Census_Income

November 19, 2022

1 Census Income Data

1.0.1 Data Set Information:

- Extraction was done by Barry Becker from the 1994 Census database. A set of reasonably clean records was extracted using the following conditions: ((AAGE>16) && (AGI>100) && (AFNLWGT>1)&& (HRSWK>0))
- Prediction task is to determine whether a person makes over 50K a year.

1.0.2 Machine Learning Models:

- Logistic Regression
- SVC Classifier
- Decision Tree Classifier
- Random Forest Classifier
- Bagging Classfier
- ExtraTreesClassifier
- Voting Classifier

1.0.3 Model Building Lifecycle

- Data Ingestion/Gathering
- EDA
- Preprocessing
- Feature Selection
- Model Building
- Evaluation of the model

[1]:
```
import pandas as pd
import numpy as np
import seaborn as sns
sns.despine(left=True, right=True, bottom=True, top=True)
sns.set_style('white')
import matplotlib.pyplot as plt
import warnings
warnings.filterwarnings('ignore')
warnings.warn('ALERT!')
warnings.warn('CAUTION!')
```

```
<Figure size 432x288 with 0 Axes>
```

[2]:
```
columns=['age','work_class','final_weigth','education','education_num','marital_status','occup
test_columns=['age','work_class','final_weigth','education','education_num','marital_status','
```

[3]:
```
df_train=pd.read_csv(r'C:
 ↪\Users\prasa\Desktop\Prasath\INeuron\FSDS_Bootcamp_2022\FSDS BootCamp\Nov␣
 ↪Machine Learning Algorithms\13th Nov Live Class Random Forest\13th Nov FSDS␣
 ↪bootcamp-20221114T151600Z-001\Assignment\dataset\adult.
 ↪data',sep=',',names=columns,header=None)
df_test=pd.read_csv(r'C:
 ↪\Users\prasa\Desktop\Prasath\INeuron\FSDS_Bootcamp_2022\FSDS BootCamp\Nov␣
 ↪Machine Learning Algorithms\13th Nov Live Class Random Forest\13th Nov FSDS␣
 ↪bootcamp-20221114T151600Z-001\Assignment\dataset\adult.
 ↪test',sep=',',names=test_columns,header=0,index_col=False)
```

[4]:
```
df_test.head()
```

[4]:

	age	work_class	final_weigth	education	education_num	marital_status	occupation	relationship	race	sex	capital_gain	capital_loss	hours_per_week	native_country
0	25	Private	226802	11th	7	Never-married	Machine-op-inspct	Own-child	Black	Male	0	0	40	United-States
1	38	Private	89814	HS-grad	9	Married-civ-spouse	Farming-fishing	Husband	White	Male	0	0	50	United-States
2	28	Local-gov	336951	Assoc-acdm	12	Married-civ-spouse	Protective-serv	Husband	White	Male	0	0	40	United-States
3	44	Private	160323	Some-college	10	Married-civ-spouse	Machine-op-inspct	Husband	Black	Male	7688	0	40	United-States
4	18	?	103497	Some-college	10	Never-married	?	Own-child	White	Female	0	0	30	United-States

[5]:
```
df_train.head()
```

[5]:

	age	work_class	final_weigth	education	education_num
0	39	State-gov	77516	Bachelors	13
1	50	Self-emp-not-inc	83311	Bachelors	13
2	38	Private	215646	HS-grad	9
3	53	Private	234721	11th	7
4	28	Private	338409	Bachelors	13

```
        marital_status          occupation    relationship    race      sex  \
0        Never-married        Adm-clerical   Not-in-family   White     Male
1   Married-civ-spouse     Exec-managerial         Husband   White     Male
2             Divorced   Handlers-cleaners   Not-in-family   White     Male
3   Married-civ-spouse   Handlers-cleaners         Husband   Black     Male
4   Married-civ-spouse      Prof-specialty            Wife   Black   Female

   capital_gain  capital_loss  hours_per_week  native_country  Income
0          2174             0              40   United-States   <=50K
1             0             0              13   United-States   <=50K
2             0             0              40   United-States   <=50K
3             0             0              40   United-States   <=50K
4             0             0              40            Cuba   <=50K
```

[6]:
```
df=pd.concat([df_train,df_test],axis=0,ignore_index=True)
```

[7]:
```
df.shape
```

[7]:
```
(48842, 15)
```

[8]:
```
df.memory_usage()
```

[8]:
```
Index               128
age              390736
work_class       390736
final_weigth     390736
education        390736
education_num    390736
marital_status   390736
occupation       390736
relationship     390736
race             390736
sex              390736
capital_gain     390736
capital_loss     390736
hours_per_week   390736
native_country   390736
Income           390736
dtype: int64
```

[9]:
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 48842 entries, 0 to 48841
Data columns (total 15 columns):
 #   Column          Non-Null Count  Dtype
---  ------          --------------  -----
 0   age             48842 non-null  int64
```

```
 1   work_class      48842 non-null  object
 2   final_weigth    48842 non-null  int64
 3   education       48842 non-null  object
 4   education_num   48842 non-null  int64
 5   marital_status  48842 non-null  object
 6   occupation      48842 non-null  object
 7   relationship    48842 non-null  object
 8   race            48842 non-null  object
 9   sex             48842 non-null  object
 10  capital_gain    48842 non-null  int64
 11  capital_loss    48842 non-null  int64
 12  hours_per_week  48842 non-null  int64
 13  native_country  48842 non-null  object
 14  Income          32561 non-null  object
dtypes: int64(6), object(9)
memory usage: 5.6+ MB
```

[10]:
```
df.describe()
```

[10]:

	age	final_weigth	education_num	capital_gain	capital_loss	hours_per_week
count	48842.000000	4.884200e+04	48842.000000	48842.000000	48842.000000	48842.000000
mean	38.643585	1.896641e+05	10.078089	1079.067626	87.502314	40.422382
std	13.710510	1.056040e+05	2.570973	7452.019058	403.004552	12.391444
min	17.000000	1.228500e+04	1.000000	0.000000	0.000000	1.000000
25%	28.000000	1.175505e+05	9.000000	0.000000	0.000000	40.000000
50%	37.000000	1.781445e+05	10.000000	0.000000	0.000000	40.000000
75%	48.000000	2.376420e+05	12.000000	0.000000	0.000000	45.000000
max	90.000000	1.490400e+06	16.000000	99999.000000	4356.000000	99.000000

[11]:
```
df.isna().sum()
```

[11]:
```
age                0
work_class         0
final_weigth       0
education          0
education_num      0
marital_status     0
occupation         0
```

