410_Final

November 30, 2023

```python
import warnings
warnings.filterwarnings('ignore')

import matplotlib.pyplot as plt

import pandas as pd
import numpy as np
from plotnine import *

from sklearn.linear_model import LinearRegression # Linear Regression Model
from sklearn.preprocessing import StandardScaler #Z-score variables
from sklearn.metrics import mean_squared_error, r2_score, accuracy_score #model
 evaluation
from sklearn.linear_model import LogisticRegression # Linear Regression Model
from sklearn.metrics import accuracy_score, confusion_matrix #model evaluation
from sklearn.cluster import KMeans # k-means
from sklearn.metrics import silhouette_score

from sklearn.model_selection import train_test_split
```

```python
df = pd.read_csv('Cleaned_Subscriber_Information.csv')
df.head()
```

```
     ID Subscription Type Subscription Event Type Purchase Store  \
  0   4           Limited        INITIAL_PURCHASE            App
  1   7           Limited                 RENEWAL            Web
  2  10           Limited                 RENEWAL            App
  3  15           Limited        INITIAL_PURCHASE            App
  4  37           Limited        INITIAL_PURCHASE            App

     Purchase Amount Currency Subscription Start Date Subscription Expiration  \
  0              0.0      NaN                 11/7/19                  2/7/20
  1             79.0      USD                 7/21/19                11/11/19
  2              0.0      NaN                 3/21/20                  3/6/21
  3              0.0      NaN                  9/1/19                  3/1/20
  4              0.0      NaN                12/29/19                 3/29/20
```

```
       Demo User Free Trial User   … Language_NED Language_PAS Language_POL  \
0           Yes              No  …              0            0            0
1           Yes              No  …              0            0            0
2            No             Yes  …              0            0            0
3           Yes              No  …              0            0            0
4           Yes              No  …              0            0            0

    Language_POR Language_RUS Language_SVE  Language_TGL  Language_TUR  \
0              0            0            0             0             0
1              0            0            0             0             0
2              0            0            0             0             0
3              0            0            0             0             0
4              0            0            0             0             0

    Language_URD  Language_VIE
0              0             0
1              0             0
2              0             0
3              0             0
4              0             0

[5 rows x 57 columns]
```

[ ]: `df.shape`

[ ]: (16540, 57)

[ ]: `np.sum(df.isnull())`

[ ]:
```
ID                          0
Subscription Type           0
Subscription Event Type     0
Purchase Store              0
Purchase Amount             0
Currency                 5808
Subscription Start Date     0
Subscription Expiration     0
Demo User                   0
Free Trial User             0
Auto Renew                  0
Country                     0
User Type                   0
Lead Platform               0
Email Subscriber            0
Push Notifications          0
Send Count                  0
Open Count                  0
```

```
Click Count                     0
Unique Open Count               0
Unique Click Count              0
ID_Count                        0
Platform_Count                  0
Activity_Type_Count             0
USD_Adjusted_Amount          5808
Subscription Duration           0
Normalized Purchase Amount      0
Language_ARA                    0
Language_CHI                    0
Language_DAR                    0
Language_DEU                    0
Language_EBR                    0
Language_ENG                    0
Language_ESC                    0
Language_ESP                    0
Language_FAR                    0
Language_FRA                    0
Language_GLE                    0
Language_GRK                    0
Language_HEB                    0
Language_HIN                    0
Language_IND                    0
Language_ITA                    0
Language_JPN                    0
Language_KIS                    0
Language_KOR                    0
Language_LAT                    0
Language_NED                    0
Language_PAS                    0
Language_POL                    0
Language_POR                    0
Language_RUS                    0
Language_SVE                    0
Language_TGL                    0
Language_TUR                    0
Language_URD                    0
Language_VIE                    0
dtype: int64
```

```python
# for i in range(len(df)):
#   if df['Purchase Amount'].iloc[i] == 'NaN':
#     df['Purchase Amount'].iloc[i] = 0

# b
# # df = df[df.PurchaseAmount != 0]
```

```
# df.shape
```

```
(16540, 57)
```

```
df.drop(['ID','Subscription Start Date', 'Subscription Expiration', 'User␣
 ↪Type', 'Currency', 'Subscription Type', 'Unique Open Count', 'Unique Click␣
 ↪Count'], axis='columns', inplace = True)

df.head()
df.shape
```

```
(16540, 49)
```

```
df.dropna(inplace = True)
df.reset_index()
```

```
       index Subscription Event Type Purchase Store   Purchase Amount  \
0          1                 RENEWAL            Web      7.900000e+01
1          5                 RENEWAL            Web      3.597000e+01
2          8        INITIAL_PURCHASE            Web      0.000000e+00
3          9        INITIAL_PURCHASE            App      4.499000e+07
4         10        INITIAL_PURCHASE            Web      1.990000e+02
…         …                    …             …              …
10727  16534        INITIAL_PURCHASE            Web      2.121300e+02
10728  16535                 RENEWAL            Web      3.597000e+01
10729  16536        INITIAL_PURCHASE            Web      1.935772e+01
10730  16537        INITIAL_PURCHASE            Web      2.121300e+02
10731  16538        INITIAL_PURCHASE            Web      4.769115e+01

       Demo User Free Trial User Auto Renew    Country Lead Platform  \
0            Yes              No       Off      Other           App
1             No              No        On      Other       Unknown
2             No              No       Off  US/Canada           Web
3             No              No        On      Other       Unknown
4            Yes              No       Off  US/Canada           App
…             …               …         …          …             …
10727         No              No       Off      Other       Unknown
10728         No              No        On      Other       Unknown
10729         No              No       Off     Europe           Web
10730         No              No       Off      Other       Unknown
10731         No              No       Off      Other       Unknown

       Email Subscriber  … Language_NED  Language_PAS  Language_POL  \
0                   Yes  …            0             0             0
1                    No  …            0             0             0
2                   Yes  …            0             0             0
```

