

ZAL_PG_churn

December 9, 2023

1 Bank Customer Churn Model

The objective of this model is to predict the customers who may leave the bank

```
[ ]: from google.colab import drive
drive.mount('/content/drive/')
```

Drive already mounted at /content/drive/; to attempt to forcibly remount, call drive.mount("/content/drive/", force_remount=True).

```
[ ]: ! pip install ydata-profiling
```

Collecting ydata-profiling

Downloading ydata_profiling-4.6.3-py2.py3-none-any.whl (357 kB)
357.6/357.6

kB 2.4 MB/s eta 0:00:00

Requirement already satisfied: scipy<1.12,>=1.4.1 in
/usr/local/lib/python3.10/dist-packages (from ydata-profiling) (1.11.4)

Requirement already satisfied: pandas!=1.4.0,<3,>1.1 in
/usr/local/lib/python3.10/dist-packages (from ydata-profiling) (1.5.3)

Requirement already satisfied: matplotlib<3.9,>=3.2 in
/usr/local/lib/python3.10/dist-packages (from ydata-profiling) (3.7.1)

Collecting pydantic>=2 (from ydata-profiling)

Downloading pydantic-2.5.2-py3-none-any.whl (381 kB)
381.9/381.9

kB 14.9 MB/s eta 0:00:00

Requirement already satisfied: PyYAML<6.1,>=5.0.0 in
/usr/local/lib/python3.10/dist-packages (from ydata-profiling) (6.0.1)

Requirement already satisfied: jinja2<3.2,>=2.11.1 in
/usr/local/lib/python3.10/dist-packages (from ydata-profiling) (3.1.2)

Collecting visions[type_image_path]==0.7.5 (from ydata-profiling)

Downloading visions-0.7.5-py3-none-any.whl (102 kB)
102.7/102.7

kB 5.9 MB/s eta 0:00:00

Requirement already satisfied: numpy<1.26,>=1.16.0 in

/usr/local/lib/python3.10/dist-packages (from ydata-profiling) (1.23.5)

Collecting htmlmin==0.1.12 (from ydata-profiling)

Downloading htmlmin-0.1.12.tar.gz (19 kB)
Preparing metadata (setup.py) ... done
Collecting phik<0.13,>=0.11.1 (from ydata-profiling)
Downloading
phik-0.12.3-cp310-cp310-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (679 kB)
679.5/679.5

kB 37.0 MB/s eta 0:00:00

Requirement already satisfied: requests<3,>=2.24.0 in
/usr/local/lib/python3.10/dist-packages (from ydata-profiling) (2.31.0)
Requirement already satisfied: tqdm<5,>=4.48.2 in
/usr/local/lib/python3.10/dist-packages (from ydata-profiling) (4.66.1)
Requirement already satisfied: seaborn<0.13,>=0.10.1 in
/usr/local/lib/python3.10/dist-packages (from ydata-profiling) (0.12.2)
Collecting multimethod<2,>=1.4 (from ydata-profiling)
Downloading multimethod-1.10-py3-none-any.whl (9.9 kB)
Requirement already satisfied: statsmodels<1,>=0.13.2 in
/usr/local/lib/python3.10/dist-packages (from ydata-profiling) (0.14.0)
Collecting typeguard<5,>=4.1.2 (from ydata-profiling)
Downloading typeguard-4.1.5-py3-none-any.whl (34 kB)
Collecting imagehash==4.3.1 (from ydata-profiling)
Downloading ImageHash-4.3.1-py2.py3-none-any.whl (296 kB)
296.5/296.5

kB 26.4 MB/s eta 0:00:00

Requirement already satisfied: wordcloud>=1.9.1 in
/usr/local/lib/python3.10/dist-packages (from ydata-profiling) (1.9.2)
Collecting dacite>=1.8 (from ydata-profiling)
Downloading dacite-1.8.1-py3-none-any.whl (14 kB)
Requirement already satisfied: numba<0.59.0,>=0.56.0 in
/usr/local/lib/python3.10/dist-packages (from ydata-profiling) (0.58.1)
Requirement already satisfied: PyWavelets in /usr/local/lib/python3.10/dist-
packages (from imagehash==4.3.1->ydata-profiling) (1.5.0)
Requirement already satisfied: pillow in /usr/local/lib/python3.10/dist-packages
(from imagehash==4.3.1->ydata-profiling) (9.4.0)
Requirement already satisfied: attrs>=19.3.0 in /usr/local/lib/python3.10/dist-
packages (from visions[type_image_path]==0.7.5->ydata-profiling) (23.1.0)
Requirement already satisfied: networkx>=2.4 in /usr/local/lib/python3.10/dist-
packages (from visions[type_image_path]==0.7.5->ydata-profiling) (3.2.1)
Collecting tangled-up-in-unicode>=0.0.4 (from
visions[type_image_path]==0.7.5->ydata-profiling)
Downloading tangled_up_in_unicode-0.2.0-py3-none-any.whl (4.7 MB)
4.7/4.7 MB

40.0 MB/s eta 0:00:00

Requirement already satisfied: MarkupSafe>=2.0 in
/usr/local/lib/python3.10/dist-packages (from jinja2<3.2,>=2.11.1->ydata-
profiling) (2.1.3)
Requirement already satisfied: contourpy>=1.0.1 in

/usr/local/lib/python3.10/dist-packages (from matplotlib<3.9,>=3.2->ydata-
 profiling) (1.2.0)
 Requirement already satisfied: cycycler>=0.10 in /usr/local/lib/python3.10/dist-
 packages (from matplotlib<3.9,>=3.2->ydata-profiling) (0.12.1)
 Requirement already satisfied: fonttools>=4.22.0 in
 /usr/local/lib/python3.10/dist-packages (from matplotlib<3.9,>=3.2->ydata-
 profiling) (4.45.1)
 Requirement already satisfied: kiwisolver>=1.0.1 in
 /usr/local/lib/python3.10/dist-packages (from matplotlib<3.9,>=3.2->ydata-
 profiling) (1.4.5)
 Requirement already satisfied: packaging>=20.0 in
 /usr/local/lib/python3.10/dist-packages (from matplotlib<3.9,>=3.2->ydata-
 profiling) (23.2)
 Requirement already satisfied: pyparsing>=2.3.1 in
 /usr/local/lib/python3.10/dist-packages (from matplotlib<3.9,>=3.2->ydata-
 profiling) (3.1.1)
 Requirement already satisfied: python-dateutil>=2.7 in
 /usr/local/lib/python3.10/dist-packages (from matplotlib<3.9,>=3.2->ydata-
 profiling) (2.8.2)
 Requirement already satisfied: llvmlite<0.42,>=0.41.0dev0 in
 /usr/local/lib/python3.10/dist-packages (from numba<0.59.0,>=0.56.0->ydata-
 profiling) (0.41.1)
 Requirement already satisfied: pytz>=2020.1 in /usr/local/lib/python3.10/dist-
 packages (from pandas!=1.4.0,<3,>1.1->ydata-profiling) (2023.3.post1)
 Requirement already satisfied: joblib>=0.14.1 in /usr/local/lib/python3.10/dist-
 packages (from phik<0.13,>=0.11.1->ydata-profiling) (1.3.2)
 Collecting annotated-types>=0.4.0 (from pydantic>=2->ydata-profiling)
 Downloading annotated_types-0.6.0-py3-none-any.whl (12 kB)
 Collecting pydantic-core==2.14.5 (from pydantic>=2->ydata-profiling)
 Downloading
 pydantic_core-2.14.5-cp310-cp310-manylinux_2_17_x86_64.manylinux2014_x86_64.whl
 (2.1 MB)