```
relationship          0
race                  0
sex                   0
capital_gain          0
capital_loss          0
hours_per_week        0
native_country        0
Income            16281
dtype: int64
```

[12]:
```
df.tail()
```

[12]:
```
       age     work_class  final_weigth   education  education_num  \
48837   39        Private        215419   Bachelors             13
48838   64              ?        321403     HS-grad              9
48839   38        Private        374983   Bachelors             13
48840   44        Private         83891   Bachelors             13
48841   35   Self-emp-inc        182148   Bachelors             13

            marital_status        occupation     relationship  \
48837             Divorced    Prof-specialty    Not-in-family
48838              Widowed                 ?   Other-relative
48839   Married-civ-spouse    Prof-specialty          Husband
48840             Divorced      Adm-clerical        Own-child
48841   Married-civ-spouse   Exec-managerial          Husband

                     race      sex  capital_gain  capital_loss  \
48837               White   Female             0             0
48838               Black     Male             0             0
48839               White     Male             0             0
48840   Asian-Pac-Islander     Male          5455             0
48841               White     Male             0             0

       hours_per_week  native_country Income
48837              36   United-States    NaN
48838              40   United-States    NaN
48839              50   United-States    NaN
48840              40   United-States    NaN
48841              60   United-States    NaN
```

[13]:
```
df['work_class'].value_counts()
```

[13]:
```
 Private             33906
 Self-emp-not-inc     3862
 Local-gov            3136
 ?                    2799
 State-gov            1981
```

```
 Self-emp-inc          1695
 Federal-gov           1432
 Without-pay             21
 Never-worked            10
Name: work_class, dtype: int64
```

[14]:
```
df[df['work_class']==' ?']
```

[14]:

	age	work_class	final_weigth	education	education_num	\
27	54	?	180211	Some-college	10	
61	32	?	293936	7th-8th	4	
69	25	?	200681	Some-college	10	
77	67	?	212759	10th	6	
106	17	?	304873	10th	6	
…	…	…	…	…	…	
48682	61	?	265201	Some-college	10	
48769	21	?	212661	Some-college	10	
48800	73	?	144872	HS-grad	9	
48812	81	?	26711	Assoc-voc	11	
48838	64	?	321403	HS-grad	9	

	marital_status	occupation	relationship	\
27	Married-civ-spouse	?	Husband	
61	Married-spouse-absent	?	Not-in-family	
69	Never-married	?	Own-child	
77	Married-civ-spouse	?	Husband	
106	Never-married	?	Own-child	
…	…	…	…	
48682	Married-civ-spouse	?	Husband	
48769	Never-married	?	Own-child	
48800	Married-civ-spouse	?	Husband	
48812	Married-civ-spouse	?	Husband	
48838	Widowed	?	Other-relative	

	race	sex	capital_gain	capital_loss	\
27	Asian-Pac-Islander	Male	0	0	
61	White	Male	0	0	
69	White	Male	0	0	
77	White	Male	0	0	
106	White	Female	34095	0	
…	…	…	…	…	
48682	White	Male	0	0	
48769	White	Female	0	0	
48800	White	Male	0	0	
48812	White	Male	2936	0	
48838	Black	Male	0	0	

```
       hours_per_week  native_country  Income
27                 60           South    >50K
61                 40               ?   <=50K
69                 40   United-States   <=50K
77                  2   United-States   <=50K
106                32   United-States   <=50K
…                 …               …     …
48682              14   United-States     NaN
48769              30   United-States     NaN
48800              25          Canada     NaN
48812              20   United-States     NaN
48838              40   United-States     NaN

[2799 rows x 15 columns]
```

[15]:
```
df['work_class'].mode().values[0]
```

[15]:
```
' Private'
```

[16]:
```
df.tail()
```

[16]:
```
       age     work_class  final_weigth   education  education_num  \
48837   39        Private        215419   Bachelors             13
48838   64              ?        321403     HS-grad              9
48839   38        Private        374983   Bachelors             13
48840   44        Private         83891   Bachelors             13
48841   35   Self-emp-inc        182148   Bachelors             13

           marital_status        occupation     relationship  \
48837            Divorced    Prof-specialty    Not-in-family
48838             Widowed                 ?   Other-relative
48839  Married-civ-spouse    Prof-specialty          Husband
48840            Divorced      Adm-clerical        Own-child
48841  Married-civ-spouse   Exec-managerial          Husband

                    race      sex  capital_gain  capital_loss  \
48837              White   Female             0             0
48838              Black     Male             0             0
48839              White     Male             0             0
48840  Asian-Pac-Islander     Male          5455             0
48841              White     Male             0             0

       hours_per_week  native_country Income
48837              36   United-States    NaN
48838              40   United-States    NaN
48839              50   United-States    NaN
48840              40   United-States    NaN
```