4

```
3               No  …          0              0              0
4               Yes …          0              0              0
…               …   …          …              …              …
10727           No  …          0              0              0
10728           No  …          0              0              0
10729           Yes …          0              0              0
10730           No  …          0              0              0
10731           No  …          0              0              0

       Language_POR  Language_RUS  Language_SVE  Language_TGL  Language_TUR  \
0                 0             0             0             0             0
1                 0             0             0             0             0
2                 0             0             0             0             0
3                 0             0             0             0             0
4                 0             0             0             0             0
…                 …             …             …             …             …
10727             0             0             0             0             0
10728             0             0             0             0             0
10729             0             0             0             0             0
10730             0             0             0             0             0
10731             0             0             0             0             0

       Language_URD  Language_VIE
0                 0             0
1                 0             0
2                 0             0
3                 0             0
4                 0             0
…                 …             …
10727             0             0
10728             0             0
10729             0             0
10730             0             0
10731             0             0

[10732 rows x 50 columns]
```

```python
for i in range(len(df)):
    if df['Push Notifications'].iloc[i] == 'Yes':
        df['Push Notifications'].iloc[i] = 1
    else:
        df['Push Notifications'].iloc[i] = 0
    if df['Email Subscriber'].iloc[i] == 'Yes':
        df['Email Subscriber'].iloc[i] = 1
    else:
        df['Email Subscriber'].iloc[i] = 0
    if df['Free Trial User'].iloc[i] == 'Yes':
```

```
            df['Free Trial User'].iloc[i] = 1
        else:
            df['Free Trial User'].iloc[i] = 0
        if df['Demo User'].iloc[i] == 'Yes':
            df['Demo User'].iloc[i] = 1
        else:
            df['Demo User'].iloc[i] = 0
        if df['Lead Platform'].iloc[i] == 'App':
            df['Lead Platform'].iloc[i] = 1
        else:
            df['Lead Platform'].iloc[i] = 0
        if df['Auto Renew'].iloc[i] == 'On':
            df['Auto Renew'].iloc[i] = 1
        else:
            df['Auto Renew'].iloc[i] = 0
        if df['Purchase Store'].iloc[i] == 'App':
            df['Purchase Store'].iloc[i] = 1
        else:
            df['Purchase Store'].iloc[i] = 0

df.reset_index()
```

```
[ ]:         index Subscription Event Type Purchase Store  Purchase Amount  \
    0             1                  RENEWAL              0     7.900000e+01
    1             5                  RENEWAL              0     3.597000e+01
    2             8        INITIAL_PURCHASE              0     0.000000e+00
    3             9        INITIAL_PURCHASE              1     4.499000e+07
    4            10        INITIAL_PURCHASE              0     1.990000e+02
    ...         ...                      ...            ...              ...
    10727     16534        INITIAL_PURCHASE              0     2.121300e+02
    10728     16535                  RENEWAL              0     3.597000e+01
    10729     16536        INITIAL_PURCHASE              0     1.935772e+01
    10730     16537        INITIAL_PURCHASE              0     2.121300e+02
    10731     16538        INITIAL_PURCHASE              0     4.769115e+01

           Demo User Free Trial User Auto Renew      Country Lead Platform  \
    0               1               0          0        Other             1
    1               0               0          1        Other             0
    2               0               0          0    US/Canada             0
    3               0               0          1        Other             0
    4               1               0          0    US/Canada             1
    ...           ...             ...        ...          ...           ...
    10727           0               0          0        Other             0
    10728           0               0          1        Other             0
    10729           0               0          0       Europe             0
    10730           0               0          0        Other             0
    10731           0               0          0        Other             0
```

```
         Email Subscriber  …  Language_NED  Language_PAS  Language_POL  \
0                       1   …             0             0             0
1                       0   …             0             0             0
2                       1   …             0             0             0
3                       0   …             0             0             0
4                       1   …             0             0             0
…                     …   …           …           …           …
10727                   0   …             0             0             0
10728                   0   …             0             0             0
10729                   1   …             0             0             0
10730                   0   …             0             0             0
10731                   0   …             0             0             0

       Language_POR  Language_RUS  Language_SVE  Language_TGL  Language_TUR  \
0                 0             0             0             0             0
1                 0             0             0             0             0
2                 0             0             0             0             0
3                 0             0             0             0             0
4                 0             0             0             0             0
…               …           …           …           …           …
10727             0             0             0             0             0
10728             0             0             0             0             0
10729             0             0             0             0             0
10730             0             0             0             0             0
10731             0             0             0             0             0

       Language_URD  Language_VIE
0                 0             0
1                 0             0
2                 0             0
3                 0             0
4                 0             0
…               …           …
10727             0             0
10728             0             0
10729             0             0
10730             0             0
10731             0             0

[10732 rows x 50 columns]
```

```python
df = pd.get_dummies(df, columns=['Subscription Event Type'])
df = pd.get_dummies(df, columns = ['Country'])
df.head()
```

7

```
[ ]:      Purchase Store  Purchase Amount  Demo User  Free Trial User  Auto Renew  \
    1                  0            79.00          1                0           0
    5                  0            35.97          0                0           1
    8                  0             0.00          0                0           0
    9                  1      44990000.00          0                0           1
    10                 0           199.00          1                0           0

        Lead Platform  Email Subscriber  Push Notifications  Send Count  Open Count  \
    1               1                 1                   1           2           0
    5               0                 0                   0           3           1
    8               0                 1                   1          17           0
    9               0                 0                   0          10           1
    10              1                 1                   1           6           1

        …  Language_SVE  Language_TGL  Language_TUR  Language_URD  Language_VIE  \
    1   …             0             0             0             0             0
    5   …             0             0             0             0             0
    8   …             0             0             0             0             0
    9   …             0             0             0             0             0
    10  …             0             0             0             0             0

        Subscription Event Type_INITIAL_PURCHASE  Subscription Event Type_RENEWAL  \
    1                                          0                                1
    5                                          0                                1
    8                                          1                                0
    9                                          1                                0
    10                                         1                                0

        Country_Europe  Country_Other  Country_US/Canada
    1                0              1                  0
    5                0              1                  0
    8                0              0                  1
    9                0              1                  0
    10               0              0                  1

    [5 rows x 52 columns]
```

```
[ ]: df.shape
```

```
[ ]: (10732, 52)
```

```
[ ]: df = df[df['Purchase Amount'] <= 1000]
```

```
[ ]: df['Email Open %'] = df['Open Count'] / df['Send Count']
     df['Email Click %'] = df['Click Count'] / df['Send Count']
```

