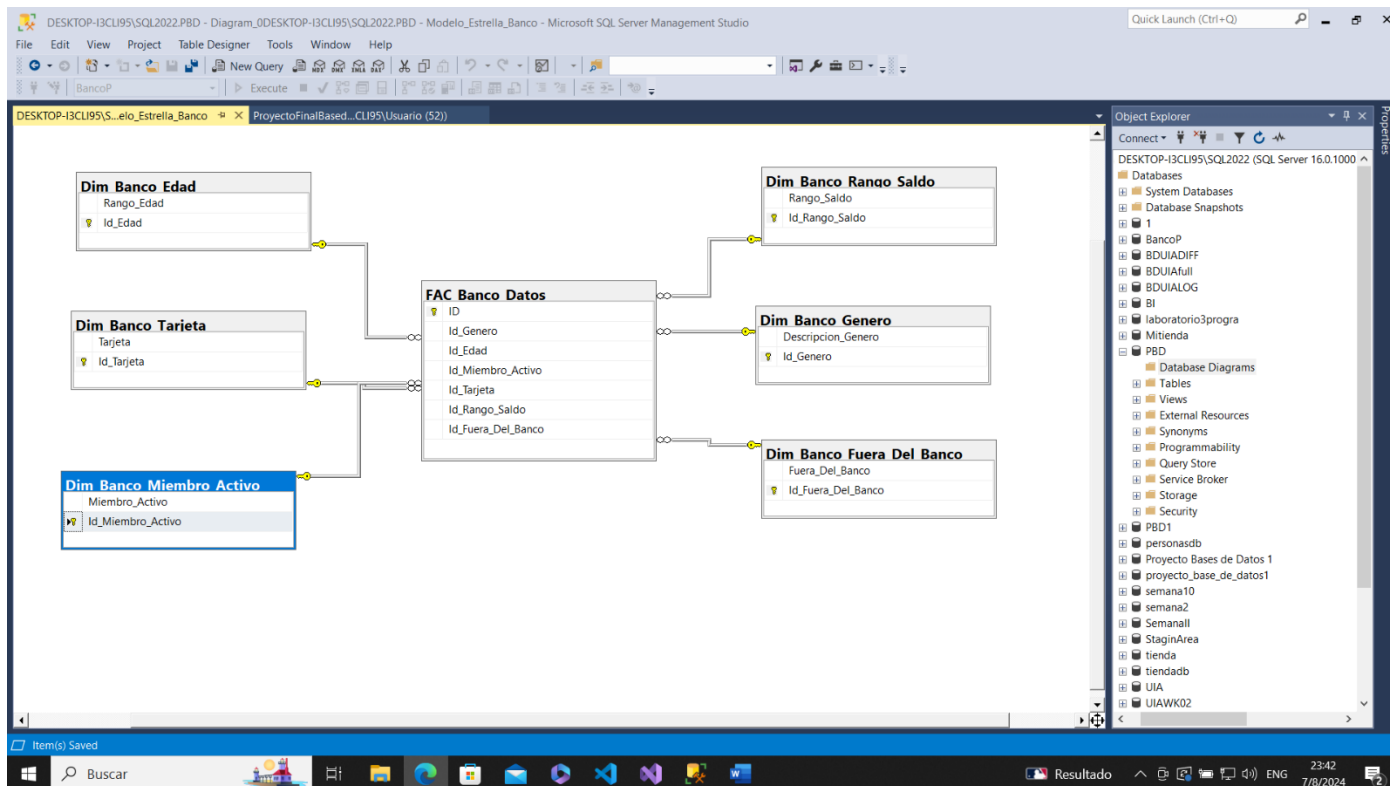
- Esta correlación también es extremadamente débil y positiva. Indica que prácticamente no existe una relación entre la calificación crediticia y el tiempo que un cliente ha estado con el banco.

tiempo_cliente vs saldo (-0.012):

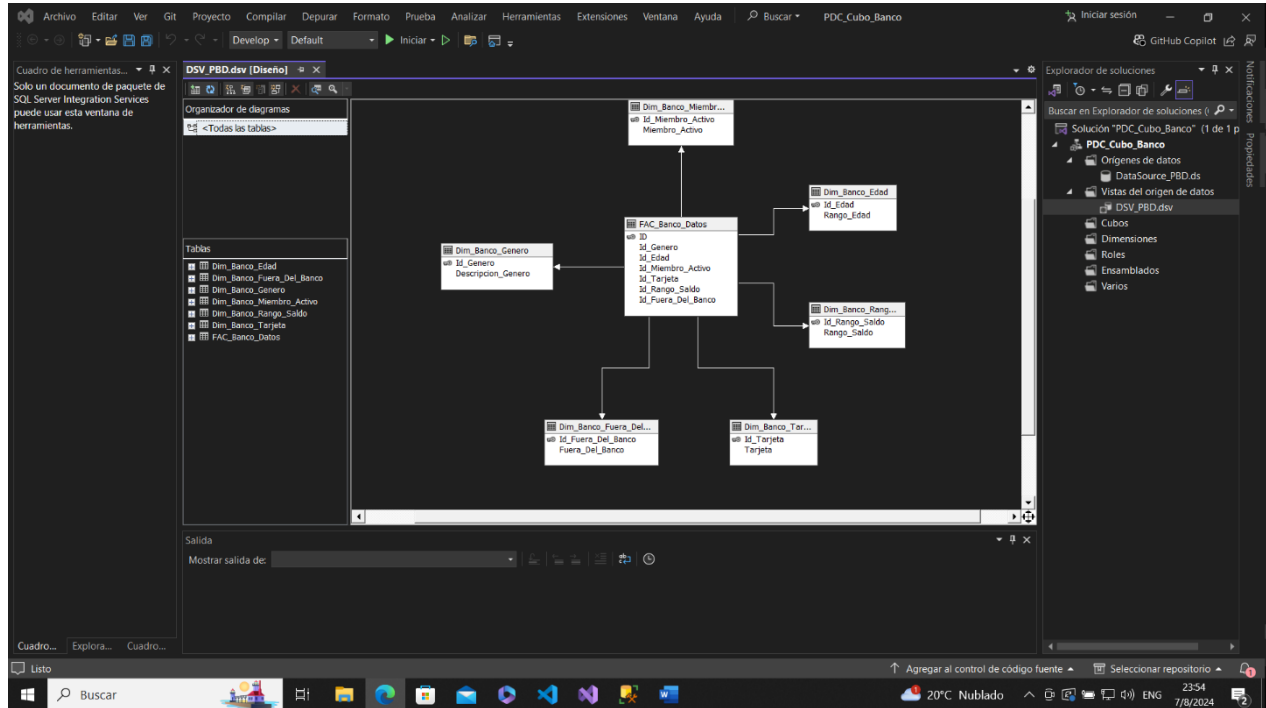
- Aquí se observa una correlación muy débil y negativa. Aunque es negativa, la relación es tan pequeña que no indica una tendencia clara entre el tiempo que un cliente ha estado con el banco y su saldo.

Como los valores están muy cerca de 0, significa que las variables no están significativamente correlacionadas. Esto implica que, en este caso, los cambios en una variable no están asociados con cambios en la otra.