```
48841              60   United-States    NaN
```

[17]:
```
df[df['work_class']==' ?']['work_class'].values
```

[17]:
```
array([' ?', ' ?', ' ?', …, ' ?', ' ?', ' ?'], dtype=object)
```

[18]:
```
df['work_class'].replace(to_replace=df[df['work_class']==' ?']['work_class'].
 ↪values,value=df['work_class'].mode().values[0],inplace=True)
```

[19]:
```
df[df['occupation']==' ?']['occupation'].values
```

[19]:
```
array([' ?', ' ?', ' ?', …, ' ?', ' ?', ' ?'], dtype=object)
```

[20]:
```
df['occupation'].mode()
```

[20]:
```
0     Prof-specialty
Name: occupation, dtype: object
```

[21]:
```
df['occupation'].replace(to_replace=df[df['occupation']==' ?']['occupation'].
 ↪values,value=df['occupation'].mode().values[0],inplace=True)
```

[22]:
```
df[df['education']==' Bachelors']
```

[22]:

	age	work_class	final_weigth	education	education_num	\
0	39	State-gov	77516	Bachelors	13	
1	50	Self-emp-not-inc	83311	Bachelors	13	
4	28	Private	338409	Bachelors	13	
9	42	Private	159449	Bachelors	13	
11	30	State-gov	141297	Bachelors	13	
…	…	…	…	…	…	
48836	33	Private	245211	Bachelors	13	
48837	39	Private	215419	Bachelors	13	
48839	38	Private	374983	Bachelors	13	
48840	44	Private	83891	Bachelors	13	
48841	35	Self-emp-inc	182148	Bachelors	13	

	marital_status	occupation	relationship	\
0	Never-married	Adm-clerical	Not-in-family	
1	Married-civ-spouse	Exec-managerial	Husband	
4	Married-civ-spouse	Prof-specialty	Wife	
9	Married-civ-spouse	Exec-managerial	Husband	
11	Married-civ-spouse	Prof-specialty	Husband	
…	…	…	…	
48836	Never-married	Prof-specialty	Own-child	
48837	Divorced	Prof-specialty	Not-in-family	
48839	Married-civ-spouse	Prof-specialty	Husband	
48840	Divorced	Adm-clerical	Own-child	
48841	Married-civ-spouse	Exec-managerial	Husband	

```
                      race      sex  capital_gain  capital_loss  \
0                    White     Male          2174             0
1                    White     Male             0             0
4                    Black   Female             0             0
9                    White     Male          5178             0
11      Asian-Pac-Islander     Male             0             0
…                      …      …             …             …
48836                White     Male             0             0
48837                White   Female             0             0
48839                White     Male             0             0
48840   Asian-Pac-Islander     Male          5455             0
48841                White     Male             0             0

       hours_per_week  native_country  Income
0                  40   United-States   <=50K
1                  13   United-States   <=50K
4                  40            Cuba   <=50K
9                  40   United-States    >50K
11                 40           India    >50K
…                   …               …      …
48836              40   United-States     NaN
48837              36   United-States     NaN
48839              50   United-States     NaN
48840              40   United-States     NaN
48841              60   United-States     NaN

[8025 rows x 15 columns]
```

[23]:
```
columns
```

[23]:
```
['age',
 'work_class',
 'final_weigth',
 'education',
 'education_num',
 'marital_status',
 'occupation',
 'relationship',
 'race',
 'sex',
 'capital_gain',
 'capital_loss',
 'hours_per_week',
 'native_country',
 'Income']
```

[24]:
```
cat_variable=df.dtypes[df.dtypes=='object'].index
```

[25]:
```
cat_variable
```

[25]:
```
Index(['work_class', 'education', 'marital_status', 'occupation',
       'relationship', 'race', 'sex', 'native_country', 'Income'],
      dtype='object')
```

[26]:
```
for i in cat_variable:
    df[i]=df[i].str.strip()
```

[27]:
```
print(f"The total number of Categorical Variables are {len(cat_variable)}\n")
for i in cat_variable:
    print(f"The {i} has the following category counts:\n")
    print(df[i].value_counts())
    print('\n')
```

```
The total number of Categorical Variables are 9

The work_class has the following category counts:

Private             36705
Self-emp-not-inc     3862
Local-gov            3136
State-gov            1981
Self-emp-inc         1695
Federal-gov          1432
Without-pay            21
Never-worked           10
Name: work_class, dtype: int64


The education has the following category counts:

HS-grad         15784
Some-college    10878
Bachelors        8025
Masters          2657
Assoc-voc        2061
11th             1812
Assoc-acdm       1601
10th             1389
7th-8th           955
Prof-school       834
9th               756
12th              657
Doctorate         594
5th-6th           509
```

```
1st-4th           247
Preschool          83
Name: education, dtype: int64


The marital_status has the following category counts:

Married-civ-spouse       22379
Never-married            16117
Divorced                  6633
Separated                 1530
Widowed                   1518
Married-spouse-absent      628
Married-AF-spouse           37
Name: marital_status, dtype: int64


The occupation has the following category counts:

Prof-specialty       8981
Craft-repair         6112
Exec-managerial      6086
Adm-clerical         5611
Sales                5504
Other-service        4923
Machine-op-inspct    3022
Transport-moving     2355
Handlers-cleaners    2072
Farming-fishing      1490
Tech-support         1446
Protective-serv       983
Priv-house-serv       242
Armed-Forces           15
Name: occupation, dtype: int64


The relationship has the following category counts:

Husband           19716
Not-in-family     12583
Own-child          7581
Unmarried          5125
Wife               2331
Other-relative     1506
Name: relationship, dtype: int64


The race has the following category counts:
```

```
White                 41762
Black                  4685
Asian-Pac-Islander     1519
Amer-Indian-Eskimo      470
Other                   406
Name: race, dtype: int64
```

```
The sex has the following category counts:

Male      32650
Female    16192
Name: sex, dtype: int64
```

```
The native_country has the following category counts:

United-States                 43832
Mexico                          951
?                               857
Philippines                     295
Germany                         206
Puerto-Rico                     184
Canada                          182
El-Salvador                     155
India                           151
Cuba                            138
England                         127
China                           122
South                           115
Jamaica                         106
Italy                           105
Dominican-Republic              103
Japan                            92
Guatemala                        88
Poland                           87
Vietnam                          86
Columbia                         85
Haiti                            75
Portugal                         67
Taiwan                           65
Iran                             59
Greece                           49
Nicaragua                        49
Peru                             46
Ecuador                          45
France                           38
```