2.1/2.1 MB

34.4 MB/s eta 0:00:00

Collecting typing-extensions>=4.6.1 (from pydantic>=2->ydata-profiling)
 Downloading typing_extensions-4.8.0-py3-none-any.whl (31 kB)
 Requirement already satisfied: charset-normalizer<4,>=2 in
 /usr/local/lib/python3.10/dist-packages (from requests<3,>=2.24.0->ydata-
 profiling) (3.3.2)
 Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.10/dist-
 packages (from requests<3,>=2.24.0->ydata-profiling) (3.6)
 Requirement already satisfied: urllib3<3,>=1.21.1 in
 /usr/local/lib/python3.10/dist-packages (from requests<3,>=2.24.0->ydata-
 profiling) (2.0.7)
 Requirement already satisfied: certifi>=2017.4.17 in
 /usr/local/lib/python3.10/dist-packages (from requests<3,>=2.24.0->ydata-
 profiling) (2023.11.17)
 Requirement already satisfied: patsy>=0.5.2 in /usr/local/lib/python3.10/dist-

```

packages (from statsmodels<1,>=0.13.2->ydata-profiling) (0.5.3)
Requirement already satisfied: six in /usr/local/lib/python3.10/dist-packages
(from patsy>=0.5.2->statsmodels<1,>=0.13.2->ydata-profiling) (1.16.0)
Building wheels for collected packages: htmlmin
  Building wheel for htmlmin (setup.py) ... done
  Created wheel for htmlmin: filename=htmlmin-0.1.12-py3-none-any.whl size=27081
sha256=150f351153c08f456c7d609e815ef08d04d1a7d8e46ca9126e7d9347d841b02e
  Stored in directory: /root/.cache/pip/wheels/dd/91/29/a79cecb328d01739e64017b6
fb9a1ab9d8cb1853098ec5966d
Successfully built htmlmin
Installing collected packages: htmlmin, typing-extensions, tangled-up-in-
unicode, multimethod, dacite, annotated-types, typeguard, pydantic-core,
imagehash, visions, pydantic, phik, ydata-profiling
  Attempting uninstall: typing-extensions
    Found existing installation: typing_extensions 4.5.0
    Uninstalling typing_extensions-4.5.0:
      Successfully uninstalled typing_extensions-4.5.0
  Attempting uninstall: pydantic
    Found existing installation: pydantic 1.10.13
    Uninstalling pydantic-1.10.13:
      Successfully uninstalled pydantic-1.10.13
ERROR: pip's dependency resolver does not currently take into account all
the packages that are installed. This behaviour is the source of the following
dependency conflicts.

lida 0.0.10 requires fastapi, which is not installed.
lida 0.0.10 requires kaleido, which is not installed.
lida 0.0.10 requires python-multipart, which is not installed.
lida 0.0.10 requires uvicorn, which is not installed.
llmx 0.0.15a0 requires cohere, which is not installed.
llmx 0.0.15a0 requires openai, which is not installed.
llmx 0.0.15a0 requires tiktoken, which is not installed.
tensorflow-probability 0.22.0 requires typing-extensions<4.6.0, but you have
typing-extensions 4.8.0 which is incompatible.

Successfully installed annotated-types-0.6.0 dacite-1.8.1 htmlmin-0.1.12
imagehash-4.3.1 multimethod-1.10 phik-0.12.3 pydantic-2.5.2 pydantic-core-2.14.5
tangled-up-in-unicode-0.2.0 typeguard-4.1.5 typing-extensions-4.8.0
visions-0.7.5 ydata-profiling-4.6.3

```

```
[ ]: pip install shap
```

```

Collecting shap
  Downloading shap-0.44.0-cp310-cp310-

```

manylinux_2_12_x86_64.manylinux2010_x86_64.manylinux_2_17_x86_64.manylinux2014_x86_64.whl (533 kB)

533.5/533.5

kB 4.8 MB/s eta 0:00:00

Requirement already satisfied: numpy in /usr/local/lib/python3.10/dist-packages (from shap) (1.23.5)
Requirement already satisfied: scipy in /usr/local/lib/python3.10/dist-packages (from shap) (1.11.4)
Requirement already satisfied: scikit-learn in /usr/local/lib/python3.10/dist-packages (from shap) (1.2.2)
Requirement already satisfied: pandas in /usr/local/lib/python3.10/dist-packages (from shap) (1.5.3)
Requirement already satisfied: tqdm>=4.27.0 in /usr/local/lib/python3.10/dist-packages (from shap) (4.66.1)
Requirement already satisfied: packaging>20.9 in /usr/local/lib/python3.10/dist-packages (from shap) (23.2)
Collecting slicer==0.0.7 (from shap)
 Downloading slicer-0.0.7-py3-none-any.whl (14 kB)
Requirement already satisfied: numba in /usr/local/lib/python3.10/dist-packages (from shap) (0.58.1)
Requirement already satisfied: cloudpickle in /usr/local/lib/python3.10/dist-packages (from shap) (2.2.1)
Requirement already satisfied: llvmlite<0.42,>=0.41.0dev0 in /usr/local/lib/python3.10/dist-packages (from numba->shap) (0.41.1)
Requirement already satisfied: python-dateutil>=2.8.1 in /usr/local/lib/python3.10/dist-packages (from pandas->shap) (2.8.2)
Requirement already satisfied: pytz>=2020.1 in /usr/local/lib/python3.10/dist-packages (from pandas->shap) (2023.3.post1)
Requirement already satisfied: joblib>=1.1.1 in /usr/local/lib/python3.10/dist-packages (from scikit-learn->shap) (1.3.2)
Requirement already satisfied: threadpoolctl>=2.0.0 in /usr/local/lib/python3.10/dist-packages (from scikit-learn->shap) (3.2.0)
Requirement already satisfied: six>=1.5 in /usr/local/lib/python3.10/dist-packages (from python-dateutil>=2.8.1->pandas->shap) (1.16.0)
Installing collected packages: slicer, shap
Successfully installed shap-0.44.0 slicer-0.0.7

```
[ ]: # Import required libraries
import pandas as pd
import matplotlib.pyplot as plt
import numpy as np
import seaborn as sns
import plotly.express as px
from ydata_profiling import ProfileReport
from sklearn.preprocessing import MinMaxScaler
from sklearn.preprocessing import StandardScaler
from sklearn.decomposition import PCA
```

```

from sklearn.model_selection import train_test_split
from xgboost import XGBRegressor
from sklearn.metrics import mean_squared_error
from sklearn.model_selection import GridSearchCV
import shap

```

```

[ ]: #dataset link='https://www.kaggle.com/datasets/mathchi/churn-for-bank-customers'
      #load the dataset

df=pd.read_csv('/content/drive/MyDrive/Colab/churn.csv')
print(df.shape)

```

(10000, 14)

2 Describe Data

```

[ ]: # Data types in dataset
df.info()

```

```

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 10000 entries, 0 to 9999
Data columns (total 14 columns):
 #   Column                Non-Null Count  Dtype
---  -
 0   RowNumber             10000 non-null  int64
 1   CustomerId            10000 non-null  int64
 2   Surname               10000 non-null  object
 3   CreditScore           10000 non-null  int64
 4   Geography             10000 non-null  object
 5   Gender                10000 non-null  object
 6   Age                   10000 non-null  int64
 7   Tenure                10000 non-null  int64
 8   Balance               10000 non-null  float64
 9   NumOfProducts         10000 non-null  int64
10   HasCrCard             10000 non-null  int64
11   IsActiveMember        10000 non-null  int64
12   EstimatedSalary       10000 non-null  float64
13   Exited                10000 non-null  int64
dtypes: float64(2), int64(9), object(3)
memory usage: 1.1+ MB

```

There are three different data types in the DataFrame: int64 (integer) - Used for columns like RowNumber, CustomerId, CreditScore, Age, Tenure, NumOfProducts, HasCrCard, IsActiveMember, and Exited. float64 (float) - Used for columns like Balance and EstimatedSalary. object - Used for categorical data such as Surname, Geography, and Gender.