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

```
[ ]:      Purchase Store  Purchase Amount Demo User Free Trial User Auto Renew  \
      1                0            79.00         1               0           0
      5                0            35.97         0               0           1
      8                0             0.00         0               0           0
      10               0           199.00         1               0           0
      12               0            34.64         0               0           0

          Lead Platform Email Subscriber Push Notifications  Send Count  Open Count  \
      1               1                1                  1           2           0
      5               0                0                  0           3           1
      8               0                1                  1          17           0
      10              1                1                  1           6           1
      12              0                0                  0          10           0

          …  Language_TUR  Language_URD  Language_VIE  \
      1   …             0             0             0
      5   …             0             0             0
      8   …             0             0             0
      10  …             0             0             0
      12  …             0             0             0

          Subscription Event Type_INITIAL_PURCHASE  Subscription Event Type_RENEWAL  \
      1                                          0                                1
      5                                          0                                1
      8                                          1                                0
      10                                         1                                0
      12                                         1                                0

          Country_Europe  Country_Other  Country_US/Canada  Email Open %  \
      1                0              1                  0      0.000000
      5                0              1                  0      0.333333
      8                0              0                  1      0.000000
      10               0              0                  1      0.166667
      12               0              1                  0      0.000000

          Email Click %
      1             0.0
      5             0.0
      8             0.0
      10            0.0
      12            0.0

      [5 rows x 54 columns]
```

```
[ ]:  # ggplot(df, aes(x='Country', y='Purchase Amount', fill = 'Country')) +
      ↪stat_summary(geom='bar')
```

```
--------------------------------------------------------------------------------
NameError                                  Traceback (most recent call last)
/usr/local/lib/python3.10/dist-packages/plotnine/mapping/evaluation.py in␣
 ↪evaluate(aesthetics, data, env)
    227                   try:
--> 228                       new_val = env.eval(col, inner_namespace=data)
    229                   except Exception as e:

/usr/local/lib/python3.10/dist-packages/patsy/eval.py in eval(self, expr,␣
 ↪source_name, inner_namespace)
    168         code = compile(expr, source_name, "eval", self.flags, False)
--> 169         return eval(code, {}, VarLookupDict([inner_namespace]

    170                                             + self._namespaces))

<string> in <module>

NameError: name 'Country' is not defined

During handling of the above exception, another exception occurred:

PlotnineError                              Traceback (most recent call last)
/usr/local/lib/python3.10/dist-packages/IPython/core/formatters.py in␣
 ↪__call__(self, obj)
    700                   type_pprinters=self.type_printers,
    701                   deferred_pprinters=self.deferred_printers)
--> 702           printer.pretty(obj)
    703           printer.flush()
    704           return stream.getvalue()

/usr/local/lib/python3.10/dist-packages/IPython/lib/pretty.py in pretty(self,␣
 ↪obj)
    392                       if cls is not object \
    393                           and callable(cls.__dict__.
 ↪get('__repr__')):
--> 394                           return _repr_pprint(obj, self, cycle)
    395
    396                   return _default_pprint(obj, self, cycle)

/usr/local/lib/python3.10/dist-packages/IPython/lib/pretty.py in␣
 ↪_repr_pprint(obj, p, cycle)
    698       """A pprint that just redirects to the normal repr function."""
    699       # Find newlines and replace them with p.break_()
--> 700       output = repr(obj)
    701       lines = output.splitlines()
    702       with p.group():
```

```
/usr/local/lib/python3.10/dist-packages/plotnine/ggplot.py in __repr__(self)
    112             Print/show the plot
    113             """
--> 114             figure = self.draw(show=True)
    115
    116             dpi = figure.get_dpi()


/usr/local/lib/python3.10/dist-packages/plotnine/ggplot.py in draw(self, show)
    222             self = deepcopy(self)
    223             with plot_context(self, show=show):
--> 224                 self._build()
    225
    226                 # setup


/usr/local/lib/python3.10/dist-packages/plotnine/ggplot.py in _build(self)
    312             # Compute aesthetics to produce data with generalised
    313             # variable names
--> 314             layers.compute_aesthetics(self)
    315
    316             # Transform data using all scales


/usr/local/lib/python3.10/dist-packages/plotnine/layer.py in␣
 ↪compute_aesthetics(self, plot)
    465     def compute_aesthetics(self, plot: Ggplot):
    466         for l in self:
--> 467             l.compute_aesthetics(plot)
    468
    469     def compute_statistic(self, layout: Layout):


/usr/local/lib/python3.10/dist-packages/plotnine/layer.py in␣
 ↪compute_aesthetics(self, plot)
    258         expression evaluation are  made in here
    259         """
--> 260         evaled = evaluate(self.mapping._starting, self.data, plot.
 ↪environment)
    261         evaled_aes = aes(**{str(col): col for col in evaled})
    262         plot.scales.add_defaults(evaled, evaled_aes)


/usr/local/lib/python3.10/dist-packages/plotnine/mapping/evaluation.py in␣
 ↪evaluate(aesthetics, data, env)
    228                     new_val = env.eval(col, inner_namespace=data)
    229                 except Exception as e:
--> 230                     raise PlotnineError(_TPL_EVAL_FAIL.format(ae, col,␣
 ↪str(e)))
    231
    232                 try:
```

```
[ ]: # ggplot(df, aes(x='Country', y='Subscription Duration', fill = 'Country')) +⌴
     ↪stat_summary(geom='bar')
```



```
[ ]: <Figure Size: (640 x 480)>
```
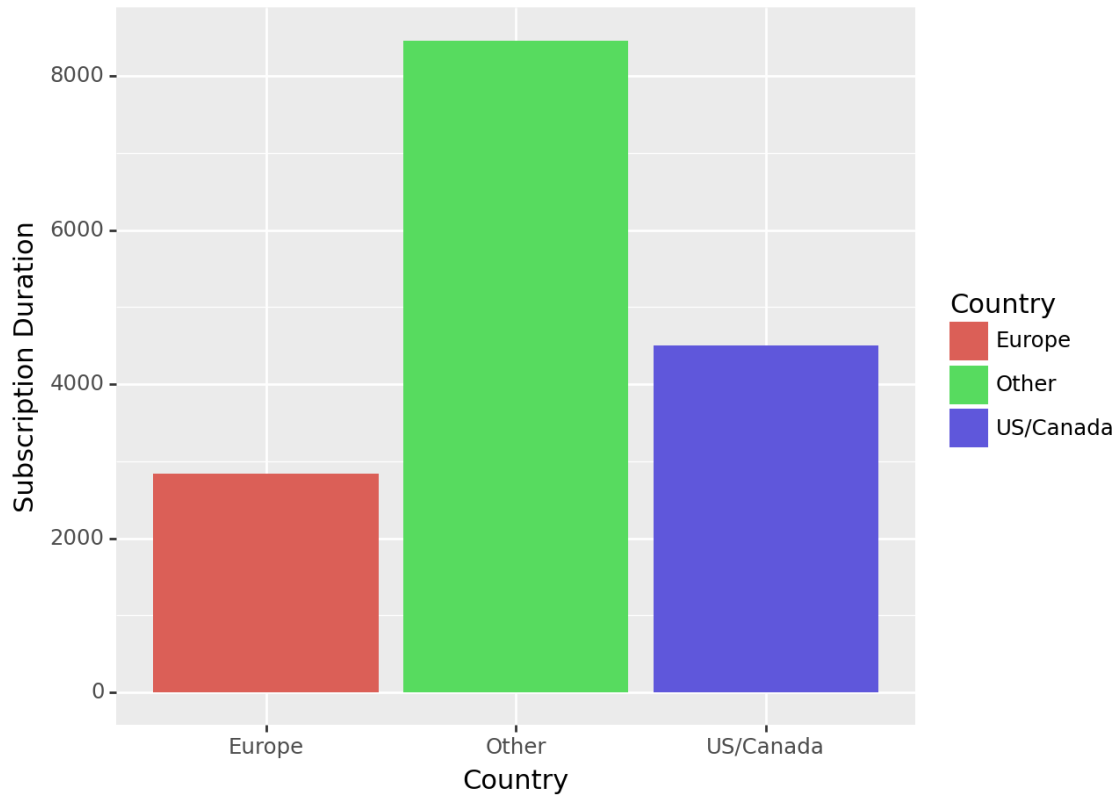
```
[ ]: df['Auto Renew'] = df['Auto Renew'].astype(float)
```

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

```
[ ]:    Purchase Store  Purchase Amount Demo User Free Trial User  Auto Renew  \
     1               0            79.00         1               0         0.0
     5               0            35.97         0               0         1.0
     8               0             0.00         0               0         0.0
     10              0           199.00         1               0         0.0
     12              0            34.64         0               0         0.0

        Lead Platform Email Subscriber Push Notifications  Send Count  Open Count  \
```

```
    1               1              1                1          2          0
    5               0              0                0          3          1
    8               0              1                1         17          0
   10               1              1                1          6          1
   12               0              0                0         10          0

        …  Language_TUR  Language_URD  Language_VIE  \
    1   …             0             0             0
    5   …             0             0             0
    8   …             0             0             0
   10   …             0             0             0
   12   …             0             0             0

        Subscription Event Type_INITIAL_PURCHASE  Subscription Event Type_RENEWAL  \
    1                                          0                                1
    5                                          0                                1
    8                                          1                                0
   10                                          1                                0
   12                                          1                                0

        Country_Europe  Country_Other  Country_US/Canada  Email Open %  \
    1                0              1                  0      0.000000
    5                0              1                  0      0.333333
    8                0              0                  1      0.000000
   10                0              0                  1      0.166667
   12                0              1                  0      0.000000

        Email Click %
    1             0.0
    5             0.0
    8             0.0
   10             0.0
   12             0.0

    [5 rows x 54 columns]
```

```python
#Instantiate Standard scalar object for Kmeans model, collate predictors
zs = StandardScaler()
predictors = ["Email Open %", "Auto Renew", 'Subscription Duration', 'Purchase␣
 ↪Amount']

# assign X variable to three inflation metrics, zscore and fit
X = df[predictors]
X[predictors] = zs.fit_transform(X[predictors])

# Set up for CV (tuning/optimizing our hyperparamter)
ks = [2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20]
```

```
sils = []

# For loop iterates through each k-value, fits a model to it and records the␣
 ↪silhouette scores of each associative model
for k in ks:
  km = KMeans(n_clusters = k)
  km.fit(X[predictors])

  sils.append(silhouette_score(X[predictors], km.predict(X[predictors])))

# create df for ggplot
sil_df = pd.DataFrame({"K": ks,
                       "silhouette": sils})

#plot of hyperparameter tuning
(ggplot(sil_df, aes(x = "K", y = "sils")) + geom_point() +
geom_line() +
theme_minimal() +
ggtitle("Silhouette for Different Ks"))
```