```
Ireland                         37
Hong                            30
Thailand                        30
Cambodia                        28
Trinadad&Tobago                 27
Laos                            23
Yugoslavia                      23
Outlying-US(Guam-USVI-etc)      23
Scotland                        21
Honduras                        20
Hungary                         19
Holand-Netherlands               1
Name: native_country, dtype: int64


The Income has the following category counts:

<=50K    24720
>50K      7841
Name: Income, dtype: int64
```

[28]:
```
df[cat_variable].describe()
```

[28]:
```
       work_class education      marital_status      occupation relationship  \
count       48842     48842               48842           48842        48842
unique          8        16                   7              14            6
top       Private   HS-grad  Married-civ-spouse  Prof-specialty      Husband
freq        36705     15784               22379            8981        19716

         race    sex native_country Income
count   48842  48842          48842  32561
unique      5      2             42      2
top     White   Male  United-States  <=50K
freq    41762  32650          43832  24720
```

[29]:
```
df[df['Income']=='<=50K']['Income'].values
```

[29]:
```
array(['<=50K', '<=50K', '<=50K', …, '<=50K', '<=50K', '<=50K'],
      dtype=object)
```

[30]:
```
df['Income'].replace(to_replace=df[df['Income']=='<=50K']['Income'].
 ↪values,value=1,inplace=True)
df['Income'].replace(to_replace=df[df['Income']=='>50K']['Income'].
 ↪values,value=0,inplace=True)
```

[31]:
```
df.head()
```

[31]:
```
   age        work_class  final_weigth  education  education_num  \
0   39         State-gov         77516  Bachelors             13
1   50  Self-emp-not-inc         83311  Bachelors             13
2   38           Private        215646    HS-grad              9
3   53           Private        234721       11th              7
4   28           Private        338409  Bachelors             13

       marital_status         occupation   relationship   race     sex  \
0       Never-married       Adm-clerical  Not-in-family  White    Male
1  Married-civ-spouse    Exec-managerial        Husband  White    Male
2            Divorced  Handlers-cleaners  Not-in-family  White    Male
3  Married-civ-spouse  Handlers-cleaners        Husband  Black    Male
4  Married-civ-spouse     Prof-specialty           Wife  Black  Female

   capital_gain  capital_loss  hours_per_week native_country  Income
0          2174             0              40  United-States     1.0
1             0             0              13  United-States     1.0
2             0             0              40  United-States     1.0
3             0             0              40  United-States     1.0
4             0             0              40           Cuba     1.0
```

[32]:
```
df.dtypes
```

[32]:
```
age                int64
work_class        object
final_weigth       int64
education         object
education_num      int64
marital_status    object
occupation        object
relationship      object
race              object
sex               object
capital_gain       int64
capital_loss       int64
hours_per_week     int64
native_country    object
Income           float64
dtype: object
```

[33]:
```
numerical_var=df.dtypes[df.dtypes!='object'].index
```

[34]:
```
print("The Statiscal Anaylsis for each variable of Numerical Columns are:\n")
for i in numerical_var:
    print(f"The {i} Feature Descriptive Statistics:\n")
    print(f"The Maximum value of {i} in {df[df['age']==df['age'].max()].age.
 ↪values[0]} feature")
```

```
    print(f"The Average value of {i} in {df[i].mean()} feature")
    print(f"The Median value of {i} in {df[i].median()} feature")
    print(f"The Minimum value of {i} in {df[df['age']==df['age'].min()].age.
↪values[0]} feature")
    print(f"The Total Number of Duplicates from {i} feature is {df[i].
↪duplicated().sum()}\n\n")
```

```
The Statiscal Anaylsis for each variable of Numerical Columns are:

The age Feature Descriptive Statistics:

The Maximum value of age in 90 feature
The Average value of age in 38.64358543876172 feature
The Median value of age in 37.0 feature
The Minimum value of age in 17 feature
The Total Number of Duplicates from age feature is 48768


The final_weigth Feature Descriptive Statistics:

The Maximum value of final_weigth in 90 feature
The Average value of final_weigth in 189664.13459727284 feature
The Median value of final_weigth in 178144.5 feature
The Minimum value of final_weigth in 17 feature
The Total Number of Duplicates from final_weigth feature is 20319


The education_num Feature Descriptive Statistics:

The Maximum value of education_num in 90 feature
The Average value of education_num in 10.078088530363212 feature
The Median value of education_num in 10.0 feature
The Minimum value of education_num in 17 feature
The Total Number of Duplicates from education_num feature is 48826


The capital_gain Feature Descriptive Statistics:

The Maximum value of capital_gain in 90 feature
The Average value of capital_gain in 1079.0676262233324 feature
The Median value of capital_gain in 0.0 feature
The Minimum value of capital_gain in 17 feature
The Total Number of Duplicates from capital_gain feature is 48719


The capital_loss Feature Descriptive Statistics:

The Maximum value of capital_loss in 90 feature
```

```
The Average value of capital_loss in 87.50231358257237 feature
The Median value of capital_loss in 0.0 feature
The Minimum value of capital_loss in 17 feature
The Total Number of Duplicates from capital_loss feature is 48743


The hours_per_week Feature Descriptive Statistics:

The Maximum value of hours_per_week in 90 feature
The Average value of hours_per_week in 40.422382375824085 feature
The Median value of hours_per_week in 40.0 feature
The Minimum value of hours_per_week in 17 feature
The Total Number of Duplicates from hours_per_week feature is 48746


The Income Feature Descriptive Statistics:

The Maximum value of Income in 90 feature
The Average value of Income in 0.7591904425539756 feature
The Median value of Income in 1.0 feature
The Minimum value of Income in 17 feature
The Total Number of Duplicates from Income feature is 48839
```

[35]:
```
df.corr()
```

[35]:
```
                     age  final_weigth  education_num  capital_gain  \
age             1.000000     -0.076628       0.030940      0.077229
final_weigth   -0.076628      1.000000      -0.038761     -0.003706
education_num   0.030940     -0.038761       1.000000      0.125146
capital_gain    0.077229     -0.003706       0.125146      1.000000
capital_loss    0.056944     -0.004366       0.080972     -0.031441
hours_per_week  0.071558     -0.013519       0.143689      0.082157
Income         -0.234037      0.009463      -0.335154     -0.223329

                capital_loss  hours_per_week    Income
age                 0.056944        0.071558 -0.234037
final_weigth       -0.004366       -0.013519  0.009463
education_num       0.080972        0.143689 -0.335154
capital_gain       -0.031441        0.082157 -0.223329
capital_loss        1.000000        0.054467 -0.150526
hours_per_week      0.054467        1.000000 -0.229689
Income             -0.150526       -0.229689  1.000000
```

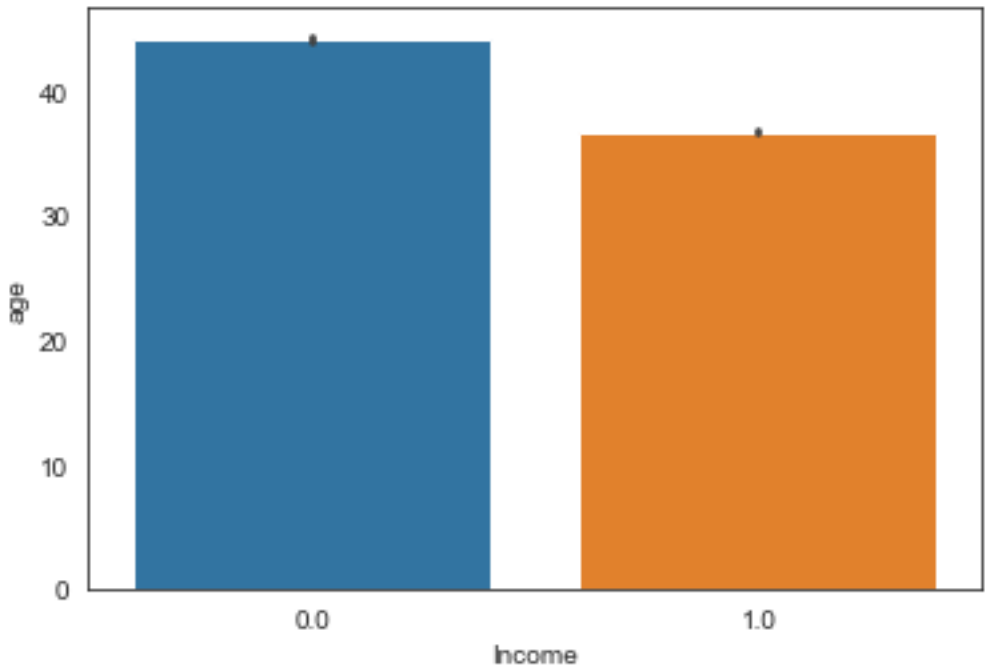
[36]:
```
num_var=numerical_var[:-1].values
```

[37]:
```
num_var
```

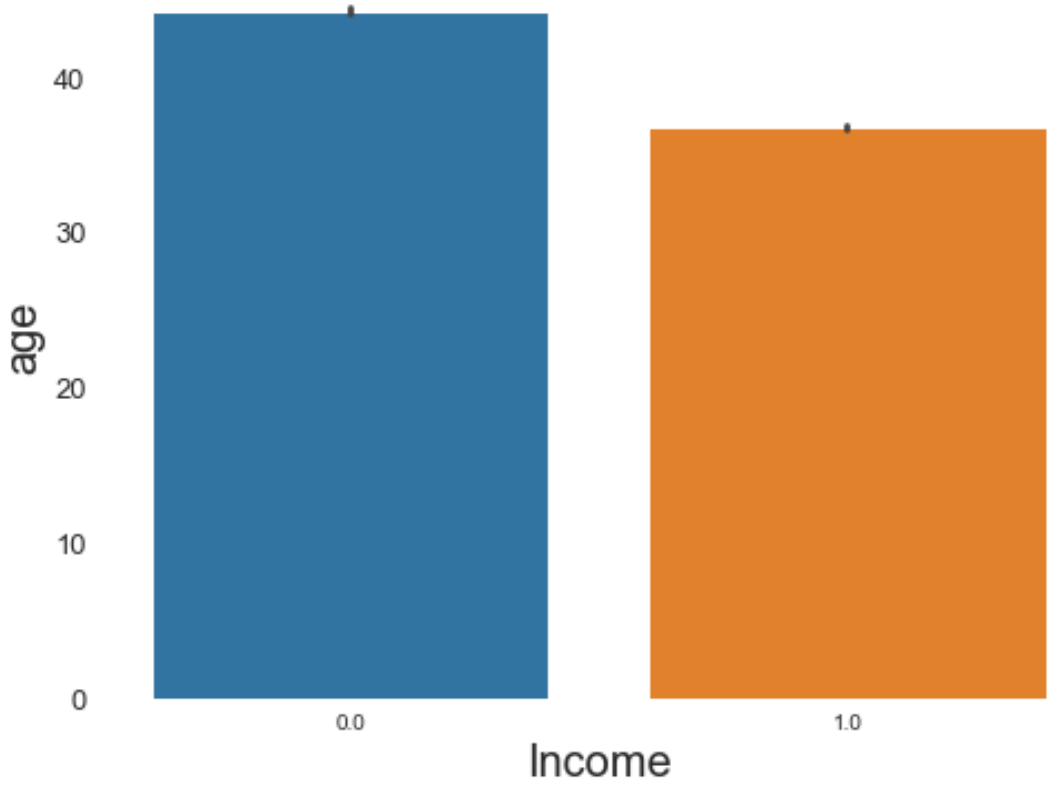
[37]:
```
array(['age', 'final_weigth', 'education_num', 'capital_gain',
       'capital_loss', 'hours_per_week'], dtype=object)
```

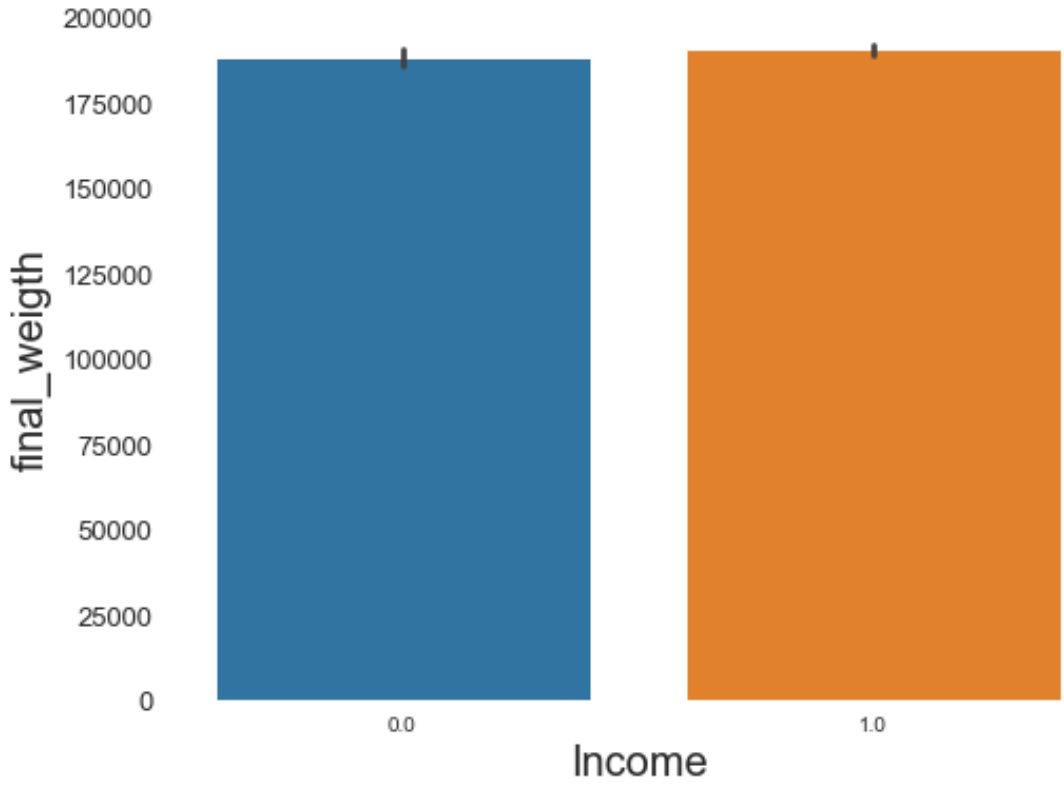
[38]:
```
sns.barplot(x=df['Income'],y=df['age'],data=df)
```

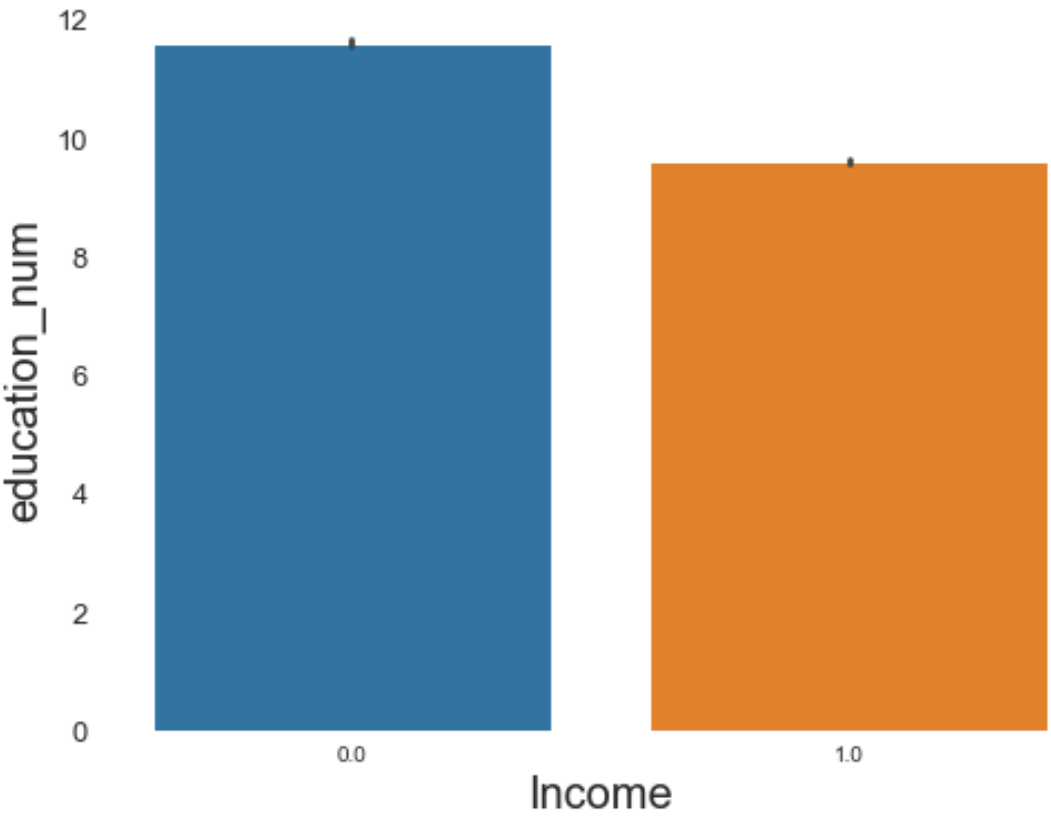
[38]:
```
<AxesSubplot: xlabel='Income', ylabel='age'>
```

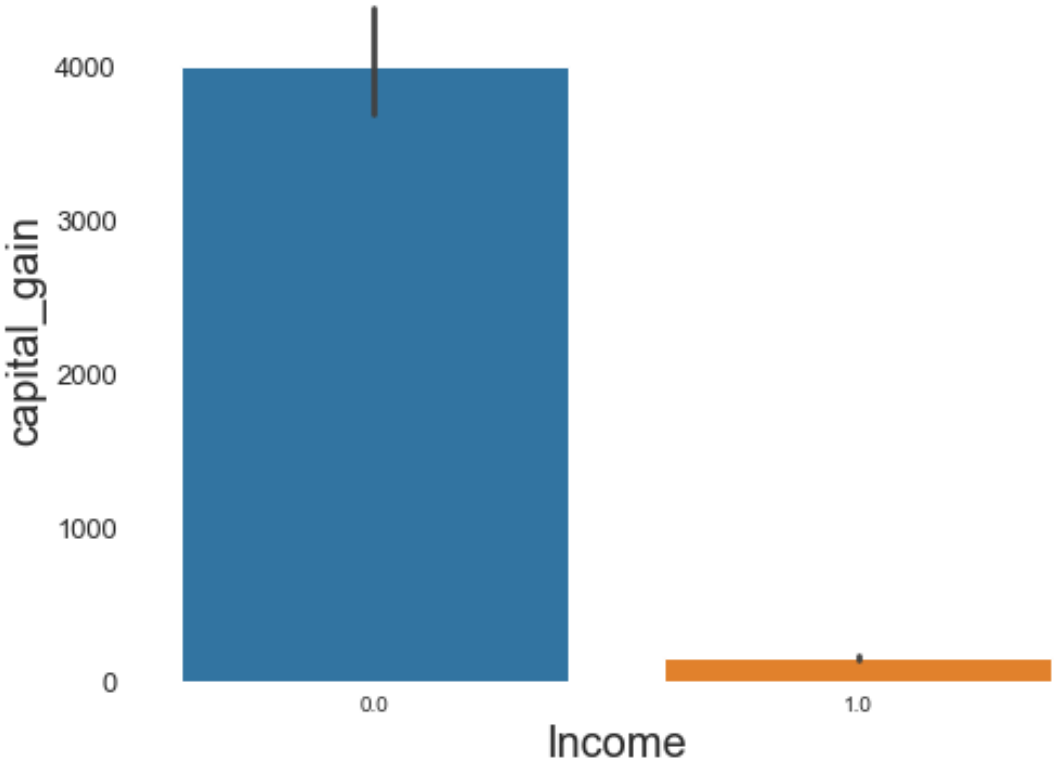


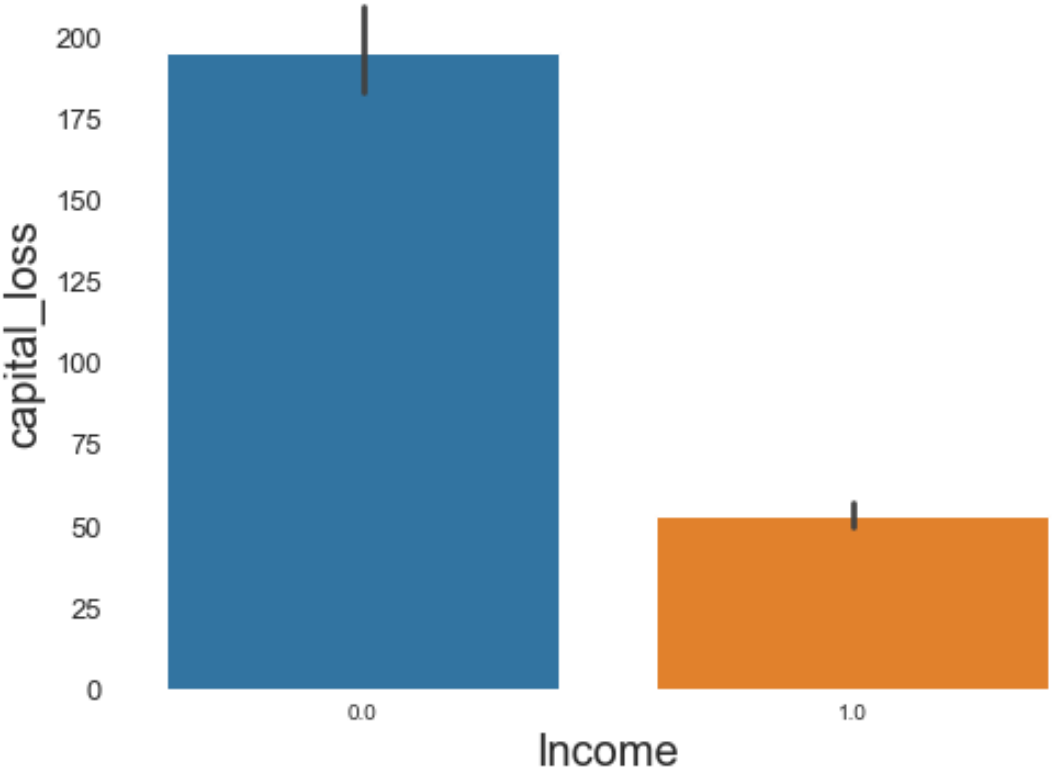