3 Data Preprocessing

```
[ ]: #Check for missing values
missing_values = df.isnull().sum()
missing_values
```

```
[ ]: RowNumber      0
      CustomerId    0
      Surname       0
      CreditScore   0
      Geography     0
      Gender        0
      Age           0
      Tenure        0
      Balance       0
      NumOfProducts 0
      HasCrCard     0
      IsActiveMember 0
      EstimatedSalary 0
      Exited        0
      dtype: int64
```

There are no missing values in any of the columns.

```
[ ]: # no. of duplicates
duplicates = len(df[df.duplicated()])
print('Number of Duplicate Entries: %d'%(duplicates))
```

Number of Duplicate Entries: 0

```
[ ]: df.head()
```

```
[ ]:   RowNumber  CustomerId  Surname  CreditScore  Geography  Gender  Age  \
0         1    15634602   Hargrave         619     France  Female   42
1         2    15647311     Hill         608      Spain  Female   41
2         3    15619304     Onio         502     France  Female   42
3         4    15701354     Boni         699     France  Female   39
4         5    15737888  Mitchell         850      Spain  Female   43

      Tenure  Balance  NumOfProducts  HasCrCard  IsActiveMember  \
0         2     0.00             1           1              1
1         1  83807.86             1           0              1
2         8 159660.80             3           1              0
3         1     0.00             2           0              0
4         2 125510.82             1           1              1

      EstimatedSalary  Exited
0         101348.88         1
```

1	112542.58	0
2	113931.57	1
3	93826.63	0
4	79084.10	0

```
[ ]: #drop features which are not required
df.drop(['RowNumber', 'CustomerId', 'Surname'], axis=1, inplace=True)
df.head()
```

```
[ ]:   CreditScore  Geography  Gender  Age  Tenure  Balance  NumOfProducts  \
0         619    France  Female  42     2     0.00             1
1         608    Spain  Female  41     1  83807.86             1
2         502    France  Female  42     8 159660.80             3
3         699    France  Female  39     1     0.00             2
4         850    Spain  Female  43     2 125510.82             1
```

	HasCrCard	IsActiveMember	EstimatedSalary	Exited
0	1	1	101348.88	1
1	0	1	112542.58	0
2	1	0	113931.57	1
3	0	0	93826.63	0
4	1	1	79084.10	0

```
[ ]: #summary statistics of the numerical columns, values are formatted with two
    ↪ decimal places
df.describe().applymap(lambda x: f"{x:0.2f}")
```

```
[ ]:   CreditScore    Age    Tenure    Balance  NumOfProducts  HasCrCard  \
count  10000.00  10000.00  10000.00  10000.00    10000.00    10000.00
mean     650.53    38.92     5.01   76485.89         1.53         0.71
std       96.65    10.49     2.89   62397.41         0.58         0.46
min      350.00    18.00     0.00     0.00         1.00         0.00
25%      584.00    32.00     3.00     0.00         1.00         0.00
50%      652.00    37.00     5.00   97198.54         1.00         1.00
75%      718.00    44.00     7.00  127644.24         2.00         1.00
max      850.00    92.00    10.00  250898.09         4.00         1.00
```

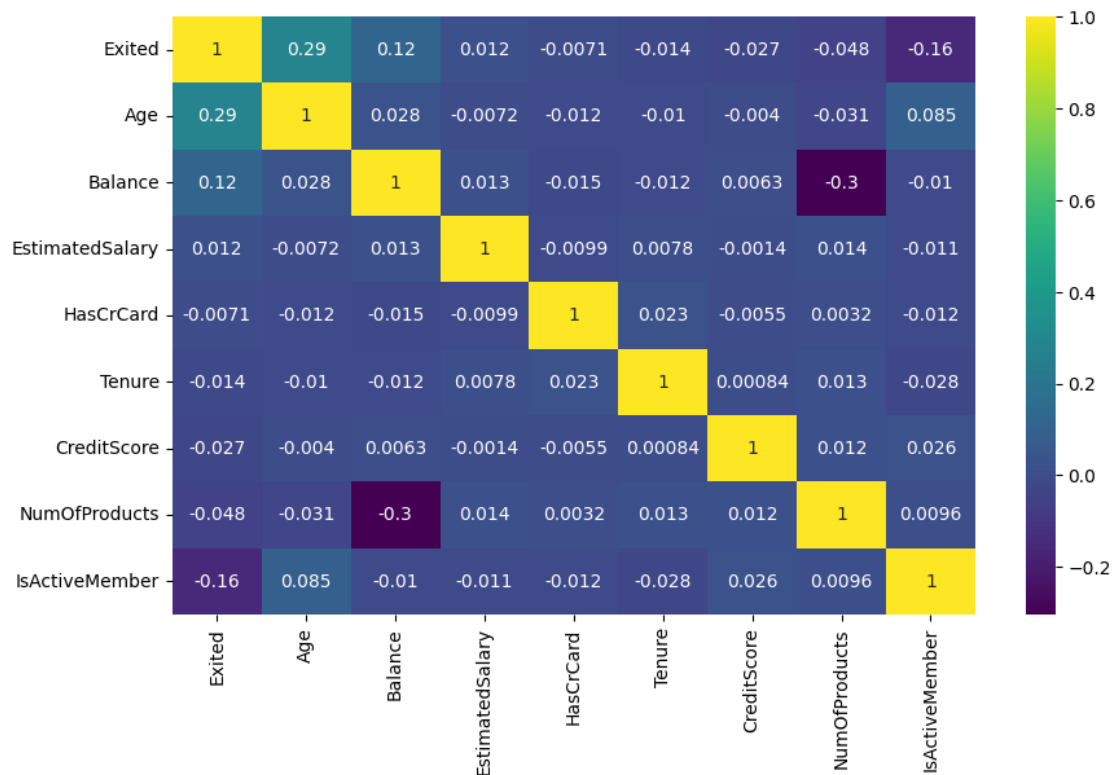
	IsActiveMember	EstimatedSalary	Exited
count	10000.00	10000.00	10000.00
mean	0.52	100090.24	0.20
std	0.50	57510.49	0.40
min	0.00	11.58	0.00
25%	0.00	51002.11	0.00
50%	1.00	100193.91	0.00
75%	1.00	149388.25	0.00
max	1.00	199992.48	1.00

4 Correlation Matrix

```
[ ]: # Exited correlation matrix
k = 10 #number of variables for heatmap
cols = df.corr().nlargest(k, 'Exited')['Exited'].index
cm = df[cols].corr()
plt.figure(figsize=(10,6))
sns.heatmap(cm, annot=True, cmap = 'viridis')
```

The default value of `numeric_only` in `DataFrame.corr` is deprecated. In a future version, it will default to `False`. Select only valid columns or specify the value of `numeric_only` to silence this warning.

```
[ ]: <Axes: >
```



5 Variance

```
[ ]: # Checking the variance of variables
variance_info = df.var()

print("variance info:")
print(variance_info)
```

```

variance info:
CreditScore      9.341860e+03
Age              1.099941e+02
Tenure           8.364673e+00
Balance          3.893436e+09
NumOfProducts   3.383218e-01
HasCrCard        2.077905e-01
IsActiveMember   2.497970e-01
EstimatedSalary  3.307457e+09
Exited           1.622225e-01
dtype: float64

```

The default value of `numeric_only` in `DataFrame.var` is deprecated. In a future version, it will default to `False`. In addition, specifying `'numeric_only=None'` is deprecated. Select only valid columns or specify the value of `numeric_only` to silence this warning.

6 pandas profiling

```

[ ]: #pandas profiling
profile = df.profile_report(
title = "Pandas Profiling Report",
correlations = {
"pearson": {"calculate": True},
"spearman": {"calculate": True},
"kendall": {"calculate": True}
},
)
#profile.to_file("pandas_profiled.html")
profile.to_notebook_iframe()

```

Output hidden; open in <https://colab.research.google.com> to view.

No duplicate values in the dataset. No missing values in the dataset. From the summary statistics and the graphs, we could conclude that the features looks Ok. There is much of correlation between any of the variables. The highest correlation is 0.29, which is between Age and Exited, which suggests a small positive relationship between Age and Customer Churn. There could be a relationship between Credit Score, Age, Balance, Number of Products with Exited respectively.

Overall, the correaltion is very weak.

Weak positive correlation with age Very weak positive correlation with balance Very weak negative correlations with number of products and membership

7 Onehot encode

```
[ ]: #Onehot encode categorical features
geog = pd.get_dummies(df['Geography'], drop_first=True)
gen = pd.get_dummies(df['Gender'], drop_first=True)
```

```
[ ]: #Concat the encoded features to the main dataframe
df = pd.concat([df, gen, geog], axis=1)

#Drop the old categorical features
df.drop(['Geography', 'Gender'], axis=1, inplace=True)
df.head()
```

```
[ ]:   CreditScore  Age  Tenure   Balance  NumOfProducts  HasCrCard  \
0         619   42     2      0.00             1           1
1         608   41     1  83807.86             1           0
2         502   42     8 159660.80             3           1
3         699   39     1      0.00             2           0
4         850   43     2 125510.82             1           1

   IsActiveMember  EstimatedSalary  Exited  Male  Germany  Spain
0                1      101348.88        1     0         0       0
1                1      112542.58        0     0         0       1
2                0      113931.57        1     0         0       0
3                0       93826.63        0     0         0       0
4                1       79084.10        0     0         0       1
```