Silhouette for Different Ks

```
[ ]: <Figure Size: (640 x 480)>
```
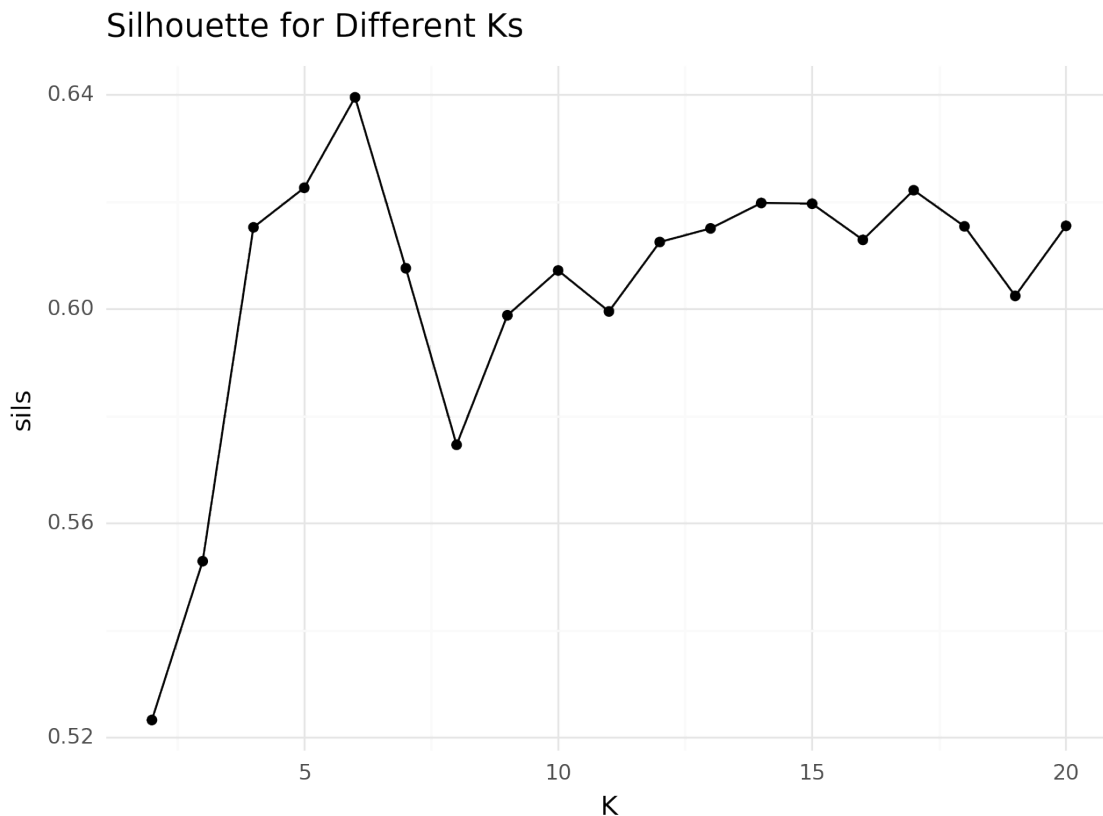
```
[ ]: #FINAL KMeans Model

     # Create new KMeans object, set hyperparameter to 3, fit the model and assign␣
      ↪membership
     km1 = KMeans(n_clusters = 4)
     km1.fit(X[predictors])
     membership = km1.predict(X[predictors])

     X['cluster'] = membership
     print("Silhouette Score (n_clusters = 4): ",sils[3])
     # ggplot(k_df, aes(x = 'USDAdjusted', y = 'Subscription Length (days)', fill =␣
      ↪'cluster')) + geom_point() + coord_cartesian(xlim=(0, 400), ylim=(0, 1500))
```

    Silhouette Score (n_clusters = 4): 0.6226684721604728

```
[ ]: k_df = df[predictors]
     k_df.head()
```

```
[ ]:     Email Open %  Auto Renew  Subscription Duration  Purchase Amount
     1       0.000000         0.0                    113            79.00
     5       0.333333         1.0                    274            35.97
     8       0.000000         0.0                    182             0.00
     10      0.166667         0.0                  28789           199.00
     12      0.000000         0.0                     91            34.64
```

```
[ ]: k_df['Cluster'] = km1.labels_

     # Calculate the average of each predictor for each cluster
     cluster_averages = k_df.groupby('Cluster').mean()
     print()
     print("Silhouette Score (Four Clusters): ",sils[3], '\n\n', cluster_averages)
```

    Silhouette Score (Four Clusters): 0.6223271836259553

Cluster	Email Open %	Auto Renew	Subscription Duration	Purchase Amount
0	0.190602	0.001397	28709.432231	203.743251
1	0.075564	1.000000	304.995298	51.891677
2	0.039036	0.000000	1002.154324	29.861447
3	0.878711	0.546701	2290.696658	57.404074

```
[ ]: ggplot(k_df, aes(x = 'Purchase Amount', y = 'Subscription Duration', fill =␣
      ↪'Cluster')) + geom_point() + coord_cartesian(xlim=(0, 400), ylim=(0, 1500))
```

[ ]: <Figure Size: (640 x 480)>
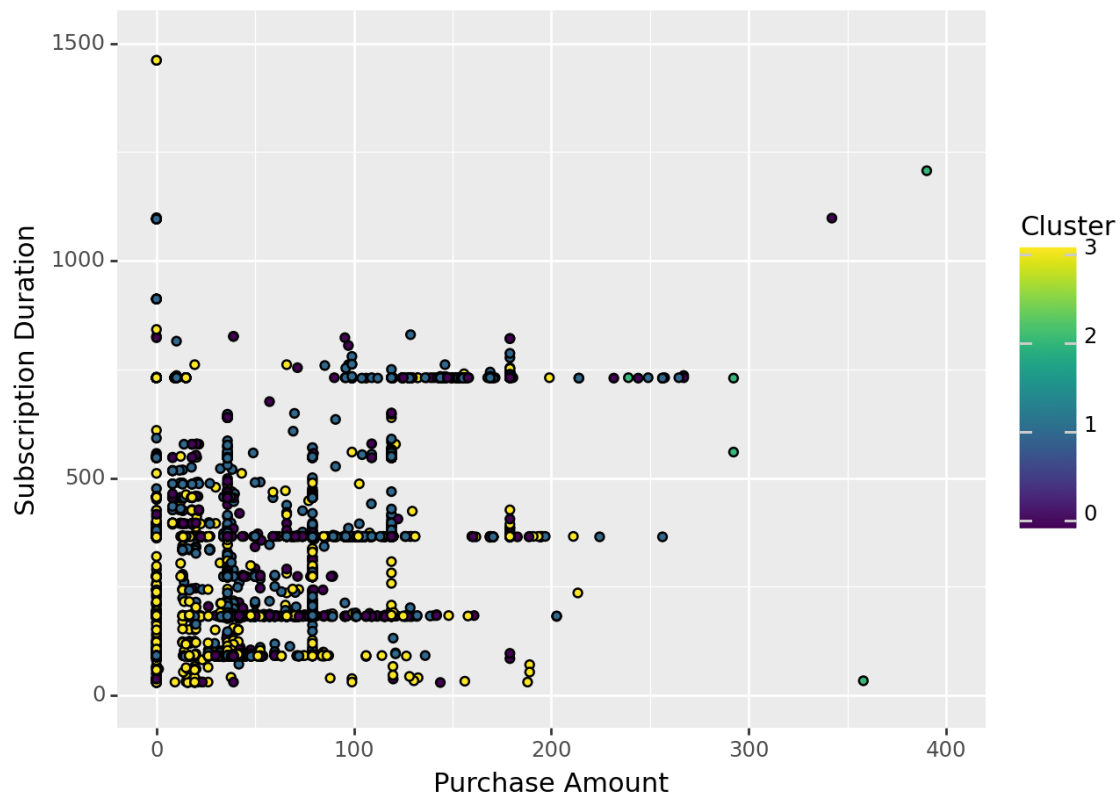
[ ]: `silhouette_score(k_df, membership)`