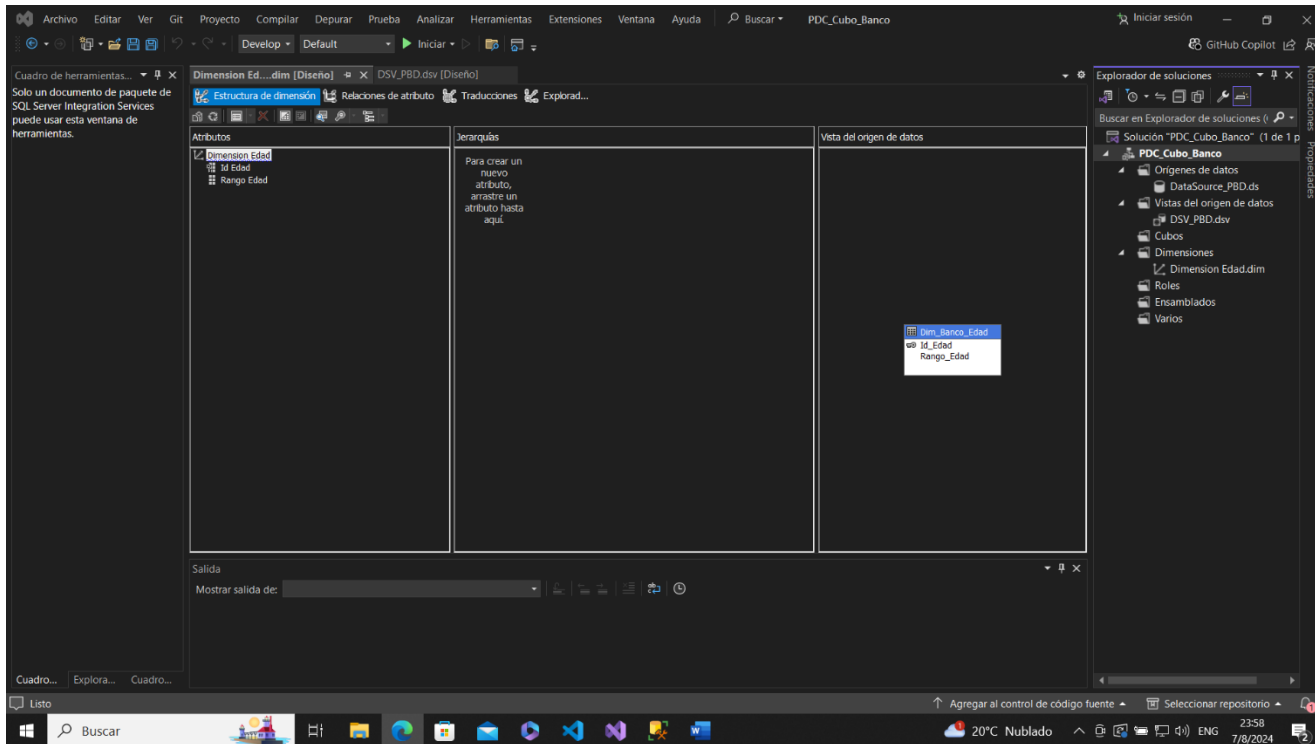
- **REQ07: Generación de las tablas de dimensiones (al menos 5)**



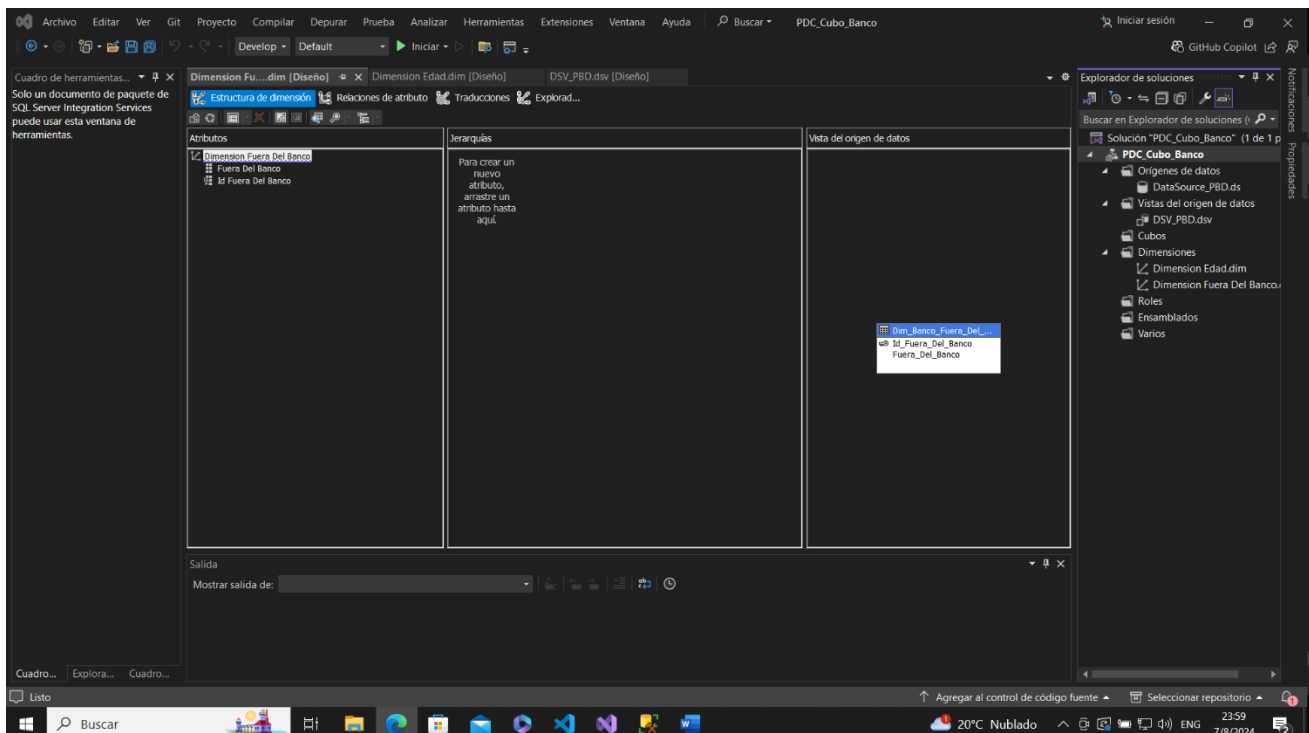
- REQ08: Generación de la tabla de hechos (al menos 1)