8 scaling

```
[ ]: # features
features = [
    'CreditScore', 'Age', 'Tenure', 'Balance', 'NumOfProducts', 'HasCrCard', 'IsActiveMember', 'EstimatedSalary',
    'Male', 'Germany', 'Spain']
```

```
[ ]: # feature scaling

# normalization dataset with MinMaxScaler
scaler = MinMaxScaler()
df[features] = scaler.fit_transform(df[features])
customer_features = df[features]
customer_features
```

```
[ ]:   CreditScore   Age  Tenure   Balance  NumOfProducts  HasCrCard  \
0         0.538  0.324324    0.2  0.000000      0.000000         1.0
1         0.516  0.310811    0.1  0.334031      0.000000         0.0
2         0.304  0.324324    0.8  0.636357      0.666667         1.0
```

3	0.698	0.283784	0.1	0.000000	0.333333	0.0
4	1.000	0.337838	0.2	0.500246	0.000000	1.0
...
9995	0.842	0.283784	0.5	0.000000	0.333333	1.0
9996	0.332	0.229730	1.0	0.228657	0.000000	1.0
9997	0.718	0.243243	0.7	0.000000	0.000000	0.0
9998	0.844	0.324324	0.3	0.299226	0.333333	1.0
9999	0.884	0.135135	0.4	0.518708	0.000000	1.0

	IsActiveMember	EstimatedSalary	Exited	Male	Germany	Spain
0	1.0	0.506735	1.0	0.0	0.0	0.0
1	1.0	0.562709	0.0	0.0	0.0	1.0
2	0.0	0.569654	1.0	0.0	0.0	0.0
3	0.0	0.469120	0.0	0.0	0.0	0.0
4	1.0	0.395400	0.0	0.0	0.0	1.0
...
9995	0.0	0.481341	0.0	1.0	0.0	0.0
9996	1.0	0.508490	0.0	1.0	0.0	0.0
9997	1.0	0.210390	1.0	0.0	0.0	0.0
9998	0.0	0.464429	1.0	1.0	1.0	0.0
9999	0.0	0.190914	0.0	0.0	0.0	0.0

[10000 rows x 12 columns]

```
[ ]: # Standardization
scaler = StandardScaler()
data_stand = scaler.fit_transform(df)
```

```
[ ]: # Checking mean of each feature
mean_values = data_stand.mean(axis=0)

# Checking standard deviation of each feature
std_dev_values = data_stand.std(axis=0)

# Displaying the mean and standard deviation
print("Mean of each feature after standardization:")
print(mean_values)

print("\nStandard Deviation of each feature after standardization:")
print(std_dev_values)
```

Mean of each feature after standardization:

```
[-5.25801624e-17  2.30393482e-16 -9.69890834e-17 -5.68434189e-17
 2.87769808e-17 -5.25801624e-17 -7.38964445e-17  8.06466005e-17
 6.67910172e-17  1.31450406e-16 -7.06990022e-17 -4.68958206e-17]
```

Standard Deviation of each feature after standardization:

```
[1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1.]
```

9 PCA

```
[ ]: # PCA

pca = PCA(n_components=4)
data_pca = pca.fit_transform(df)
```

```
[ ]: print("Explained Variance Ratio:")
      print(pca.explained_variance_ratio_)
```

```
Explained Variance Ratio:
[0.1815528  0.1587977  0.15568359 0.13271483]
```

```
[ ]: print("Cumulative Explained Variance:")
      print(np.cumsum(pca.explained_variance_ratio_))
```

```
Cumulative Explained Variance:
[0.1815528  0.3403505  0.49603409 0.62874892]
```

10 Train Test Split

```
[ ]: # Splitting the data into training and testing sets
X_train, X_test, y_train, y_test = train_test_split(df.drop('Exited', axis=1),
                                                    df['Exited'],
                                                    test_size=0.2, random_state=42)

# Initializing the XGBoost model
model = XGBRegressor()
# Training the model on the training data
model.fit(X_train, y_train)
# Predicting on the test data
y_pred = model.predict(X_test)
# Evaluating the model
mse = mean_squared_error(y_test, y_pred)
print(f"Mean Squared Error: {mse}")
```

```
Mean Squared Error: 0.11328349743109671
```

```
[ ]: # Example of hyperparameter optimization for XGBoost
param_grid = {
    'n_estimators': [100, 200, 300],
    'learning_rate': [0.01, 0.1, 0.2],
    'max_depth': [3, 4, 5]
}
# Grid search with cross-validation
```

```

grid_search = GridSearchCV(estimator=XGBRegressor(), param_grid=param_grid,
    ↳scoring='neg_mean_squared_error', cv=5)
grid_search.fit(X_train, y_train)

# Best parameters
best_params = grid_search.best_params_
print(f"Best parameters: {best_params}")

```

Best parameters: {'learning_rate': 0.1, 'max_depth': 4, 'n_estimators': 100}

```

[ ]: # Best parameters from grid search
best_params = {'n_estimators': 100, 'learning_rate': 0.1, 'max_depth': 4} #
    ↳Replace with your actual best_params

# Initialize XGBoost model with the best parameters
best_model = XGBRegressor(**best_params)

# Train the model on the training data
best_model.fit(X_train, y_train)

# Predict on the test data
y_pred_best = best_model.predict(X_test)

# Evaluate the model
mse_best = mean_squared_error(y_test, y_pred_best)
print(f"Mean Squared Error with Best Parameters: {mse_best}")

```

Mean Squared Error with Best Parameters: 0.09759047410572504

The model's predictions are close to the actual values, with relatively small errors.

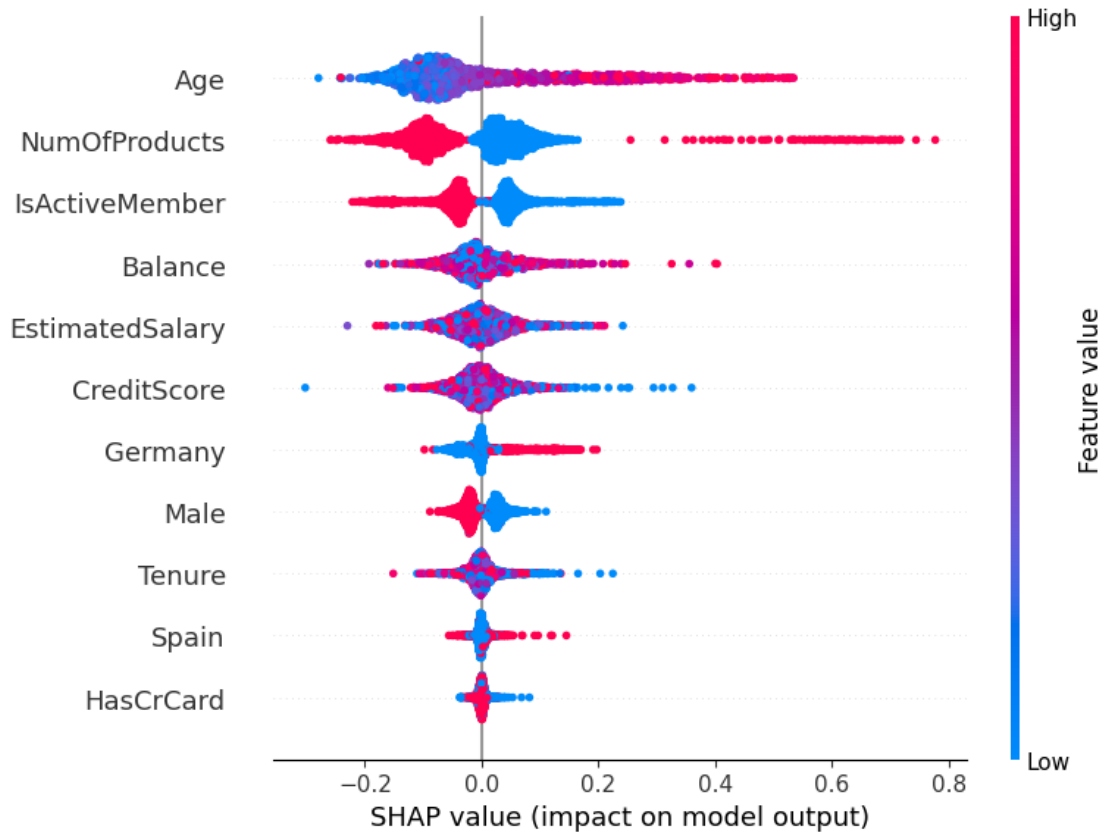
11 SHAP values

```

[ ]: # Create a SHAP explainer for the XGBoost model
explainer = shap.Explainer(model)
# Calculate SHAP values for the test data
shap_values = explainer.shap_values(X_test)
# Visualize SHAP values using a summary plot
shap.summary_plot(shap_values, X_test)