[ ]: 0.0761439246325425

[ ]: `df.info()`

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 10042 entries, 1 to 16538
Data columns (total 55 columns):
 #   Column                Non-Null Count  Dtype
---  ------                --------------  -----
 0   Purchase Store        10042 non-null  object
 1   Purchase Amount       10042 non-null  float64
 2   Demo User             10042 non-null  object
 3   Free Trial User       10042 non-null  object
 4   Auto Renew            10042 non-null  float64
 5   Lead Platform         10042 non-null  object
 6   Email Subscriber      10042 non-null  object
 7   Push Notifications    10042 non-null  object
```

```
8    Send Count                                  10042 non-null  int64
9    Open Count                                  10042 non-null  int64
10   Click Count                                 10042 non-null  int64
11   ID_Count                                    10042 non-null  int64
12   Platform_Count                              10042 non-null  int64
13   Activity_Type_Count                         10042 non-null  int64
14   USD_Adjusted_Amount                         10042 non-null  float64
15   Subscription Duration                       10042 non-null  int64
16   Normalized Purchase Amount                  10042 non-null  float64
17   Language_ARA                                10042 non-null  int64
18   Language_CHI                                10042 non-null  int64
19   Language_DAR                                10042 non-null  int64
20   Language_DEU                                10042 non-null  int64
21   Language_EBR                                10042 non-null  int64
22   Language_ENG                                10042 non-null  int64
23   Language_ESC                                10042 non-null  int64
24   Language_ESP                                10042 non-null  int64
25   Language_FAR                                10042 non-null  int64
26   Language_FRA                                10042 non-null  int64
27   Language_GLE                                10042 non-null  int64
28   Language_GRK                                10042 non-null  int64
29   Language_HEB                                10042 non-null  int64
30   Language_HIN                                10042 non-null  int64
31   Language_IND                                10042 non-null  int64
32   Language_ITA                                10042 non-null  int64
33   Language_JPN                                10042 non-null  int64
34   Language_KIS                                10042 non-null  int64
35   Language_KOR                                10042 non-null  int64
36   Language_LAT                                10042 non-null  int64
37   Language_NED                                10042 non-null  int64
38   Language_PAS                                10042 non-null  int64
39   Language_POL                                10042 non-null  int64
40   Language_POR                                10042 non-null  int64
41   Language_RUS                                10042 non-null  int64
42   Language_SVE                                10042 non-null  int64
43   Language_TGL                                10042 non-null  int64
44   Language_TUR                                10042 non-null  int64
45   Language_URD                                10042 non-null  int64
46   Language_VIE                                10042 non-null  int64
47   Subscription Event Type_INITIAL_PURCHASE    10042 non-null  uint8
48   Subscription Event Type_RENEWAL             10042 non-null  uint8
49   Country_Europe                              10042 non-null  uint8
50   Country_Other                               10042 non-null  uint8
51   Country_US/Canada                           10042 non-null  uint8
52   Email Open %                                10042 non-null  float64
53   Email Impression %                          10042 non-null  float64
54   Cluster                                     10042 non-null  int32
dtypes: float64(6), int32(1), int64(37), object(6), uint8(5)
```

```
memory usage: 4.2+ MB
```

```python
import pandas as pd
import matplotlib.pyplot as plt
from sklearn.ensemble import RandomForestRegressor
from sklearn.model_selection import train_test_split

predictors = ['Subscription Event Type_INITIAL_PURCHASE', 'Subscription Event␣
 ↪Type_RENEWAL', 'Demo User', 'Free Trial User',
                'Auto Renew', 'Email Subscriber', 'Push Notifications',␣
 ↪'Email Open %','ID_Count','Activity_Type_Count']

# Assuming df and predictors are defined as in the previous example

X = df[predictors]
y = df['Purchase Amount']

# Splitting the data
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2,␣
 ↪random_state=42)

# Training the model
rf_model = RandomForestRegressor(n_estimators=500, random_state=42)
rf_model.fit(X_train, y_train)

# Extracting feature importances
feature_importances = rf_model.feature_importances_

# Creating a DataFrame for feature importances
importances_df = pd.DataFrame({'Feature': predictors, 'Importance':␣
 ↪feature_importances})

# Sorting the DataFrame by importance
importances_df = importances_df.sort_values(by='Importance', ascending=False)

# Plotting
plt.figure(figsize=(10, 6))
plt.barh(importances_df['Feature'], importances_df['Importance'])
plt.xlabel('Importance')
plt.ylabel('Feature')
plt.title('Feature Importances in Predicting Purchase Amount')
plt.gca().invert_yaxis()  # To display the highest importance at the top
plt.show()
```

```python
importances_df
```

```
[ ]:                                    Feature   Importance
     7                               Email Open %     0.261219
     8                                  ID_Count      0.214252
     9                         Activity_Type_Count    0.125590
     4                                Auto Renew      0.123324
     5                            Email Subscriber    0.075059
     6                           Push Notifications   0.064074
     1        Subscription Event Type_RENEWAL         0.040134
     2                                 Demo User      0.038420
     3                            Free Trial User     0.029263
     0   Subscription Event Type_INITIAL_PURCHASE     0.028664
```

```python
importances_df['Cumulative Importance'] = importances_df['Importance'].cumsum()

plt.figure(figsize=(10, 6))
plt.plot(importances_df['Feature'], importances_df['Cumulative Importance'],
    'b-')
plt.xticks(rotation=45, ha='right')
plt.xlabel('Feature')
plt.ylabel('Cumulative Importance')
plt.title('Cumulative Feature Importances')
plt.show()
```
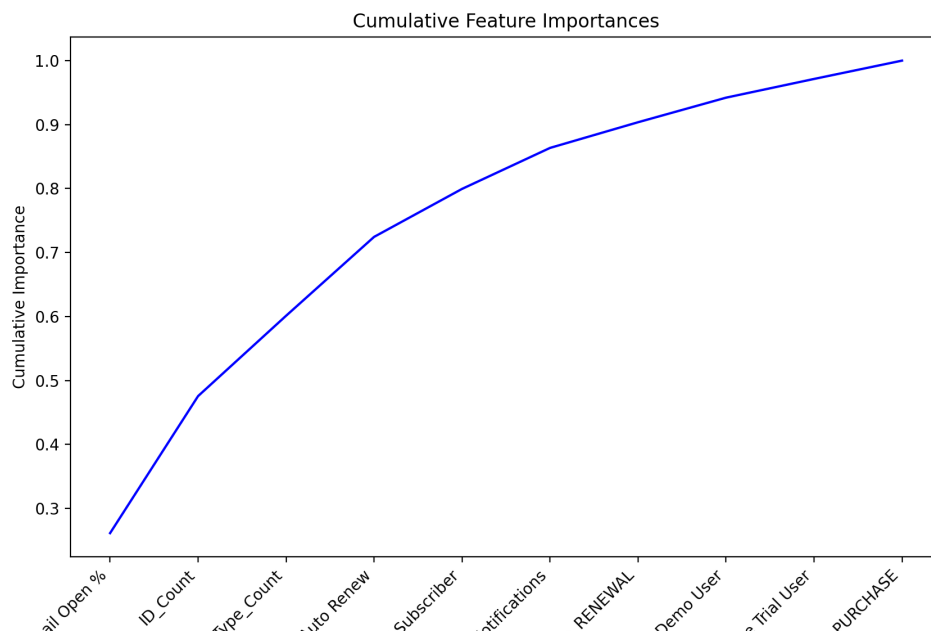

[ ]:

```
[ ]:  # from sklearn import plot_partial_dependence

      # # Selecting top N features (e.g., top 3 for simplicity)
      # top_features = importances_df['Feature'][:3]

      # fig, ax = plt.subplots(figsize=(10, 6))
      # plot_partial_dependence(rf_model, features=top_features, X=X_train, ax=ax)
      # plt.show()
```

```
      ---------------------------------------------------------------------------
      ImportError                               Traceback (most recent call last)
      <ipython-input-41-1c2d6a675269> in <cell line: 1>()
      ----> 1 from sklearn import plot_partial_dependence
            2
            3 # Selecting top N features (e.g., top 3 for simplicity)
            4 top_features = importances_df['Feature'][:3]
            5

      ImportError: cannot import name 'plot_partial_dependence' from 'sklearn' (/usr/
        ↪local/lib/python3.10/dist-packages/sklearn/__init__.py)

      ---------------------------------------------------------------------------
      NOTE: If your import is failing due to a missing package, you can
      manually install dependencies using either !pip or !apt.
      To view examples of installing some common dependencies, click the
      "Open Examples" button below.
      ---------------------------------------------------------------------------
```

```
[ ]:  from sklearn.metrics import mean_squared_error, r2_score

      # Predicting on the test set
      y_pred = rf_model.predict(X_test)

      # Calculating R-squared
      r_squared = r2_score(y_test, y_pred)
      print(f"R-squared: {r_squared}")

      # Calculating Mean Squared Error
```

```
mse = mean_squared_error(y_test, y_pred)
print(f"Mean Squared Error: {mse}")
```

R-squared: 0.2233741036283422
Mean Squared Error: 4607.406134153336

```
[ ]: import pandas as pd
     import matplotlib.pyplot as plt
     from sklearn.ensemble import RandomForestRegressor
     from sklearn.model_selection import train_test_split

     predictors = ['Subscription Event Type_INITIAL_PURCHASE', 'Subscription Event␣
      ↪Type_RENEWAL', 'Demo User', 'Free Trial User',
                       'Auto Renew', 'Email Subscriber', 'Push Notifications',␣
      ↪'Email Open %','ID_Count','Activity_Type_Count']

     # Assuming df and predictors are defined as in the previous example

     X = df[predictors]
     y = df['Subscription Duration']

     # Splitting the data
     X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2,␣
      ↪random_state=42)

     # Training the model
     rf_model = RandomForestRegressor(n_estimators=500, random_state=42)
     rf_model.fit(X_train, y_train)

     # Extracting feature importances
     feature_importances = rf_model.feature_importances_

     # Creating a DataFrame for feature importances
     importances_df = pd.DataFrame({'Feature': predictors, 'Importance':␣
      ↪feature_importances})

     # Sorting the DataFrame by importance
     importances_df = importances_df.sort_values(by='Importance', ascending=False)

     # Plotting
     plt.figure(figsize=(10, 6))
     plt.barh(importances_df['Feature'], importances_df['Importance'])
     plt.xlabel('Importance')
     plt.ylabel('Feature')
     plt.title('Feature Importances in Predicting Subscription Duration')
     plt.gca().invert_yaxis()  # To display the highest importance at the top
     plt.show()
```

```python
# mse = mean_squared_error(y_test, y_pred)
# print(f"Mean Squared Error: {mse}")
```

```
[ ]: importances_df
```

```python
from sklearn.metrics import mean_squared_error, r2_score

# Predicting on the test set
y_pred = rf_model.predict(X_test)

# Calculating R-squared
r_squared = r2_score(y_test, y_pred)
print(f"R-squared: {r_squared}")

# Calculating Mean Squared Error
mse = mean_squared_error(y_test, y_pred)
print(f"Mean Squared Error: {mse}")
```

```
R-squared: 0.3668368024253542
Mean Squared Error: 89418568.10766277
```

```python
df.head()
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 10042 entries, 1 to 16538
Data columns (total 55 columns):
 #   Column                 Non-Null Count  Dtype
---  ------                 --------------  -----
 0   Purchase Store         10042 non-null  object
 1   Purchase Amount        10042 non-null  float64
 2   Demo User              10042 non-null  object
 3   Free Trial User        10042 non-null  object
 4   Auto Renew             10042 non-null  float64
 5   Lead Platform          10042 non-null  object
 6   Email Subscriber       10042 non-null  object
 7   Push Notifications     10042 non-null  object
 8   Send Count             10042 non-null  int64
 9   Open Count             10042 non-null  int64
 10  Click Count            10042 non-null  int64
 11  ID_Count               10042 non-null  int64
 12  Platform_Count         10042 non-null  int64
 13  Activity_Type_Count    10042 non-null  int64
 14  USD_Adjusted_Amount    10042 non-null  float64
 15  Subscription Duration  10042 non-null  int64
```

```
16   Normalized Purchase Amount                10042 non-null   float64
17   Language_ARA                              10042 non-null   int64
18   Language_CHI                              10042 non-null   int64
19   Language_DAR                              10042 non-null   int64
20   Language_DEU                              10042 non-null   int64
21   Language_EBR                              10042 non-null   int64
22   Language_ENG                              10042 non-null   int64
23   Language_ESC                              10042 non-null   int64
24   Language_ESP                              10042 non-null   int64
25   Language_FAR                              10042 non-null   int64
26   Language_FRA                              10042 non-null   int64
27   Language_GLE                              10042 non-null   int64
28   Language_GRK                              10042 non-null   int64
29   Language_HEB                              10042 non-null   int64
30   Language_HIN                              10042 non-null   int64
31   Language_IND                              10042 non-null   int64
32   Language_ITA                              10042 non-null   int64
33   Language_JPN                              10042 non-null   int64
34   Language_KIS                              10042 non-null   int64
35   Language_KOR                              10042 non-null   int64
36   Language_LAT                              10042 non-null   int64
37   Language_NED                              10042 non-null   int64
38   Language_PAS                              10042 non-null   int64
39   Language_POL                              10042 non-null   int64
40   Language_POR                              10042 non-null   int64
41   Language_RUS                              10042 non-null   int64
42   Language_SVE                              10042 non-null   int64
43   Language_TGL                              10042 non-null   int64
44   Language_TUR                              10042 non-null   int64
45   Language_URD                              10042 non-null   int64
46   Language_VIE                              10042 non-null   int64
47   Subscription Event Type_INITIAL_PURCHASE  10042 non-null   uint8
48   Subscription Event Type_RENEWAL           10042 non-null   uint8
49   Country_Europe                            10042 non-null   uint8
50   Country_Other                             10042 non-null   uint8
51   Country_US/Canada                         10042 non-null   uint8
52   Email Open %                              10042 non-null   float64
53   Email Impression %                        10042 non-null   float64
54   Cluster                                   10042 non-null   int32
dtypes: float64(6), int32(1), int64(37), object(6), uint8(5)
memory usage: 4.2+ MB
```

```python
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.ensemble import RandomForestClassifier
from sklearn.metrics import accuracy_score, classification_report
#'Purchase Amount'
```

```python
# Assuming 'df' is your DataFrame
# , 'Subscription Duration'
predictors = ['Purchase Amount', 'Subscription Duration', 'Demo User', 'Free␣
 ↪Trial User', 'Email Open %', 'ID_Count',
              'Activity_Type_Count', 'Push Notifications', 'Click Count']
X = df[predictors]
y = df['Auto Renew']  # Assuming 'Auto Renew' is already in numeric format

# Splitting data into training and testing sets
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2,␣
 ↪random_state=42)