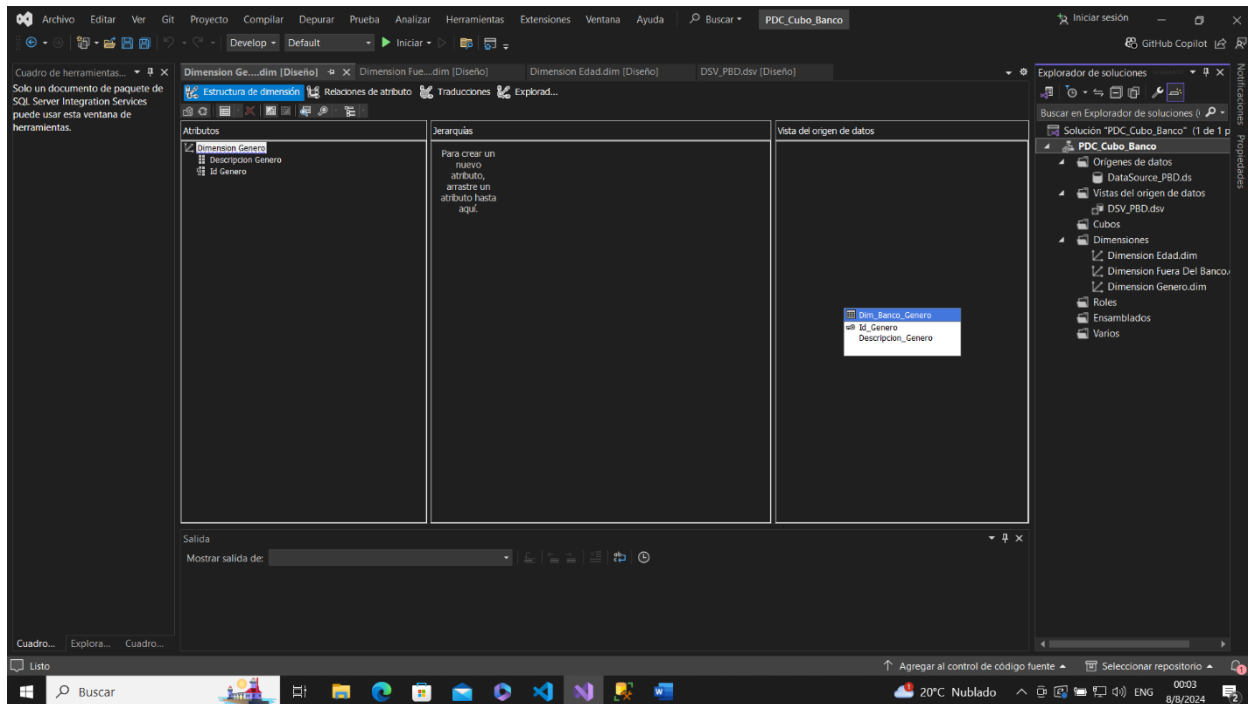
Creación de dimensión Edad



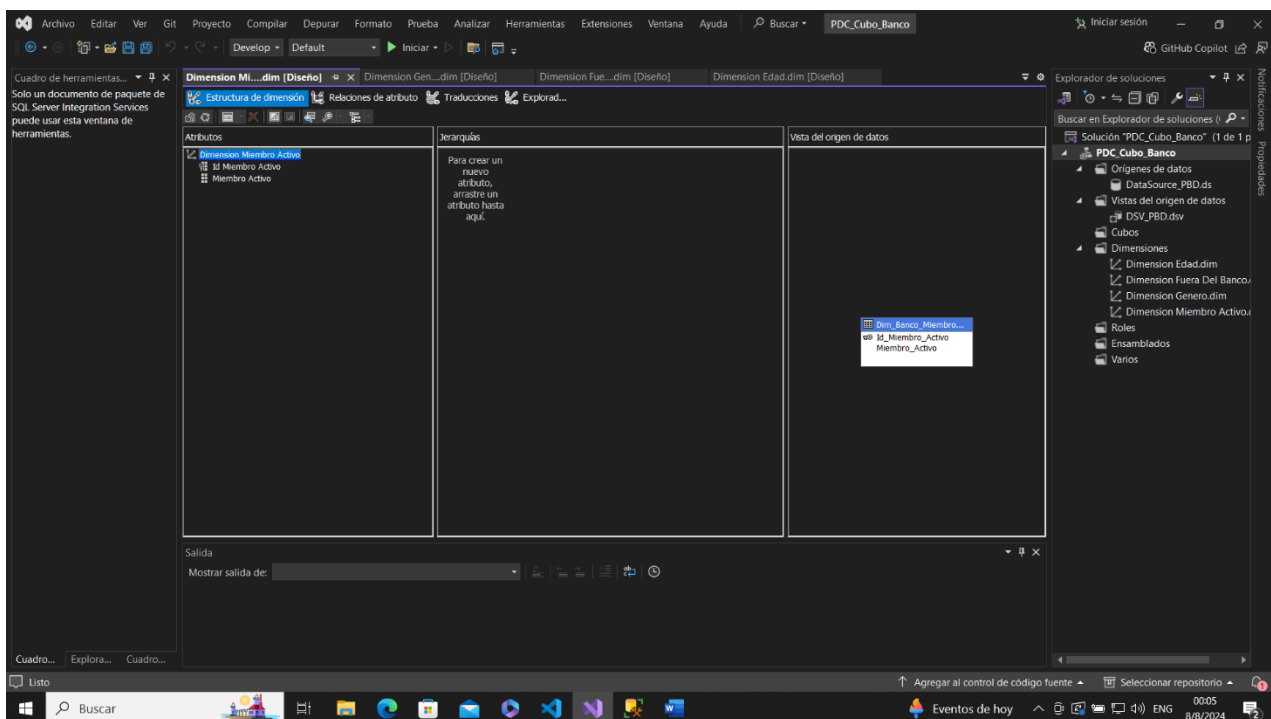
Creación de dimensión Fuera del banco



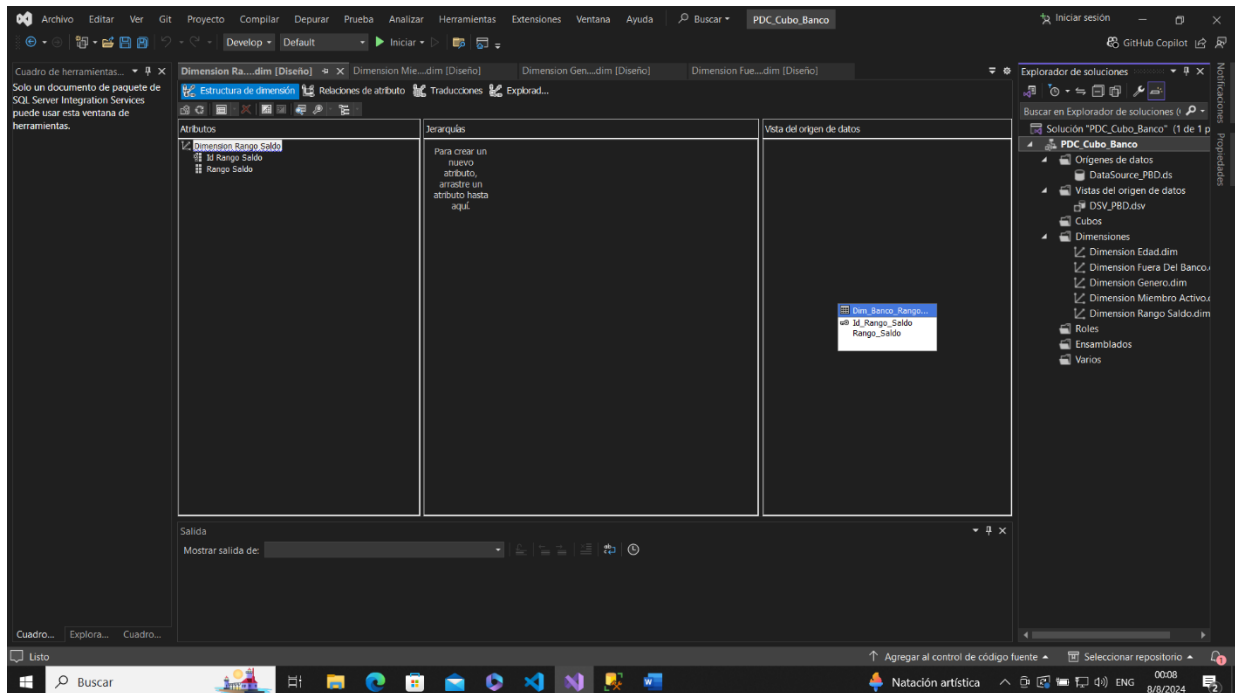
Creación de dimensión Genero



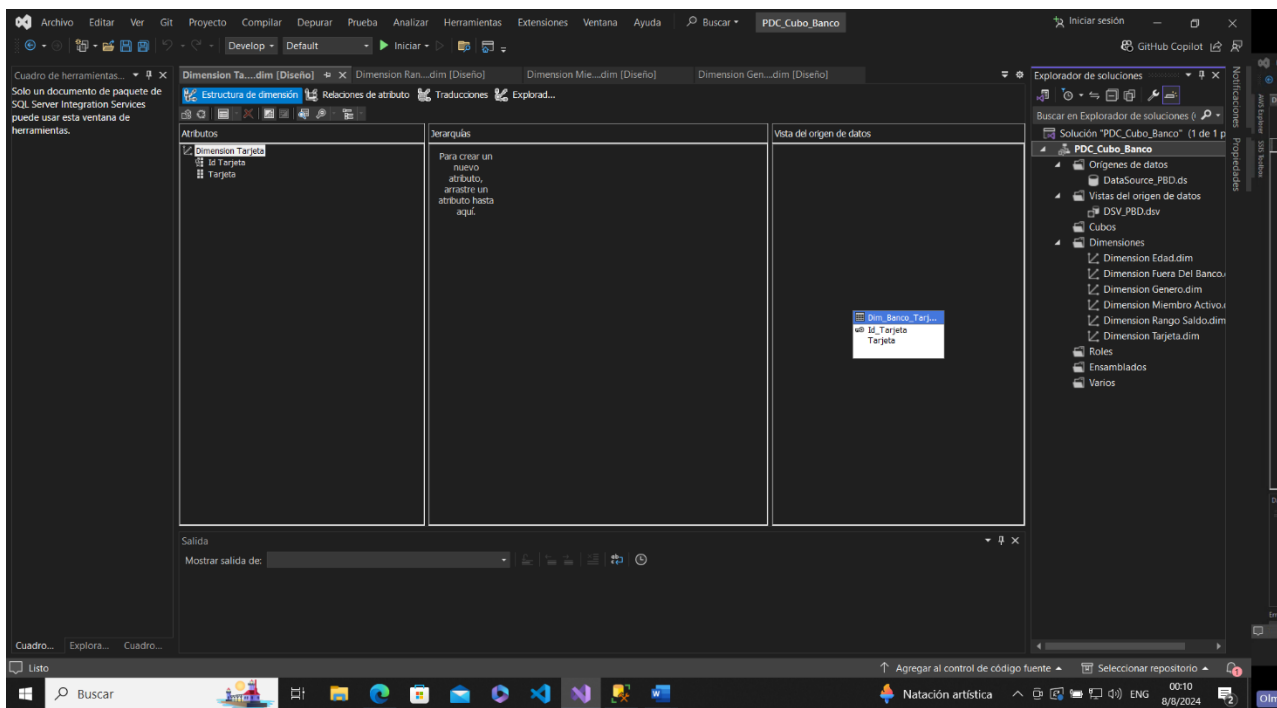
Creación de dimensión Miembro Activo



Creación de Dimensión Rango Saldo

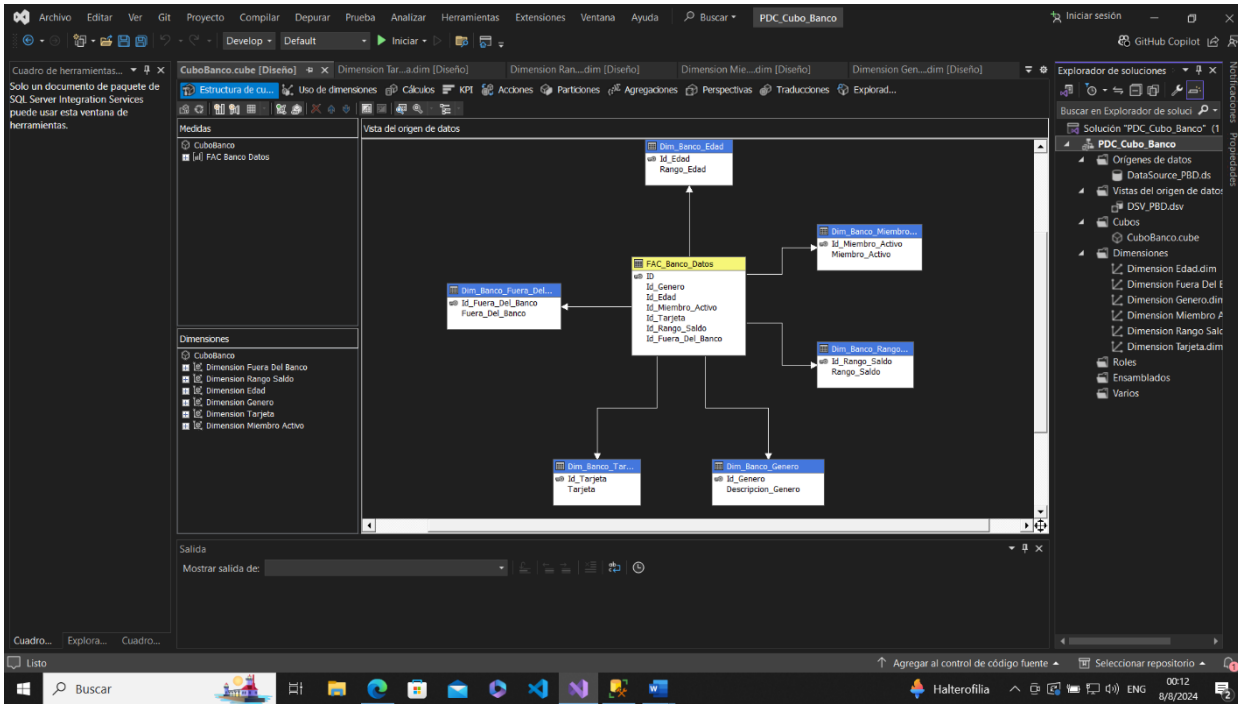


Creación de Dimensión Tarjeta

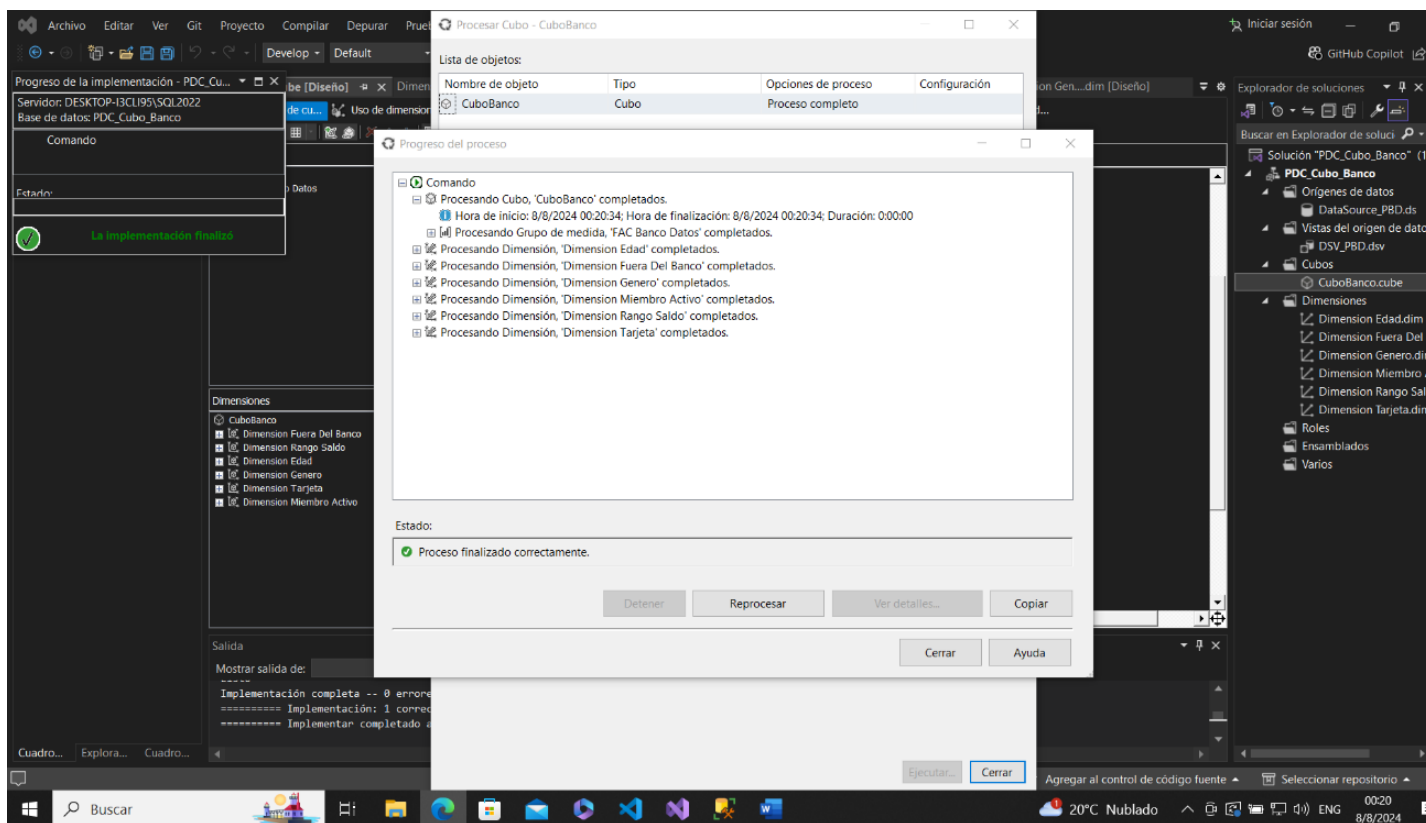
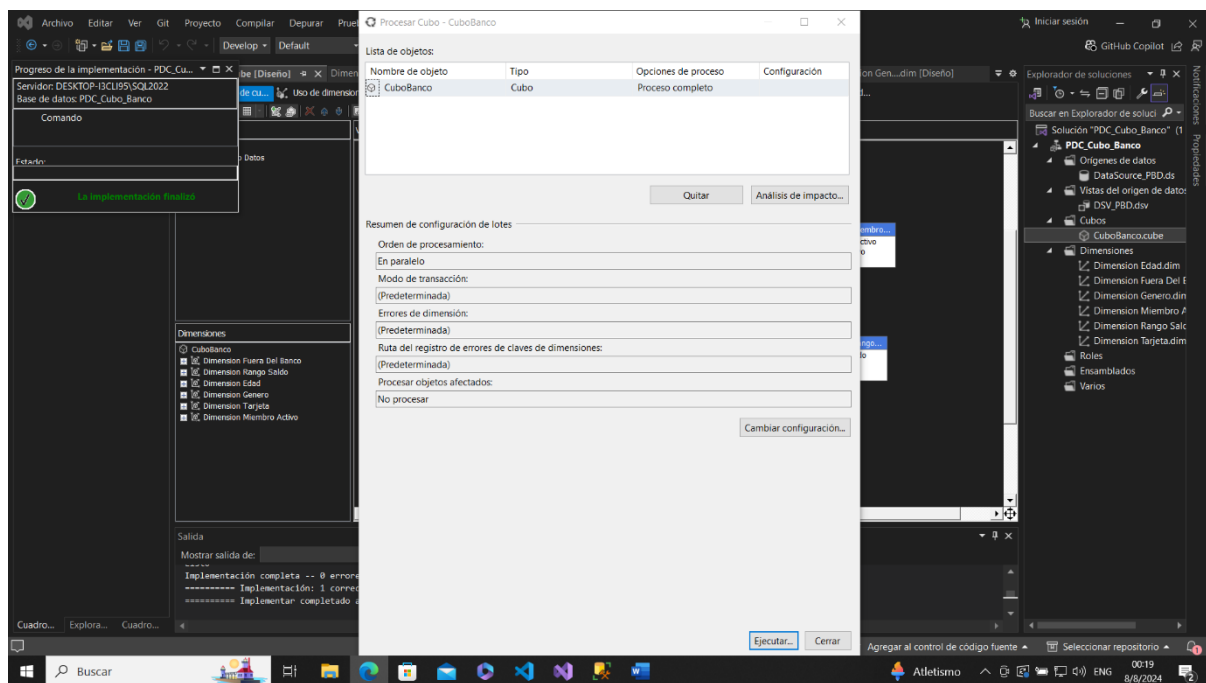


- **REQ09: Generación del cubo multidimensional**

Creacion del Cubo llamado CuboBanco



Mensaje de que el Deployment fue exitoso



- REQ10: Presentación de los resultados obtenidos: Para el EDA se debe utilizar RMarkdown

Datos, Variables

```
rio/Desktop/ProyectoFinal_Base-de-Datos-2.html
-r-de-Datos-2.html Open in Browser Find Publi
```

```
print(datos.head())
```

```