```

[07:54:44] WARNING: /workspace/src/c_api/c_api.cc:1240: Saving into deprecated binary model format, please consider using `json` or `ubj`. Model format will default to JSON in XGBoost 2.2 if not specified.



```
[32]: !sudo apt-get install texlive-xetex texlive-fonts-recommended_
      ↪texlive-plain-generic
```

Reading package lists... Done

Building dependency tree... Done

Reading state information... Done

The following additional packages will be installed:

```
dvisvgm fonts-droid-fallback fonts-lato fonts-lmodern fonts-noto-mono
fonts-texgyre fonts-urw-base35 libapache-pom-java libcommons-logging-java
libcommons-parent-java libfontbox-java libfontenc1 libgs9 libgs9-common
libidn12 libijs-0.35 libjbig2dec0 libkpathsea6 libpdfbox-java libptexenc1
libruby3.0 libsynchronet2 libteckit0 libtexlua53 libtexluajit2 libwoff1
libzip-0-13 lmodern poppler-data preview-latex-style rake ruby
ruby-net-telnet ruby-rubygems ruby-webrick ruby-xmlrpc ruby3.0
rubygems-integration tlutils teckit tex-common tex-gyre texlive-base
texlive-binaries texlive-latex-base texlive-latex-extra
texlive-latex-recommended texlive-pictures tipa xfonts-encodings
xfonts-utils
```

Suggested packages:

```
fonts-noto fonts-freefont-otf | fonts-freefont-ttf libavalon-framework-java
libcommons-logging-java-doc libexcalibur-logkit-java liblog4j1.2-java
```

```

poppler-utils ghostscript fonts-japanese-mincho | fonts-ipafont-mincho
fonts-japanese-gothic | fonts-ipafont-gothic fonts-arphic-ukai
fonts-arphic-uming fonts-nanum ri ruby-dev bundler debhelper gv
| postscript-viewer perl-tk xpdf | pdf-viewer xzdec
texlive-fonts-recommended-doc texlive-latex-base-doc python3-pygments
icc-profiles libfile-which-perl libspreadsheet-parseexcel-perl
texlive-latex-extra-doc texlive-latex-recommended-doc texlive-luatex
texlive-pstricks dot2tex prerex texlive-pictures-doc vprerex
default-jre-headless tipa-doc

```

The following NEW packages will be installed:

```

dvisvgm fonts-droid-fallback fonts-lato fonts-lmodern fonts-noto-mono
fonts-texgyre fonts-urw-base35 libapache-pom-java libcommons-logging-java
libcommons-parent-java libfontbox-java libfontenc1 libgs9 libgs9-common
libidn12 libijs-0.35 libjbig2dec0 libkpathsea6 libpdfbox-java libptexenc1
libruby3.0 libsynchronet2 libteckit0 libtexlua53 libtexlua53-luatex libwoff1
libzip-0-13 lmodern poppler-data preview-latex-style rake ruby
ruby-net-telnet ruby-rubygems ruby-webrick ruby-xmlrpc ruby3.0
rubygems-integration tlutils teckit tex-common tex-gyre texlive-base
texlive-binaries texlive-fonts-recommended texlive-latex-base
texlive-latex-extra texlive-latex-recommended texlive-pictures
texlive-plain-generic texlive-xetex tipa xfonts-encodings xfonts-utils

```

0 upgraded, 54 newly installed, 0 to remove and 15 not upgraded.

Need to get 182 MB of archives.

After this operation, 571 MB of additional disk space will be used.

Get:1 <http://archive.ubuntu.com/ubuntu> jammy/main amd64 fonts-droid-fallback all 1:6.0.1r16-1.1build1 [1,805 kB]

Get:2 <http://archive.ubuntu.com/ubuntu> jammy/main amd64 fonts-lato all 2.0-2.1 [2,696 kB]

Get:3 <http://archive.ubuntu.com/ubuntu> jammy/main amd64 poppler-data all 0.4.11-1 [2,171 kB]

Get:4 <http://archive.ubuntu.com/ubuntu> jammy/universe amd64 tex-common all 6.17 [33.7 kB]

Get:5 <http://archive.ubuntu.com/ubuntu> jammy/main amd64 fonts-urw-base35 all 20200910-1 [6,367 kB]

Get:6 <http://archive.ubuntu.com/ubuntu> jammy-updates/main amd64 libgs9-common all 9.55.0~dfsg1-0ubuntu5.5 [752 kB]

Get:7 <http://archive.ubuntu.com/ubuntu> jammy-updates/main amd64 libidn12 amd64 1.38-4ubuntu1 [60.0 kB]

Get:8 <http://archive.ubuntu.com/ubuntu> jammy/main amd64 libijs-0.35 amd64 0.35-15build2 [16.5 kB]

Get:9 <http://archive.ubuntu.com/ubuntu> jammy/main amd64 libjbig2dec0 amd64 0.19-3build2 [64.7 kB]

Get:10 <http://archive.ubuntu.com/ubuntu> jammy-updates/main amd64 libgs9 amd64 9.55.0~dfsg1-0ubuntu5.5 [5,030 kB]

Get:11 <http://archive.ubuntu.com/ubuntu> jammy-updates/main amd64 libkpathsea6 amd64 2021.20210626.59705-1ubuntu0.1 [60.3 kB]

Get:12 <http://archive.ubuntu.com/ubuntu> jammy/main amd64 libwoff1 amd64 1.0.2-1build4 [45.2 kB]