# Creating and training the Random Forest Classifier
rf_classifier = RandomForestClassifier(n_estimators=500, random_state=42)
rf_classifier.fit(X_train, y_train)

# Predicting on the test set
y_pred = rf_classifier.predict(X_test)

# Evaluating the model
accuracy = accuracy_score(y_test, y_pred)
print(f"Accuracy: {accuracy}",'\n')
print(classification_report(y_test, y_pred))
```

```python
[ ]: # Feature Importance Plot
     import matplotlib.pyplot as plt
     import pandas as pd

     # Assuming rf_classifier is your trained RandomForestClassifier model
     feature_importances = rf_classifier.feature_importances_
     features = pd.DataFrame({'Feature': predictors, 'Importance':␣
      ↪feature_importances}).sort_values(by='Importance', ascending=False)

     plt.figure(figsize=(13, 6))
     plt.barh(features['Feature'], features['Importance'])
     plt.xlabel('Importance')
     plt.ylabel('Feature')
     plt.title('Feature Importances in Predicting Customer Churn')
     plt.gca().invert_yaxis()
     plt.show()
```

```python
[ ]: features
```

```python
[ ]: # Confusion Matrix
     from sklearn.metrics import confusion_matrix
     import seaborn as sns
```

```
cm = confusion_matrix(y_test, y_pred)
plt.figure(figsize=(8, 6))
sns.heatmap(cm, annot=True, fmt='d', cmap='Blues')
plt.title('Confusion Matrix')
plt.ylabel('Actual label')
plt.xlabel('Predicted label')
plt.show()
```
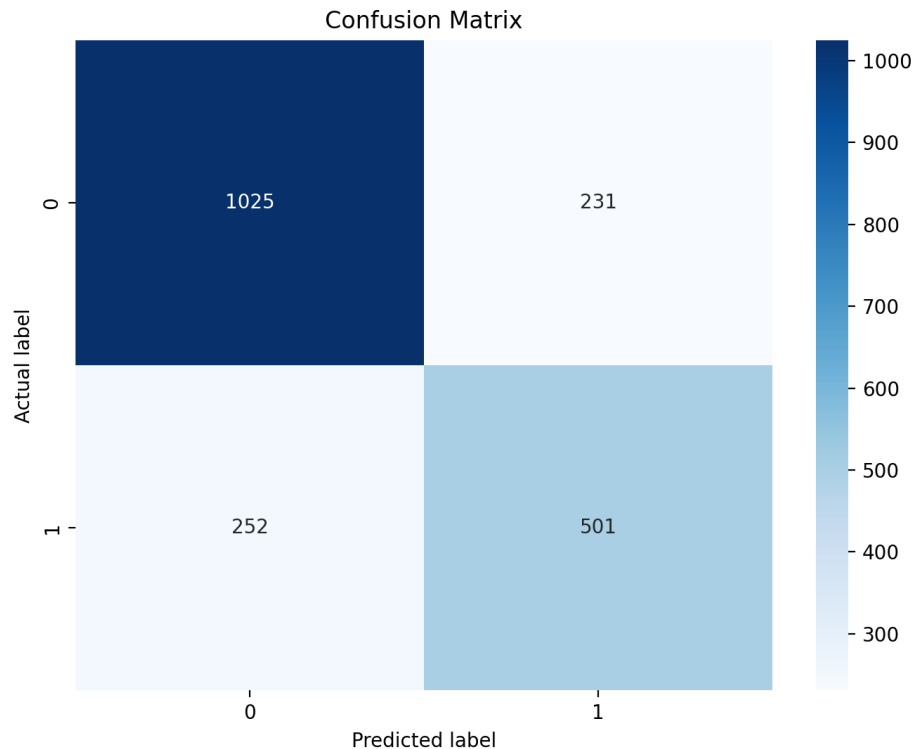


```
[ ]: from sklearn.metrics import roc_curve, auc

     # Compute ROC curve and ROC area for each class
     fpr, tpr, _ = roc_curve(y_test, rf_classifier.predict_proba(X_test)[:, 1])
     roc_auc = auc(fpr, tpr)

     plt.figure()
     plt.plot(fpr, tpr, color='darkorange', lw=2, label=f'ROC curve (area = {roc_auc:
       →.2f})')
     plt.plot([0, 1], [0, 1], color='navy', lw=2, linestyle='--')
     plt.xlim([0.0, 1.0])
     plt.ylim([0.0, 1.05])
```

```python
plt.xlabel('False Positive Rate')
plt.ylabel('True Positive Rate')
plt.title('Receiver Operating Characteristic')
plt.legend(loc='lower right')
plt.show()
```
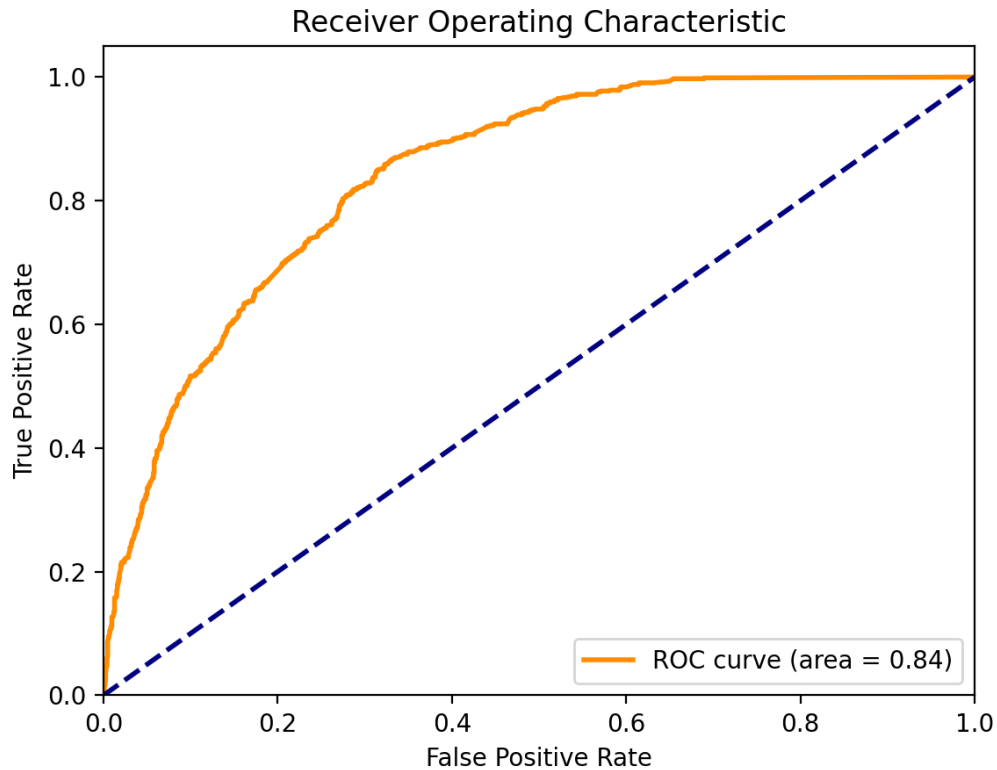


```python
# Download as PDF with Jupyter Notebook

# doesn't show this cells output when downloading PDF
!pip install gwpy &> /dev/null

# installing necessary files
!apt-get install texlive texlive-xetex texlive-latex-extra pandoc
!sudo apt-get update
!sudo apt-get install texlive-xetex texlive-fonts-recommended␣
 ↪texlive-plain-generic

# installing pypandoc
!pip install pypandoc
```

```python
# connecting your google drive
from google.colab import drive
drive.mount('/content/drive')

# copying your file over. Change "Class6-Completed.ipynb" to whatever your file
 ↪is called (see top of notebook)
!cp "drive/My Drive/Colab Notebooks/410_Final.ipynb" ./