##           Apellido Cal_Crediticia  Genero  Edad  Tiempo_Cliente \
## ID
## 15634602 Hargrave          619.0 Femenino 41-50          2
## 15647311 Hill            608.0 Femenino 41-50          1
## 15619304 Onio            502.0 Femenino 41-50          8
## 15701354 Boni            699.0 Femenino 31-40          1
## 15737888 Mitchell        850.0 Femenino 41-50          2
##
##           Saldo Tarjeta Miembro_Activo  Salario Fuera_Del_Banco \
## ID
## 15634602      0.00      Sí            Sí 101348.88          Sí
## 15647311 83807.86      No            Sí 112542.58          No
## 15619304 159660.80      Sí            No 113931.57          Sí
## 15701354      0.00      No            No 93826.63          No
## 15737888 125510.82      Sí            Sí 79084.1          No
##
##           RangoSaldo
## ID
## 15634602      0 - 50179.62
## 15647311 50179.62 - 100359.24
## 15619304 150538.86 - 200718.48
## 15701354      0 - 50179.62
## 15737888 100359.24 - 150538.86
```

“Discribe “estadístico sobre estas (promedio, mediana, desviación estándar)

```
# Mostrar descripción completa
print(datos.describe())
```

```
##           Cal_Crediticia  Tiempo_Cliente          Saldo
## count      10000.000000      10000.000000  10000.000000
## mean         650.528800         5.012800    76485.889288
## std          96.653299         2.892174    62397.405202
## min          350.000000         0.000000     0.000000
## 25%          584.000000         3.000000     0.000000
## 50%          652.000000         5.000000    97198.540000
## 75%          718.000000         7.000000   127644.240000
## max          850.000000        10.000000   250898.090000
```

```
numeric_data = datos.select_dtypes(include=[np.number])
print("Datos numéricos:")
```

```
## Datos numéricos:
```

```
print(numeric_data.head())
```

```
##           Cal_Crediticia  Tiempo_Cliente          Saldo
## ID
## 15634602          619.0          2          0.00
## 15647311          608.0          1   83807.86
## 15619304          502.0          8  159660.80
## 15701354          699.0          1          0.00
## 15737888          850.0          2  125510.82
```