Get:13 <http://archive.ubuntu.com/ubuntu> jammy/universe amd64 dvisvgm amd64 2.13.1-1 [1,221 kB]
Get:14 <http://archive.ubuntu.com/ubuntu> jammy/universe amd64 fonts-lmodern all 2.004.5-6.1 [4,532 kB]
Get:15 <http://archive.ubuntu.com/ubuntu> jammy/main amd64 fonts-noto-mono all 20201225-1build1 [397 kB]
Get:16 <http://archive.ubuntu.com/ubuntu> jammy/universe amd64 fonts-texgyre all 20180621-3.1 [10.2 MB]
Get:17 <http://archive.ubuntu.com/ubuntu> jammy/universe amd64 libapache-pom-java all 18-1 [4,720 B]
Get:18 <http://archive.ubuntu.com/ubuntu> jammy/universe amd64 libcommons-parent-java all 43-1 [10.8 kB]
Get:19 <http://archive.ubuntu.com/ubuntu> jammy/universe amd64 libcommons-logging-java all 1.2-2 [60.3 kB]
Get:20 <http://archive.ubuntu.com/ubuntu> jammy/main amd64 libfontenc1 amd64 1:1.1.4-1build3 [14.7 kB]
Get:21 <http://archive.ubuntu.com/ubuntu> jammy-updates/main amd64 libptexenc1 amd64 2021.20210626.59705-1ubuntu0.1 [39.1 kB]
Get:22 <http://archive.ubuntu.com/ubuntu> jammy/main amd64 rubygems-integration all 1.18 [5,336 B]
Get:23 <http://archive.ubuntu.com/ubuntu> jammy-updates/main amd64 ruby3.0 amd64 3.0.2-7ubuntu2.4 [50.1 kB]
Get:24 <http://archive.ubuntu.com/ubuntu> jammy/main amd64 ruby-rubygems all 3.3.5-2 [228 kB]
Get:25 <http://archive.ubuntu.com/ubuntu> jammy/main amd64 ruby amd64 1:3.0~exp1 [5,100 B]
Get:26 <http://archive.ubuntu.com/ubuntu> jammy/main amd64 rake all 13.0.6-2 [61.7 kB]
Get:27 <http://archive.ubuntu.com/ubuntu> jammy/main amd64 ruby-net-telnet all 0.1.1-2 [12.6 kB]
Get:28 <http://archive.ubuntu.com/ubuntu> jammy/universe amd64 ruby-webrick all 1.7.0-3 [51.8 kB]
Get:29 <http://archive.ubuntu.com/ubuntu> jammy-updates/main amd64 ruby-xmlrpc all 0.3.2-1ubuntu0.1 [24.9 kB]
Get:30 <http://archive.ubuntu.com/ubuntu> jammy-updates/main amd64 libruby3.0 amd64 3.0.2-7ubuntu2.4 [5,113 kB]
Get:31 <http://archive.ubuntu.com/ubuntu> jammy-updates/main amd64 libsinctex2 amd64 2021.20210626.59705-1ubuntu0.1 [55.5 kB]
Get:32 <http://archive.ubuntu.com/ubuntu> jammy/universe amd64 libteckit0 amd64 2.5.11+ds1-1 [421 kB]
Get:33 <http://archive.ubuntu.com/ubuntu> jammy-updates/main amd64 libtexlua53 amd64 2021.20210626.59705-1ubuntu0.1 [120 kB]
Get:34 <http://archive.ubuntu.com/ubuntu> jammy-updates/main amd64 libtexluajit2 amd64 2021.20210626.59705-1ubuntu0.1 [267 kB]
Get:35 <http://archive.ubuntu.com/ubuntu> jammy/universe amd64 libzip-0-13 amd64 0.13.72+dfsg.1-1.1 [27.0 kB]
Get:36 <http://archive.ubuntu.com/ubuntu> jammy/main amd64 xfonts-encodings all 1:1.0.5-0ubuntu2 [578 kB]

```

Get:37 http://archive.ubuntu.com/ubuntu jammy/main amd64 xfonts-utils amd64
1:7.7+6build2 [94.6 kB]
Get:38 http://archive.ubuntu.com/ubuntu jammy/universe amd64 lmodern all
2.004.5-6.1 [9,471 kB]
Get:39 http://archive.ubuntu.com/ubuntu jammy/universe amd64 preview-latex-style
all 12.2-1ubuntu1 [185 kB]
Get:40 http://archive.ubuntu.com/ubuntu jammy/main amd64 t1utils amd64
1.41-4build2 [61.3 kB]
Get:41 http://archive.ubuntu.com/ubuntu jammy/universe amd64 teckit amd64
2.5.11+ds1-1 [699 kB]
Get:42 http://archive.ubuntu.com/ubuntu jammy/universe amd64 tex-gyre all
20180621-3.1 [6,209 kB]
Get:43 http://archive.ubuntu.com/ubuntu jammy-updates/universe amd64 texlive-
binaries amd64 2021.20210626.59705-1ubuntu0.1 [9,848 kB]
Get:44 http://archive.ubuntu.com/ubuntu jammy/universe amd64 texlive-base all
2021.20220204-1 [21.0 MB]
Get:45 http://archive.ubuntu.com/ubuntu jammy/universe amd64 texlive-fonts-
recommended all 2021.20220204-1 [4,972 kB]
Get:46 http://archive.ubuntu.com/ubuntu jammy/universe amd64 texlive-latex-base
all 2021.20220204-1 [1,128 kB]
Get:47 http://archive.ubuntu.com/ubuntu jammy/universe amd64 libfontbox-java all
1:1.8.16-2 [207 kB]
Get:48 http://archive.ubuntu.com/ubuntu jammy/universe amd64 libpdfbox-java all
1:1.8.16-2 [5,199 kB]
Get:49 http://archive.ubuntu.com/ubuntu jammy/universe amd64 texlive-latex-
recommended all 2021.20220204-1 [14.4 MB]
Get:50 http://archive.ubuntu.com/ubuntu jammy/universe amd64 texlive-pictures
all 2021.20220204-1 [8,720 kB]
Get:51 http://archive.ubuntu.com/ubuntu jammy/universe amd64 texlive-latex-extra
all 2021.20220204-1 [13.9 MB]
Get:52 http://archive.ubuntu.com/ubuntu jammy/universe amd64 texlive-plain-
generic all 2021.20220204-1 [27.5 MB]
Get:53 http://archive.ubuntu.com/ubuntu jammy/universe amd64 tipa all 2:1.3-21
[2,967 kB]
Get:54 http://archive.ubuntu.com/ubuntu jammy/universe amd64 texlive-xetex all
2021.20220204-1 [12.4 MB]
Fetched 182 MB in 3s (64.6 MB/s)
debconf: unable to initialize frontend: Dialog
debconf: (No usable dialog-like program is installed, so the dialog based
frontend cannot be used. at /usr/share/perl5/Debconf/FrontEnd/Dialog.pm line 78,
<> line 54.)
debconf: falling back to frontend: Readline
debconf: unable to initialize frontend: Readline
debconf: (This frontend requires a controlling tty.)
debconf: falling back to frontend: Teletype
dpkg-preconfigure: unable to re-open stdin:
Selecting previously unselected package fonts-droid-fallback.
(Reading database ... 120899 files and directories currently installed.)

```

```

Preparing to unpack .../00-fonts-droid-fallback_1%3a6.0.1r16-1.1build1_all.deb
...
Unpacking fonts-droid-fallback (1:6.0.1r16-1.1build1) ...
Selecting previously unselected package fonts-lato.
Preparing to unpack .../01-fonts-lato_2.0-2.1_all.deb ...
Unpacking fonts-lato (2.0-2.1) ...
Selecting previously unselected package poppler-data.
Preparing to unpack .../02-poppler-data_0.4.11-1_all.deb ...
Unpacking poppler-data (0.4.11-1) ...
Selecting previously unselected package tex-common.
Preparing to unpack .../03-tex-common_6.17_all.deb ...
Unpacking tex-common (6.17) ...
Selecting previously unselected package fonts-urw-base35.
Preparing to unpack .../04-fonts-urw-base35_20200910-1_all.deb ...
Unpacking fonts-urw-base35 (20200910-1) ...
Selecting previously unselected package libgs9-common.
Preparing to unpack .../05-libgs9-common_9.55.0~dfsg1-0ubuntu5.5_all.deb ...
Unpacking libgs9-common (9.55.0~dfsg1-0ubuntu5.5) ...
Selecting previously unselected package libidn12:amd64.
Preparing to unpack .../06-libidn12_1.38-4ubuntu1_amd64.deb ...
Unpacking libidn12:amd64 (1.38-4ubuntu1) ...
Selecting previously unselected package libijs-0.35:amd64.
Preparing to unpack .../07-libijs-0.35_0.35-15build2_amd64.deb ...
Unpacking libijs-0.35:amd64 (0.35-15build2) ...
Selecting previously unselected package libjbig2dec0:amd64.
Preparing to unpack .../08-libjbig2dec0_0.19-3build2_amd64.deb ...
Unpacking libjbig2dec0:amd64 (0.19-3build2) ...
Selecting previously unselected package libgs9:amd64.
Preparing to unpack .../09-libgs9_9.55.0~dfsg1-0ubuntu5.5_amd64.deb ...
Unpacking libgs9:amd64 (9.55.0~dfsg1-0ubuntu5.5) ...
Selecting previously unselected package libkpathsea6:amd64.
Preparing to unpack .../10-libkpathsea6_2021.20210626.59705-1ubuntu0.1_amd64.deb
...
Unpacking libkpathsea6:amd64 (2021.20210626.59705-1ubuntu0.1) ...
Selecting previously unselected package libwoff1:amd64.
Preparing to unpack .../11-libwoff1_1.0.2-1build4_amd64.deb ...
Unpacking libwoff1:amd64 (1.0.2-1build4) ...
Selecting previously unselected package dvisvgm.
Preparing to unpack .../12-dvisvgm_2.13.1-1_amd64.deb ...
Unpacking dvisvgm (2.13.1-1) ...
Selecting previously unselected package fonts-lmodern.
Preparing to unpack .../13-fonts-lmodern_2.004.5-6.1_all.deb ...
Unpacking fonts-lmodern (2.004.5-6.1) ...
Selecting previously unselected package fonts-noto-mono.
Preparing to unpack .../14-fonts-noto-mono_20201225-1build1_all.deb ...
Unpacking fonts-noto-mono (20201225-1build1) ...
Selecting previously unselected package fonts-texgyre.
Preparing to unpack .../15-fonts-texgyre_20180621-3.1_all.deb ...