# Again, replace "Class6-Completed.ipynb" to whatever your file is called (see
 ↪top of notebook)
!jupyter nbconvert --to PDF "410_Final.ipynb"
```

Reading package lists… Done
Building dependency tree… Done
Reading state information… Done
pandoc is already the newest version (2.9.2.1-3ubuntu2).
pandoc set to manually installed.
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 libsynctex2 libteckit0 libtexlua53
libtexluajit2 libwoff1
  libzzip-0-13 lmodern poppler-data preview-latex-style rake ruby ruby-net-
telnet ruby-rubygems
  ruby-webrick ruby-xmlrpc ruby3.0 rubygems-integration t1utils teckit tex-
common tex-gyre
  texlive-base texlive-binaries texlive-fonts-recommended texlive-latex-base
  texlive-latex-recommended texlive-pictures texlive-plain-generic 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 libsynctex2 libteckit0 libtexlua53
libtexluajit2 libwoff1
  libzzip-0-13 lmodern poppler-data preview-latex-style rake ruby ruby-net-
telnet ruby-rubygems
  ruby-webrick ruby-xmlrpc ruby3.0 rubygems-integration t1utils teckit tex-
common tex-gyre texlive
  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, 55 newly installed, 0 to remove and 15 not upgraded.
Need to get 182 MB of archives.
After this operation, 572 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 libsynctex2
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 libzzip-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 texlive-latex-
recommended all 2021.20220204-1 [14.4 MB]
Get:48 http://archive.ubuntu.com/ubuntu jammy/universe amd64 texlive all
2021.20220204-1 [14.3 kB]
Get:49 http://archive.ubuntu.com/ubuntu jammy/universe amd64 libfontbox-java all
1:1.8.16-2 [207 kB]
Get:50 http://archive.ubuntu.com/ubuntu jammy/universe amd64 libpdfbox-java all
1:1.8.16-2 [5,199 kB]
Get:51 http://archive.ubuntu.com/ubuntu jammy/universe amd64 texlive-pictures
all 2021.20220204-1 [8,720 kB]
Get:52 http://archive.ubuntu.com/ubuntu jammy/universe amd64 texlive-latex-extra
all 2021.20220204-1 [13.9 MB]
Get:53 http://archive.ubuntu.com/ubuntu jammy/universe amd64 texlive-plain-
generic all 2021.20220204-1 [27.5 MB]
Get:54 http://archive.ubuntu.com/ubuntu jammy/universe amd64 tipa all 2:1.3-21
[2,967 kB]
Get:55 http://archive.ubuntu.com/ubuntu jammy/universe amd64 texlive-xetex all
2021.20220204-1 [12.4 MB]
Fetched 182 MB in 10s (17.8 MB/s)
Extracting templates from packages: 100%
Preconfiguring packages …
Selecting previously unselected package fonts-droid-fallback.
(Reading database … 120882 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 libsynctex2:amd64.
Preparing to unpack …/30-libsynctex2_2021.20210626.59705-1ubuntu0.1_amd64.deb
…
Unpacking libsynctex2: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 libzzip-0-13:amd64.
Preparing to unpack …/34-libzzip-0-13_0.13.72+dfsg.1-1.1_amd64.deb …
Unpacking libzzip-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 texlive-latex-recommended.
Preparing to unpack …/46-texlive-latex-recommended_2021.20220204-1_all.deb …
Unpacking texlive-latex-recommended (2021.20220204-1) …
Selecting previously unselected package texlive.
Preparing to unpack …/47-texlive_2021.20220204-1_all.deb …
```

```
Unpacking texlive (2021.20220204-1) …
Selecting previously unselected package libfontbox-java.
Preparing to unpack …/48-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 …/49-libpdfbox-java_1%3a1.8.16-2_all.deb …
Unpacking libpdfbox-java (1:1.8.16-2) …
Selecting previously unselected package texlive-pictures.
Preparing to unpack …/50-texlive-pictures_2021.20220204-1_all.deb …
Unpacking texlive-pictures (2021.20220204-1) …
Selecting previously unselected package texlive-latex-extra.
Preparing to unpack …/51-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 …/52-texlive-plain-generic_2021.20220204-1_all.deb …
Unpacking texlive-plain-generic (2021.20220204-1) …
Selecting previously unselected package tipa.
Preparing to unpack …/53-tipa_2%3a1.3-21_all.deb …
Unpacking tipa (2:1.3-21) …
Selecting previously unselected package texlive-xetex.
Preparing to unpack …/54-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 libzzip-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) …
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 t1utils (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-xmlrpc (0.3.2-1ubuntu0.1) …
Setting up libsynctex2: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
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 (2021.20220204-1) …
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/libtbbmalloc_proxy.so.2 is not a symbolic
link

/sbin/ldconfig.real: /usr/local/lib/libtbbbind_2_0.so.3 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/libtbbbind.so.3 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/libtbbmalloc.so.2 is not a symbolic link


Processing triggers for tex-common (6.17) …
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.
Get:1 http://security.ubuntu.com/ubuntu jammy-security InRelease [110 kB]
Get:2 https://cloud.r-project.org/bin/linux/ubuntu jammy-cran40/ InRelease
[3,626 B]
Get:3 https://developer.download.nvidia.com/compute/cuda/repos/ubuntu2204/x86_64
InRelease [1,581 B]
Hit:4 http://archive.ubuntu.com/ubuntu jammy InRelease
Get:5 https://cloud.r-project.org/bin/linux/ubuntu jammy-cran40/ Packages [47.6
kB]
Get:6 http://archive.ubuntu.com/ubuntu jammy-updates InRelease [119 kB]
Get:7 http://security.ubuntu.com/ubuntu jammy-security/restricted amd64 Packages
[1,494 kB]
Get:8 https://developer.download.nvidia.com/compute/cuda/repos/ubuntu2204/x86_64
Packages [632 kB]
Get:9 http://security.ubuntu.com/ubuntu jammy-security/main amd64 Packages
[1,265 kB]
Hit:10 https://ppa.launchpadcontent.net/c2d4u.team/c2d4u4.0+/ubuntu jammy
InRelease
Hit:11 https://ppa.launchpadcontent.net/deadsnakes/ppa/ubuntu jammy InRelease
Get:12 http://archive.ubuntu.com/ubuntu jammy-backports InRelease [109 kB]
Hit:13 https://ppa.launchpadcontent.net/graphics-drivers/ppa/ubuntu jammy
InRelease
Hit:14 https://ppa.launchpadcontent.net/ubuntugis/ppa/ubuntu jammy InRelease
Get:15 http://archive.ubuntu.com/ubuntu jammy-updates/main amd64 Packages [1,535
kB]
Get:16 http://archive.ubuntu.com/ubuntu jammy-updates/restricted amd64 Packages
[1,520 kB]
```

```
Get:17 http://archive.ubuntu.com/ubuntu jammy-updates/universe amd64 Packages
[1,292 kB]
Get:18 http://archive.ubuntu.com/ubuntu jammy-backports/universe amd64 Packages
[32.6 kB]
Fetched 8,159 kB in 1s (5,506 kB/s)
Reading package lists… Done
Reading package lists… Done
Building dependency tree… Done
Reading state information… Done
texlive-fonts-recommended is already the newest version (2021.20220204-1).
texlive-fonts-recommended set to manually installed.
texlive-plain-generic is already the newest version (2021.20220204-1).
texlive-plain-generic set to manually installed.
texlive-xetex is already the newest version (2021.20220204-1).
0 upgraded, 0 newly installed, 0 to remove and 17 not upgraded.
Collecting pypandoc
  Downloading pypandoc-1.12-py3-none-any.whl (20 kB)
```