Datos Numericos

```
C:/Users/Usuario/Desktop/ProyectoFinal_Base-de-Datos-2.html
oyectoFinal_Base-de-Datos-2.html | Open in Browser | Find | Publish

## Datos numéricos:

print(numeric_data.head())

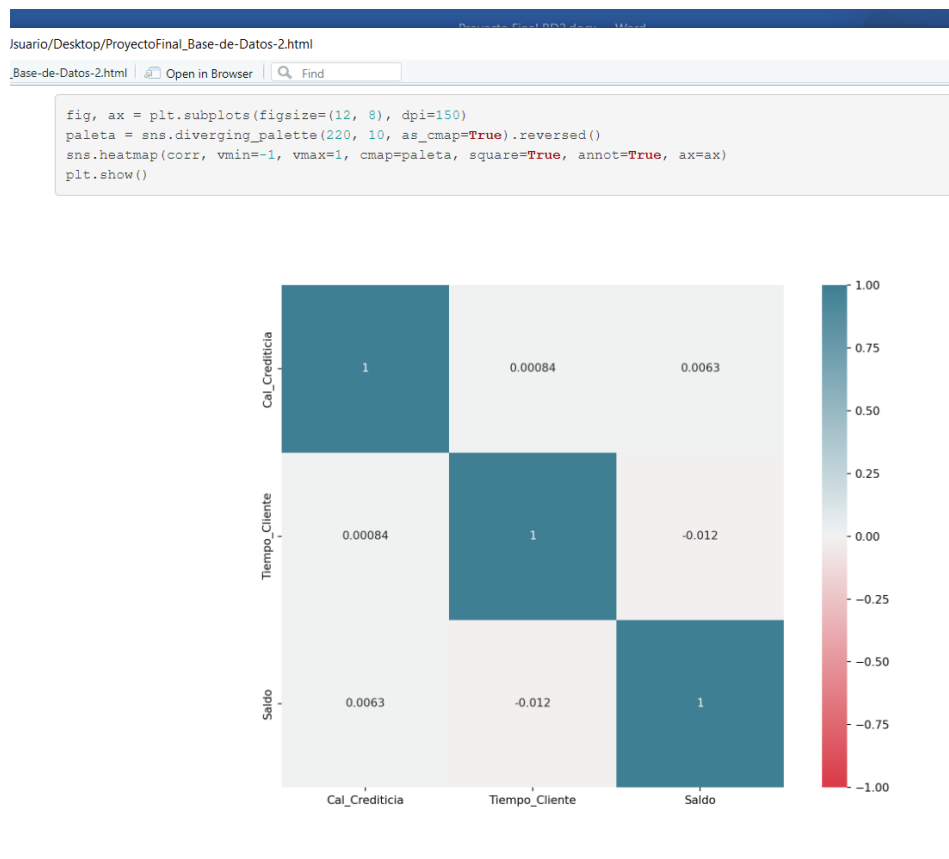
##           Cal_Crediticia  Tiempo_Cliente      Saldo
## ID
## 15634602             619.0             2         0.00
## 15647311             608.0             1    83807.86
## 15619304             502.0             8   159660.80
## 15701354             699.0             1         0.00
## 15737888             850.0             2   125510.82

corr = numeric_data.corr()
print(corr)

##           Cal_Crediticia  Tiempo_Cliente      Saldo
## Cal_Crediticia      1.000000      0.000842  0.006268
## Tiempo_Cliente      0.000842      1.000000 -0.012254
## Saldo               0.006268     -0.012254  1.000000

fig, ax = plt.subplots(figsize=(12, 8), dpi=150)
paleta = sns.diverging_palette(220, 10, as_cmap=True).reversed()
sns.heatmap(corr, vmin=-1, vmax=1, cmap=paleta, square=True, annot=True, ax=ax)
plt.show()
```

gráfico de calor y las correlaciones entre las variables



cal_crediticia vs saldo (0.0063):

- Este valor indica una correlación muy débil y positiva entre la calificación crediticia y el saldo. Esto sugiere que no hay una relación significativa entre estas dos variables.

cal_crediticia vs tiempo_cliente (0.00084):

- Esta correlación también es extremadamente débil y positiva. Indica que prácticamente no existe una relación entre la calificación crediticia y el tiempo que un cliente ha estado con el banco.

tiempo_cliente vs saldo (-0.012):

- Aquí se observa una correlación muy débil y negativa. Aunque es negativa, la relación es tan pequeña que no indica una tendencia clara entre el tiempo que un cliente ha estado con el banco y su saldo.

Como los valores están muy cerca de 0, significa que las variables no están significativamente correlacionadas. Esto implica que, en este caso, los cambios en una variable no están asociados con cambios en la otra.

Datos escalados

```
datos_scaled = pd.DataFrame(StandardScaler().fit_transform(numeric_data), columns=numeric_data.columns, index=numeric_data.index)
print("Datos escalados:")
```

```
## Datos escalados:
```

```
print(datos_scaled)
```

```
##           Cal_Crediticia  Tiempo_Cliente      Saldo
## ID
## 15634602      -0.326221      -1.041760 -1.225848
## 15647311      -0.440036      -1.387538  0.117350
## 15619304      -1.536794       1.032908  1.333053
## 15701354       0.501521      -1.387538 -1.225848
## 15737888       2.063884      -1.041760  0.785728
## ...           ...           ...           ...
## 15606229       1.246488      -0.004426 -1.225848
## 15569892      -1.391939       1.724464 -0.306379
## 15584532       0.604988       0.687130 -1.225848
## 15682355       1.256835      -0.695982 -0.022608
## 15628319       1.463771      -0.350204  0.859965
##
## [10000 rows x 3 columns]
```

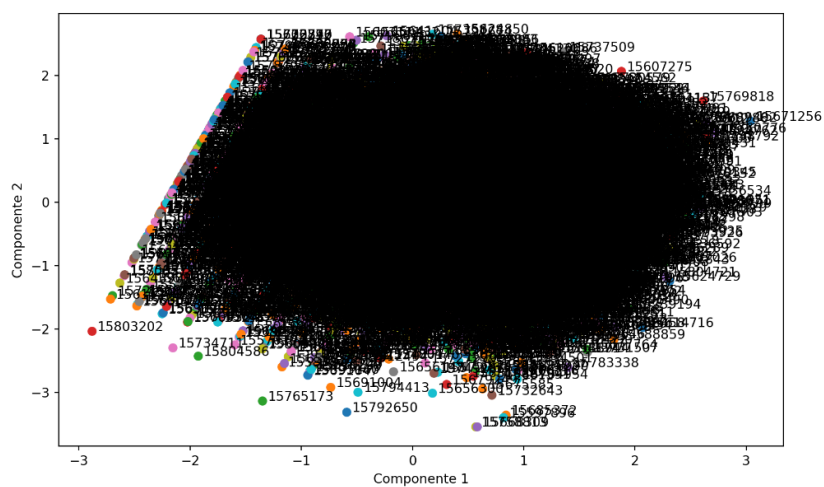
Componentes

[de-Datos-2.html](#)

Browser Find

```
pca_sintransponser = PCA(n_components=2)
componentes = pca_sintransponser.fit_transform(datos_scaled)

fig, ax = plt.subplots(nrows=1, ncols=1, figsize=(10, 6), dpi=200)
ax.set_xlabel("Componente 1")
ax.set_ylabel("Componente 2")
for i in range(componentes.shape[0]):
    x = componentes[i, 0]
    y = componentes[i, 1]
    ax.scatter(x, y)
    ax.annotate(text=datos_scaled.index[i], xy=(x + 0.05, y))
plt.show()
```



¿Cuál es el perfil demográfico de los clientes que tienden a abandonar el banco?