```

```

Unpacking fonts-texgyre (20180621-3.1) ...
Selecting previously unselected package libapache-pom-java.
Preparing to unpack .../16-libapache-pom-java_18-1_all.deb ...
Unpacking libapache-pom-java (18-1) ...
Selecting previously unselected package libcommons-parent-java.
Preparing to unpack .../17-libcommons-parent-java_43-1_all.deb ...
Unpacking libcommons-parent-java (43-1) ...
Selecting previously unselected package libcommons-logging-java.
Preparing to unpack .../18-libcommons-logging-java_1.2-2_all.deb ...
Unpacking libcommons-logging-java (1.2-2) ...
Selecting previously unselected package libfontenc1:amd64.
Preparing to unpack .../19-libfontenc1_1%3a1.1.4-1build3_amd64.deb ...
Unpacking libfontenc1:amd64 (1:1.1.4-1build3) ...
Selecting previously unselected package libptexenc1:amd64.
Preparing to unpack .../20-libptexenc1_2021.20210626.59705-1ubuntu0.1_amd64.deb
...
Unpacking libptexenc1:amd64 (2021.20210626.59705-1ubuntu0.1) ...
Selecting previously unselected package rubygems-integration.
Preparing to unpack .../21-rubygems-integration_1.18_all.deb ...
Unpacking rubygems-integration (1.18) ...
Selecting previously unselected package ruby3.0.
Preparing to unpack .../22-ruby3.0_3.0.2-7ubuntu2.4_amd64.deb ...
Unpacking ruby3.0 (3.0.2-7ubuntu2.4) ...
Selecting previously unselected package ruby-rubygems.
Preparing to unpack .../23-ruby-rubygems_3.3.5-2_all.deb ...
Unpacking ruby-rubygems (3.3.5-2) ...
Selecting previously unselected package ruby.
Preparing to unpack .../24-ruby_1%3a3.0~exp1_amd64.deb ...
Unpacking ruby (1:3.0~exp1) ...
Selecting previously unselected package rake.
Preparing to unpack .../25-rake_13.0.6-2_all.deb ...
Unpacking rake (13.0.6-2) ...
Selecting previously unselected package ruby-net-telnet.
Preparing to unpack .../26-ruby-net-telnet_0.1.1-2_all.deb ...
Unpacking ruby-net-telnet (0.1.1-2) ...
Selecting previously unselected package ruby-webrick.
Preparing to unpack .../27-ruby-webrick_1.7.0-3_all.deb ...
Unpacking ruby-webrick (1.7.0-3) ...
Selecting previously unselected package ruby-xmlrpc.
Preparing to unpack .../28-ruby-xmlrpc_0.3.2-1ubuntu0.1_all.deb ...
Unpacking ruby-xmlrpc (0.3.2-1ubuntu0.1) ...
Selecting previously unselected package libruby3.0:amd64.
Preparing to unpack .../29-libruby3.0_3.0.2-7ubuntu2.4_amd64.deb ...
Unpacking libruby3.0:amd64 (3.0.2-7ubuntu2.4) ...
Selecting previously unselected package libsyntax2:amd64.
Preparing to unpack .../30-libsyntax2_2021.20210626.59705-1ubuntu0.1_amd64.deb
...
Unpacking libsyntax2:amd64 (2021.20210626.59705-1ubuntu0.1) ...

```

```

Selecting previously unselected package libteckit0:amd64.
Preparing to unpack .../31-libteckit0_2.5.11+ds1-1_amd64.deb ...
Unpacking libteckit0:amd64 (2.5.11+ds1-1) ...
Selecting previously unselected package libtexlua53:amd64.
Preparing to unpack .../32-libtexlua53_2021.20210626.59705-1ubuntu0.1_amd64.deb
...
Unpacking libtexlua53:amd64 (2021.20210626.59705-1ubuntu0.1) ...
Selecting previously unselected package libtexluajit2:amd64.
Preparing to unpack
.../33-libtexluajit2_2021.20210626.59705-1ubuntu0.1_amd64.deb ...
Unpacking libtexluajit2:amd64 (2021.20210626.59705-1ubuntu0.1) ...
Selecting previously unselected package libzip-0-13:amd64.
Preparing to unpack .../34-libzip-0-13_0.13.72+dfsg.1-1.1_amd64.deb ...
Unpacking libzip-0-13:amd64 (0.13.72+dfsg.1-1.1) ...
Selecting previously unselected package xfonts-encodings.
Preparing to unpack .../35-xfonts-encodings_1%3a1.0.5-0ubuntu2_all.deb ...
Unpacking xfonts-encodings (1:1.0.5-0ubuntu2) ...
Selecting previously unselected package xfonts-utils.
Preparing to unpack .../36-xfonts-utils_1%3a7.7+6build2_amd64.deb ...
Unpacking xfonts-utils (1:7.7+6build2) ...
Selecting previously unselected package lmodern.
Preparing to unpack .../37-lmodern_2.004.5-6.1_all.deb ...
Unpacking lmodern (2.004.5-6.1) ...
Selecting previously unselected package preview-latex-style.
Preparing to unpack .../38-preview-latex-style_12.2-1ubuntu1_all.deb ...
Unpacking preview-latex-style (12.2-1ubuntu1) ...
Selecting previously unselected package t1utils.
Preparing to unpack .../39-t1utils_1.41-4build2_amd64.deb ...
Unpacking t1utils (1.41-4build2) ...
Selecting previously unselected package teckit.
Preparing to unpack .../40-teckit_2.5.11+ds1-1_amd64.deb ...
Unpacking teckit (2.5.11+ds1-1) ...
Selecting previously unselected package tex-gyre.
Preparing to unpack .../41-tex-gyre_20180621-3.1_all.deb ...
Unpacking tex-gyre (20180621-3.1) ...
Selecting previously unselected package texlive-binaries.
Preparing to unpack .../42-texlive-
binaries_2021.20210626.59705-1ubuntu0.1_amd64.deb ...
Unpacking texlive-binaries (2021.20210626.59705-1ubuntu0.1) ...
Selecting previously unselected package texlive-base.
Preparing to unpack .../43-texlive-base_2021.20220204-1_all.deb ...
Unpacking texlive-base (2021.20220204-1) ...
Selecting previously unselected package texlive-fonts-recommended.
Preparing to unpack .../44-texlive-fonts-recommended_2021.20220204-1_all.deb ...
Unpacking texlive-fonts-recommended (2021.20220204-1) ...
Selecting previously unselected package texlive-latex-base.
Preparing to unpack .../45-texlive-latex-base_2021.20220204-1_all.deb ...
Unpacking texlive-latex-base (2021.20220204-1) ...

```

```

Selecting previously unselected package libfontbox-java.
Preparing to unpack .../46-libfontbox-java_1%3a1.8.16-2_all.deb ...
Unpacking libfontbox-java (1:1.8.16-2) ...
Selecting previously unselected package libpdfbox-java.
Preparing to unpack .../47-libpdfbox-java_1%3a1.8.16-2_all.deb ...
Unpacking libpdfbox-java (1:1.8.16-2) ...
Selecting previously unselected package texlive-latex-recommended.
Preparing to unpack .../48-texlive-latex-recommended_2021.20220204-1_all.deb ...
Unpacking texlive-latex-recommended (2021.20220204-1) ...
Selecting previously unselected package texlive-pictures.
Preparing to unpack .../49-texlive-pictures_2021.20220204-1_all.deb ...
Unpacking texlive-pictures (2021.20220204-1) ...
Selecting previously unselected package texlive-latex-extra.
Preparing to unpack .../50-texlive-latex-extra_2021.20220204-1_all.deb ...
Unpacking texlive-latex-extra (2021.20220204-1) ...
Selecting previously unselected package texlive-plain-generic.
Preparing to unpack .../51-texlive-plain-generic_2021.20220204-1_all.deb ...
Unpacking texlive-plain-generic (2021.20220204-1) ...
Selecting previously unselected package tipa.
Preparing to unpack .../52-tipa_2%3a1.3-21_all.deb ...
Unpacking tipa (2:1.3-21) ...
Selecting previously unselected package texlive-xetex.
Preparing to unpack .../53-texlive-xetex_2021.20220204-1_all.deb ...
Unpacking texlive-xetex (2021.20220204-1) ...
Setting up fonts-lato (2.0-2.1) ...
Setting up fonts-noto-mono (20201225-1build1) ...
Setting up libwoff1:amd64 (1.0.2-1build4) ...
Setting up libtexlua53:amd64 (2021.20210626.59705-1ubuntu0.1) ...
Setting up libijs-0.35:amd64 (0.35-15build2) ...
Setting up libtexluajit2:amd64 (2021.20210626.59705-1ubuntu0.1) ...
Setting up libfontbox-java (1:1.8.16-2) ...
Setting up rubygems-integration (1.18) ...
Setting up libzip-0-13:amd64 (0.13.72+dfsg.1-1.1) ...
Setting up fonts-urw-base35 (20200910-1) ...
Setting up poppler-data (0.4.11-1) ...
Setting up tex-common (6.17) ...
debconf: unable to initialize frontend: Dialog
debconf: (No usable dialog-like program is installed, so the dialog based
frontend cannot be used. at /usr/share/perl5/Debconf/FrontEnd/Dialog.pm line
78.)
debconf: falling back to frontend: Readline
update-language: texlive-base not installed and configured, doing nothing!
Setting up libfontenc1:amd64 (1:1.1.4-1build3) ...
Setting up libjbig2dec0:amd64 (0.19-3build2) ...
Setting up libteckit0:amd64 (2.5.11+ds1-1) ...
Setting up libapache-pom-java (18-1) ...
Setting up ruby-net-telnet (0.1.1-2) ...
Setting up xfonts-encodings (1:1.0.5-0ubuntu2) ...

