

Group Name : Solo

Name: Vishnu Priya Malchetti

Email: [vishnupriyam.sapbasis@gmail.com](mailto:vishnupriyam.sapbasis@gmail.com)

Country: Ireland

Specialization: Data Analyst

***Problem description:***

We are determined to help XYZ Bank improve its cross-selling strategies and enhance customer engagement. The bank offers a wide array of financial products and services, including savings accounts, credit cards, mortgages, loans, and investment options. However, we've observed that many of our customers have limited product adoption and aren't fully utilizing the range of services available to them. To tackle this challenge head-on, we plan to implement customer segmentation techniques to gain deeper insights into our customer base. By dividing our customers into distinct groups based on their demographics, financial behavior, and product usage patterns, we hope to identify specific customer segments that are more likely to use products and services. Armed with this valuable information, we aim to create personalized marketing strategies and tailored cross-selling initiatives to boost customer satisfaction and encourage higher product adoption. As part of our data analysis team, the objective is to thoroughly analyze the extensive customer dataset provided by XYZ Bank and conduct a comprehensive customer segmentation analysis. The dataset includes detailed information about each customer, such as age, gender, income, transaction history, product holdings, and tenure with our bank.

***Data Cleansing***

Missing values in the training dataset

```
[55]: #missing values checking
df1.isnull().sum()
```

```
[55]: fecha_datos      0
      ncodpers        0
      ind_empleado    27734
      pais_residencia 27734
      sexo            27804
      age             0
      fecha_alta      27734
      ind_nuevo       27734
      antiguedad      0
      indrel          27734
      ult_fec_cli_1t   13622516
      indrel_1mes     149781
      tiprel_1mes     149781
      indresi         27734
      indext          27734
      conyuemp        13645501
      canal_entrada   186126
      indfall         27734
```

Using mean()

```
memory usage: 712160
[60]: df1['indrel']=df1['indrel'].fillna(df1['indrel'].mean())
[61]: df1['tipodom']=df1['tipodom'].fillna(df1['tipodom'].mean())
[62]: df1['cod_prov']=df1['cod_prov'].fillna(df1['cod_prov'].mean())
[64]: df1['ind_actividad_cliente']=df1['ind_actividad_cliente'].fillna(df1['ind_actividad_cliente'].mean())
[65]: df1['renta']=df1['renta'].fillna(df1['renta'].mean())
[68]: df1['ind_nomina_ult1']=df1['ind_nomina_ult1'].fillna(df1['ind_nomina_ult1'].mean())
[70]: df1['ind_nom_pens_ult1']=df1['ind_nom_pens_ult1'].fillna(df1['ind_nom_pens_ult1'].mean())
[71]: df1.isnull().sum()

[71]: fecha_datos      0
      ncodpers        0
      ind_empleado    27734
      pais_residencia 27734
      sexo            27804
      age             0
      fecha_alta      27734
```

```
[75]: df_new=df1.dropna()
```

```
[76]: df_new.isnull().sum()
```

```
[76]: fecha_dato      0
      ncodpers      0
      ind_empleado  0
      pais_residencia 0
      sexo          0
      age          0
      fecha_alta    0
      ind_nuevo     0
      antiguedad    0
      indrel        0
      ult_fec_cli_1t 0
      indrel_1mes   0
      tiprel_1mes   0
      indresi       0
      indext        0
      conyuemp      0
      canal_entrada 0
      indfall       0
      tipodom       0
      cod_prov      0
      nomprov       0
      ind_actividad_cliente 0
```