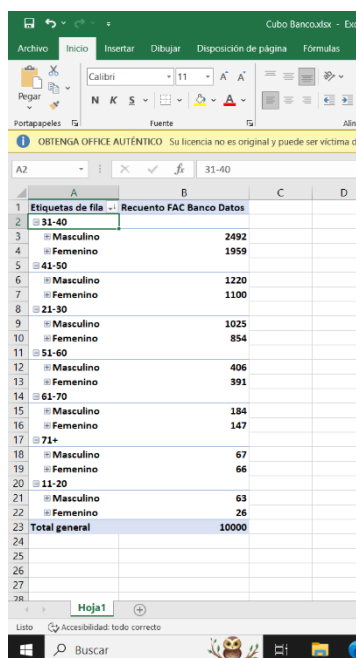
Género: Masculino, Femenino.

Edad: Rango de edades.

Fuera Del Banco: Dentro o fuera del banco en su totalidad.

Justificación: Identificar el perfil demográfico de los clientes que abandonan el banco puede ayudar a comprender si ciertos grupos tienen una mayor tendencia a hacerlo, lo que permitiría diseñar estrategias de retención específicas.

Objetivo: Descubrir patrones demográficos entre los clientes que abandonan el banco para desarrollar campañas de retención personalizadas y reducir la tasa de abandono



Cubo Bancario - Exo

Archivo Inicio Insertar Dibujar Disposición de página Fórmulas

Calibri 11 A A

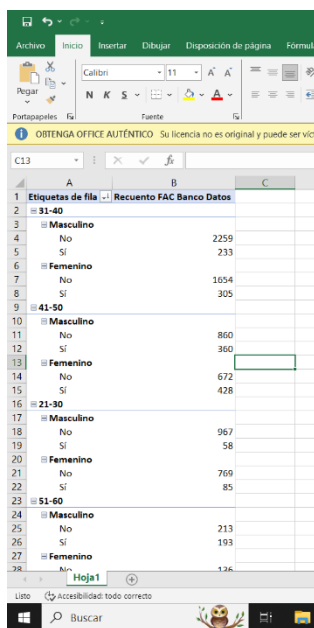
Portapapeles Pegar Fuente Alinea

OBTENGA OFFICE AUTÉNTICO Su licencia no es original y puede ser víctima de

	A	B	C	D
1	Etiquetas de fila	Recuento FAC Banco Datos		
2	31-40			
3	Masculino	2492		
4	Femenino	1959		
5	41-50			
6	Masculino	1220		
7	Femenino	1100		
8	21-30			
9	Masculino	1025		
10	Femenino	854		
11	51-60			
12	Masculino	406		
13	Femenino	391		
14	61-70			
15	Masculino	184		
16	Femenino	147		
17	71+			
18	Masculino	67		
19	Femenino	66		
20	11-20			
21	Masculino	63		
22	Femenino	26		
23	Total general	10000		
24				
25				
26				
27				
28				

Hoja1

Buscar



Cubo Bancario - Exo

Archivo Inicio Insertar Dibujar Disposición de página Fórmulas

Calibri 11 A A

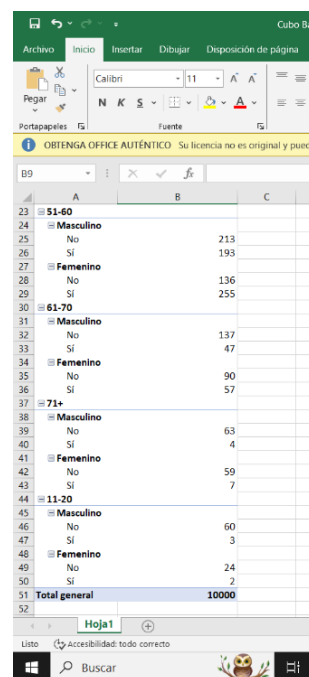
Portapapeles Pegar Fuente Alinea

OBTENGA OFFICE AUTÉNTICO Su licencia no es original y puede ser víctima de

	A	B	C
1	Etiquetas de fila	Recuento FAC Banco Datos	
2	31-40		
3	Masculino		
4	No	2259	
5	Si	233	
6	Femenino		
7	No	1654	
8	Si	305	
9	41-50		
10	Masculino		
11	No	860	
12	Si	360	
13	Femenino		
14	No	672	
15	Si	428	
16	21-30		
17	Masculino		
18	No	967	
19	Si	58	
20	Femenino		
21	No	769	
22	Si	85	
23	51-60		
24	Masculino		
25	No	213	
26	Si	193	
27	Femenino		
28	No	136	

Hoja1

Buscar



Cubo Bancario - Exo

Archivo Inicio Insertar Dibujar Disposición de página Fórmulas

Calibri 11 A A

Portapapeles Pegar Fuente Alinea

OBTENGA OFFICE AUTÉNTICO Su licencia no es original y puede ser víctima de

	A	B	C
23	51-60		
24	Masculino		
25	No	213	
26	Si	193	
27	Femenino		
28	No	136	
29	Si	255	
30	61-70		
31	Masculino		
32	No	137	
33	Si	47	
34	Femenino		
35	No	90	
36	Si	57	
37	71+		
38	Masculino		
39	No	63	
40	Si	4	
41	Femenino		
42	No	59	
43	Si	7	
44	11-20		
45	Masculino		
46	No	60	
47	Si	3	
48	Femenino		
49	No	24	
50	Si	2	
51	Total general	10000	
52			

Hoja1

Buscar

	A	B	C
1	Etiquetas de fila	Recuento FAC Banco Datos	
2	Masculino		
3	No	4559	
4	Sí	898	
5	Femenino		
6	No	3404	
7	Sí	1139	
8	Total general	10000	
9			
10			
11			

Respuesta:

Las Personas entre 31-40 años son las que más abandonan el banco, curiosamente, en cada rango de edades las mujeres lideran este fenómeno, lo que indica que ellas son las que más salen de la entidad financiera sin importar su edad, lo que genera una gran inquietud.

Plan de Retención: Para abordar esta situación, se podría desarrollar un programa de beneficios personalizado para clientas femeninas, especialmente aquellas en el rango de edad identificado. Esto podría incluir:

1. **Ofertas y Promociones Especiales:** Crear productos financieros diseñados para mujeres, como cuentas de ahorro con tasas de interés competitivas, préstamos con condiciones favorables para emprendimientos femeninos, o seguros con cobertura extendida para ellas y sus familias.
2. **Asesoría Financiera Personalizada:** Implementar servicios de consultoría financiera que se centren en las metas y desafíos específicos que enfrentan las mujeres en diferentes etapas de su vida, como la planificación de la jubilación, la educación de los hijos, o la gestión de deudas.
3. **Fidelización y Reconocimiento:** Desarrollar un programa de fidelización que premie la permanencia de las clientas, ofreciendo puntos, descuentos o beneficios exclusivos por mantener sus cuentas activas y realizar transacciones regulares.