```

```

Setting up tlutils (1.41-4build2) ...
Setting up libidn12:amd64 (1.38-4ubuntu1) ...
Setting up fonts-texgyre (20180621-3.1) ...
Setting up libkpathsea6:amd64 (2021.20210626.59705-1ubuntu0.1) ...
Setting up ruby-webrick (1.7.0-3) ...
Setting up fonts-lmodern (2.004.5-6.1) ...
Setting up fonts-droid-fallback (1:6.0.1r16-1.1build1) ...
Setting up ruby-xmllrpc (0.3.2-1ubuntu0.1) ...
Setting up libsynchronet2:amd64 (2021.20210626.59705-1ubuntu0.1) ...
Setting up libgs9-common (9.55.0~dfsg1-0ubuntu5.5) ...
Setting up teckit (2.5.11+ds1-1) ...
Setting up libpdfbox-java (1:1.8.16-2) ...
Setting up libgs9:amd64 (9.55.0~dfsg1-0ubuntu5.5) ...
Setting up preview-latex-style (12.2-1ubuntu1) ...
Setting up libcommons-parent-java (43-1) ...
Setting up dvisvgm (2.13.1-1) ...
Setting up libcommons-logging-java (1.2-2) ...
Setting up xfonts-utils (1:7.7+6build2) ...
Setting up libptexenc1:amd64 (2021.20210626.59705-1ubuntu0.1) ...
Setting up texlive-binaries (2021.20210626.59705-1ubuntu0.1) ...
update-alternatives: using /usr/bin/xdvi-xaw to provide /usr/bin/xdvi.bin
(xdvi.bin) in auto mode
update-alternatives: using /usr/bin/bibtex.original to provide /usr/bin/bibtex
(bibtex) in auto mode
Setting up lmodern (2.004.5-6.1) ...
Setting up texlive-base (2021.20220204-1) ...
/usr/bin/ucfr
/usr/bin/ucfr
/usr/bin/ucfr
/usr/bin/ucfr
mktexlsr: Updating /var/lib/texmf/ls-R-TEXLIVEDIST...
mktexlsr: Updating /var/lib/texmf/ls-R-TEXMFMAIN...
mktexlsr: Updating /var/lib/texmf/ls-R...
mktexlsr: Done.
tl-paper: setting paper size for dvips to a4:
/var/lib/texmf/dvips/config/config-paper.ps
tl-paper: setting paper size for dvipdfmx to a4:
/var/lib/texmf/dvipdfmx/dvipdfmx-paper.cfg
tl-paper: setting paper size for xdvi to a4: /var/lib/texmf/xdvi/XDvi-paper
tl-paper: setting paper size for pdftex to a4: /var/lib/texmf/tex/generic/tex-
ini-files/pdftexconfig.tex
debconf: unable to initialize frontend: Dialog
debconf: (No usable dialog-like program is installed, so the dialog based
frontend cannot be used. at /usr/share/perl5/Debconf/FrontEnd/Dialog.pm line
78.)
debconf: falling back to frontend: Readline
Setting up tex-gyre (20180621-3.1) ...
Setting up texlive-plain-generic (2021.20220204-1) ...

```

```

Setting up texlive-latex-base (2021.20220204-1) ...
Setting up texlive-latex-recommended (2021.20220204-1) ...
Setting up texlive-pictures (2021.20220204-1) ...
Setting up texlive-fonts-recommended (2021.20220204-1) ...
Setting up tipa (2:1.3-21) ...
Setting up texlive-latex-extra (2021.20220204-1) ...
Setting up texlive-xetex (2021.20220204-1) ...
Setting up rake (13.0.6-2) ...
Setting up libruby3.0:amd64 (3.0.2-7ubuntu2.4) ...
Setting up ruby3.0 (3.0.2-7ubuntu2.4) ...
Setting up ruby (1:3.0~exp1) ...
Setting up ruby-rubygems (3.3.5-2) ...
Processing triggers for man-db (2.10.2-1) ...
Processing triggers for fontconfig (2.13.1-4.2ubuntu5) ...
Processing triggers for libc-bin (2.35-0ubuntu3.4) ...
/sbin/ldconfig.real: /usr/local/lib/libtbbbind_2_0.so.3 is not a symbolic link

/sbin/ldconfig.real: /usr/local/lib/libtbbbind.so.3 is not a symbolic link

/sbin/ldconfig.real: /usr/local/lib/libtbbmalloc.so.2 is not a symbolic link

/sbin/ldconfig.real: /usr/local/lib/libtbb.so.12 is not a symbolic link

/sbin/ldconfig.real: /usr/local/lib/libtbbbind_2_5.so.3 is not a symbolic link

/sbin/ldconfig.real: /usr/local/lib/libtbbmalloc_proxy.so.2 is not a symbolic
link

Processing triggers for tex-common (6.17) ...
debconf: unable to initialize frontend: Dialog
debconf: (No usable dialog-like program is installed, so the dialog based
frontend cannot be used. at /usr/share/perl5/Debconf/FrontEnd/Dialog.pm line
78.)
debconf: falling back to frontend: Readline
Running updmap-sys. This may take some time... done.
Running mktexlsr /var/lib/texmf ... done.
Building format(s) --all.
    This may take some time... done.

```

```
[31]: !jupyter nbconvert --to pdf /content/drive/MyDrive/Colab/ZAL_PG_churn.ipynb
```

```

[NbConvertApp] Converting notebook
/content/drive/MyDrive/Colab/ZAL_PG_churn.ipynb to pdf
[NbConvertApp] Support files will be in ZAL_PG_churn_files/
[NbConvertApp] Making directory ./ZAL_PG_churn_files
[NbConvertApp] Making directory ./ZAL_PG_churn_files
[NbConvertApp] Writing 66646 bytes to notebook.tex
[NbConvertApp] Building PDF

```



```

Traceback (most recent call last):
  File "/usr/local/bin/jupyter-nbconvert", line 8, in <module>
    sys.exit(main())
  File "/usr/local/lib/python3.10/dist-packages/jupyter_core/application.py",
line 280, in launch_instance
    super().launch_instance(argv=argv, **kwargs)
  File "/usr/local/lib/python3.10/dist-
packages/traitlets/config/application.py", line 992, in launch_instance
    app.start()
  File "/usr/local/lib/python3.10/dist-packages/nbconvert/nbconvertapp.py", line
423, in start
    self.convert_notebooks()
  File "/usr/local/lib/python3.10/dist-packages/nbconvert/nbconvertapp.py", line
597, in convert_notebooks
    self.convert_single_notebook(notebook_filename)
  File "/usr/local/lib/python3.10/dist-packages/nbconvert/nbconvertapp.py", line
560, in convert_single_notebook
    output, resources = self.export_single_notebook(
  File "/usr/local/lib/python3.10/dist-packages/nbconvert/nbconvertapp.py", line
488, in export_single_notebook
    output, resources = self.exporter.from_filename(
  File "/usr/local/lib/python3.10/dist-
packages/nbconvert/exporters/exporter.py", line 189, in from_filename
    return self.from_file(f, resources=resources, **kw)
  File "/usr/local/lib/python3.10/dist-
packages/nbconvert/exporters/exporter.py", line 206, in from_file
    return self.from_notebook_node(
  File "/usr/local/lib/python3.10/dist-packages/nbconvert/exporters/pdf.py",
line 194, in from_notebook_node
    self.run_latex(tex_file)
  File "/usr/local/lib/python3.10/dist-packages/nbconvert/exporters/pdf.py",
line 164, in run_latex
    return self.run_command(
  File "/usr/local/lib/python3.10/dist-packages/nbconvert/exporters/pdf.py",
line 111, in run_command
    raise OSError(
OSError: xelatex not found on PATH, if you have not installed xelatex you may
need to do so. Find further instructions at
https://nbconvert.readthedocs.io/en/latest/install.html#installing-tex.

```