¿Cómo influye el balance de la cuenta en la permanencia o salida de los clientes?

- **Variables:**
 1. **Saldo** Cantidad de dinero disponible en la cuenta.
 2. **Edad:** Rango de edades.
 3. **Fuera Del Banco:** Dentro o fuera del banco en su totalidad.
- **Justificación:** El balance de la cuenta puede ser un indicador importante de la estabilidad financiera y la relación del cliente con el banco. Comprender cómo varía el balance entre los clientes que permanecen y los que abandonan puede ayudar a identificar riesgos financieros.
- **Objetivo:** Analizar si existe una correlación entre el Saldo de la cuenta y la probabilidad de abandono, influyen en esta relación.

Total, de Abandonos: 2037 personas

Abandonos totales en rango de 100.359 a 150.538: 987 personas

A	B	C	D
Etiquetas de fila	Recuento FAC Banco Datos		
100359.24 - 150538.86			
31-40			
No	1429		
Si	262		
41-50			
No	534		
Si	378		
21-30			
No	597		
Si	76		
51-60			
No	129		
Si	207		
61-70			
No	80		
Si	57		
71+			
No	41		
Si	5		
11-20			
No	31		
Si	2		
0 - 50179.62	3692		
50179.62 - 100359.24	1533		
150538.86 - 200718.48	915		
200718.48 - 250898.09	32		
Total general	10000		

A	B	C	D	E
Etiquetas de fila	Recuento FAC Banco Datos			
No	7963			
Si	2037			
Total general	10000			

la mayoría de los clientes que han abandonado el banco se encuentran en el rango de saldo de 100,359 a 150,538, representando casi la mitad de los abandonos totales. Este hallazgo sugiere que, para este grupo de clientes, la cantidad de dinero en sus cuentas no es el principal problema, sino posiblemente la falta de incentivos atractivos que justifiquen su permanencia en el banco.

Plan de Retención: Para retener a estos clientes, es esencial diseñar estrategias que no solo mejoren las condiciones financieras, sino que también fortalezcan su compromiso con el banco. Algunas propuestas incluyen:

1. **Mejora de Tasas de Interés:** Introducir tasas de interés más competitivas para las cuentas de ahorro en este rango de saldo, permitiendo que los clientes vean un crecimiento tangible en sus ahorros. Ofrecer productos de inversión con un bajo riesgo pero con rendimientos atractivos podría también ser una opción para mantener a estos clientes interesados en dejar su dinero en el banco.
2. **Beneficios Exclusivos:** Implementar un programa de beneficios exclusivo para este segmento, que podría incluir descuentos en servicios financieros, acceso a líneas de crédito preferenciales, o la posibilidad de participar en sorteos y promociones.

especiales. Este tipo de incentivos podría hacer que los clientes perciban un mayor valor al mantener su saldo dentro de este rango.

¿Qué impacto tiene la posesión de una tarjeta de crédito en la retención de clientes?

- **Variables:**
 1. **Tarjeta de crédito:** Si el cliente tiene o no tarjeta de crédito.
 2. **Género:** Masculino, Femenino.
 3. **Fuera Del Banco:** Dentro o fuera del banco en su totalidad.
- **Justificación:** La posesión de una tarjeta de crédito puede estar relacionada con la lealtad del cliente y su satisfacción con los servicios bancarios. Analizar esta relación puede proporcionar información valiosa sobre cómo mejorar la retención de clientes.
- **Objetivo:** Evaluar si tener una tarjeta de crédito está asociado con una mayor o menor probabilidad de abandono.

	A	B	C	D	E	F	G
1	Etiquetas de fila	Recuento FAC Banco Datos					
2	Masculino						
3	No						
4	No	1325					
5	Sí	269					
6	Sí			Fuera del Banco Con Tarjeta	1424		
7	No	3234		Fuera del Banco sin Tarjeta	613		
8	Sí	629					
9	Femenino			Dentro del Banco Con Tarjeta	5631		
10	No			Dentro del Banco sin Tarjeta	2332		
11	No	1007					
12	Sí	344					
13	Sí						
14	No	2397					
15	Sí	795					
16	Total general	10000					
17							
18							
19							
20							

El análisis indica una tendencia preocupante, de las 2,037 personas que han abandonado el banco, 1,424 poseían una tarjeta, lo que sugiere que tener una tarjeta por sí sola no es

suficiente para garantizar la permanencia de los clientes. Esto resalta la necesidad de mejorar la percepción de valor asociada a la tarjeta para los clientes actuales y futuros.

Plan de Retención:

Para abordar este desafío, es crucial desarrollar un plan integral que no solo incentive a los clientes actuales a mantener su tarjeta, sino que también motive a aquellos que aún no la tienen a solicitar una. Algunas estrategias a considerar incluyen:

1. Revisión y Mejora de Beneficios de la Tarjeta:

- **Recompensas Personalizadas:** Introducir programas de recompensas más personalizados y atractivos que reflejen los patrones de gasto y preferencias de los clientes. Esto podría incluir puntos que se acumulen por cada transacción y que puedan ser canjeados por descuentos en compras, viajes, entretenimiento, entre otros.
- **Cashback Atractivo:** Implementar un sistema de cashback con tasas atractivas en categorías específicas, como compras diarias, combustible o restaurantes, donde los clientes puedan ver un retorno inmediato en sus gastos.

Conclusión

El análisis de los clientes que han abandonado el banco revela patrones preocupantes, especialmente entre ciertos grupos demográficos y segmentos financieros. Las mujeres entre 31 y 40 años destacan como las más propensas a abandonar la entidad, sugiriendo un descontento que trasciende la edad y se manifiesta a través de un número significativo de salidas. Asimismo, un considerable grupo de clientes con saldos elevados también opta por dejar el banco, lo que sugiere que, a pesar de tener una situación financiera aparentemente estable, estos clientes no encuentran incentivos suficientes para permanecer.

Finalmente, la posesión de una tarjeta de crédito, aunque común entre los clientes, no se traduce en una lealtad a la entidad, indicando que la propuesta de valor de las tarjetas no está alineada con las expectativas de los usuarios.

Recomendación:

Para mitigar la tasa de abandono y fortalecer la retención, es crucial implementar estrategias diferenciadas y específicas para los grupos identificados:

1. **Programa de beneficios personalizados para mujeres:** Dada la alta tasa de abandono entre las clientas, especialmente en el rango de edad de 31-40 años, se recomienda diseñar un programa de fidelización enfocado en este grupo. Este programa podría incluir beneficios exclusivos, asesoramiento financiero personalizado, y opciones de productos financieros adaptados a sus necesidades.
2. **Incentivos y programas de fidelización para clientes con saldos elevados:** Considerando que muchos clientes con altos saldos abandonan el banco, es esencial ofrecerles beneficios adicionales que vayan más allá de las condiciones financieras estándar. Esto podría incluir tasas preferenciales, acceso a servicios exclusivos, y programas de inversión atractivos.
3. **Revalorizar la oferta de tarjetas de crédito:** Dado que la posesión de una tarjeta no garantiza la lealtad del cliente, es necesario repensar el valor percibido de estas. Se podrían introducir programas de recompensas más atractivos, descuentos exclusivos en comercios seleccionados, y servicios adicionales vinculados a la tarjeta que hagan más atractiva su tenencia.

Referencias

Badole, S. (Mayo de 2024). *Kaggle*. Obtenido de <https://www.kaggle.com/datasets/saurabhbadole/bank-customer-churn-prediction-dataset>