[39]:
```
for i in num_var:
    fig, ax = plt.subplots(1,1, figsize=(8, 6))
    sns.barplot(x=df['Income'],y=df[i],data=df)
    plt.xlabel(f'Income', fontsize=20)
    plt.ylabel(f'{i}',fontsize=20)
    plt.yticks(fontsize=13)
    plt.box(False)
```

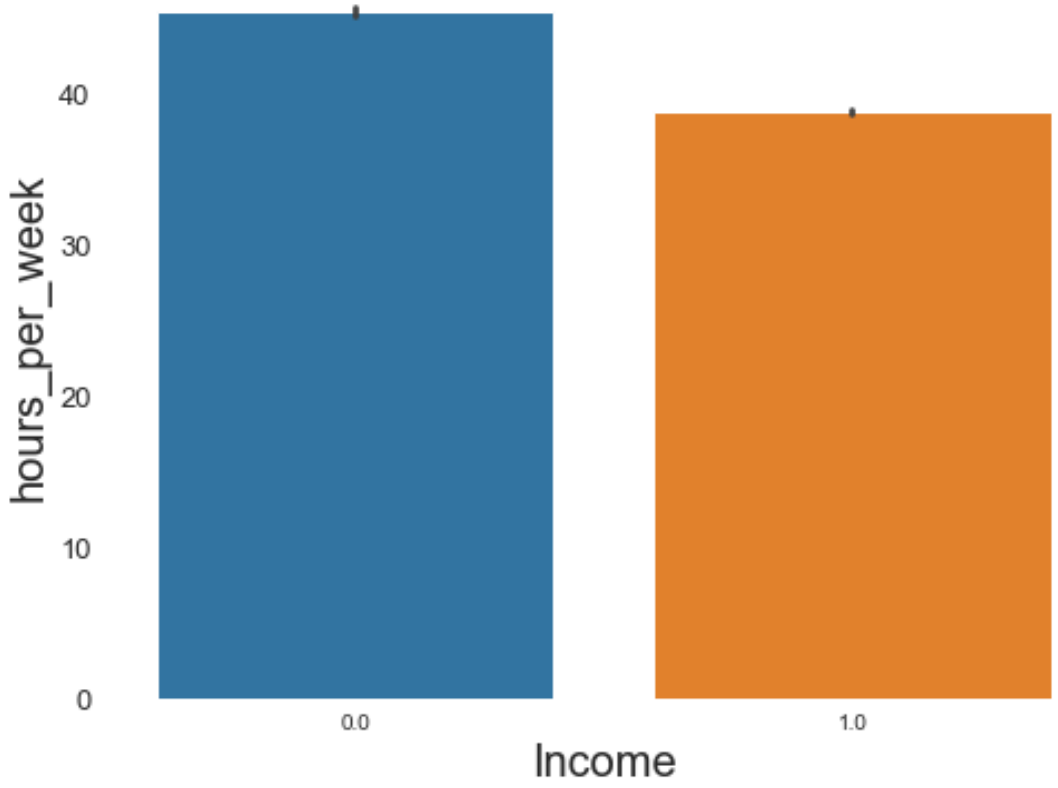

40
30
20
10
0
age
0.0
1.0
Income


200000
175000
150000
125000
100000
75000
50000
25000
0
final_weigth
0.0
1.0
Income

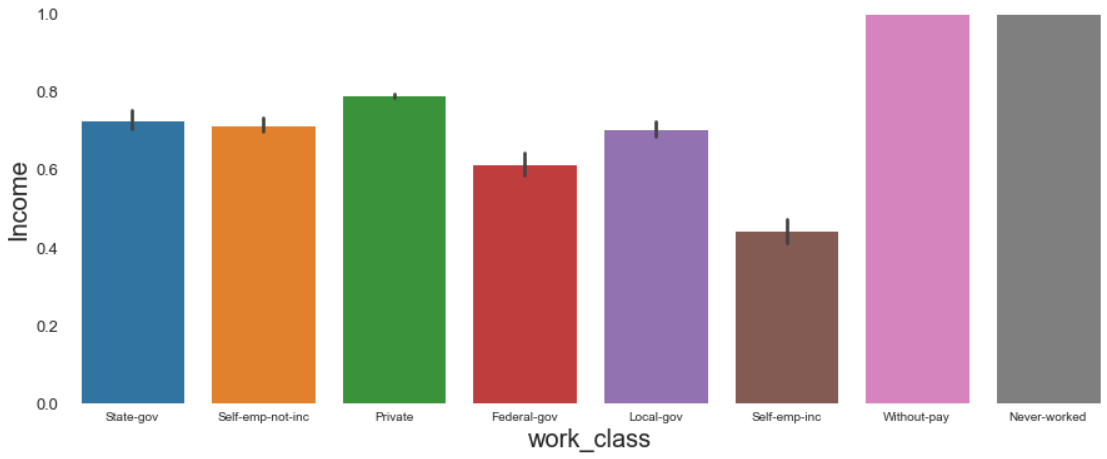

12
10
8
6
4
2
0
education_num
0.0
1.0
Income

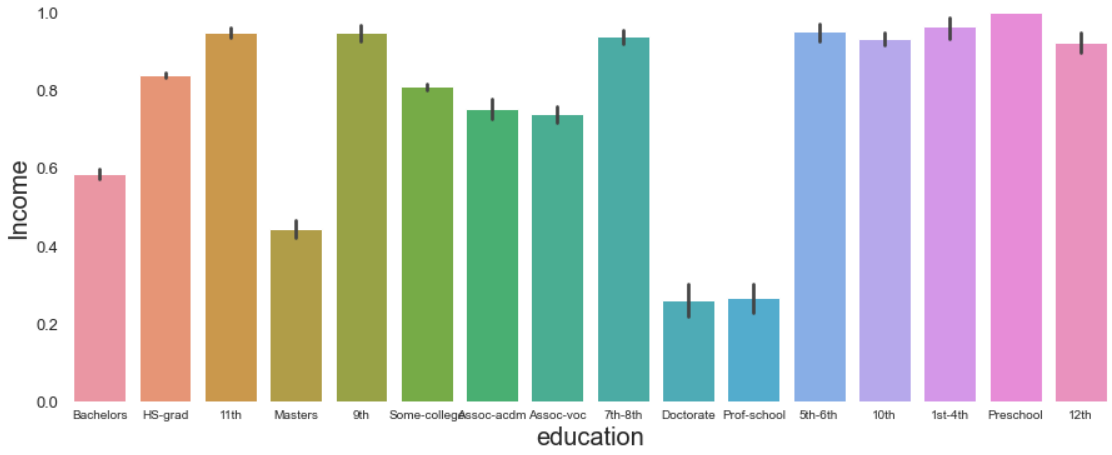

4000
3000
2000
1000
0
capital_gain
0.0
1.0
Income

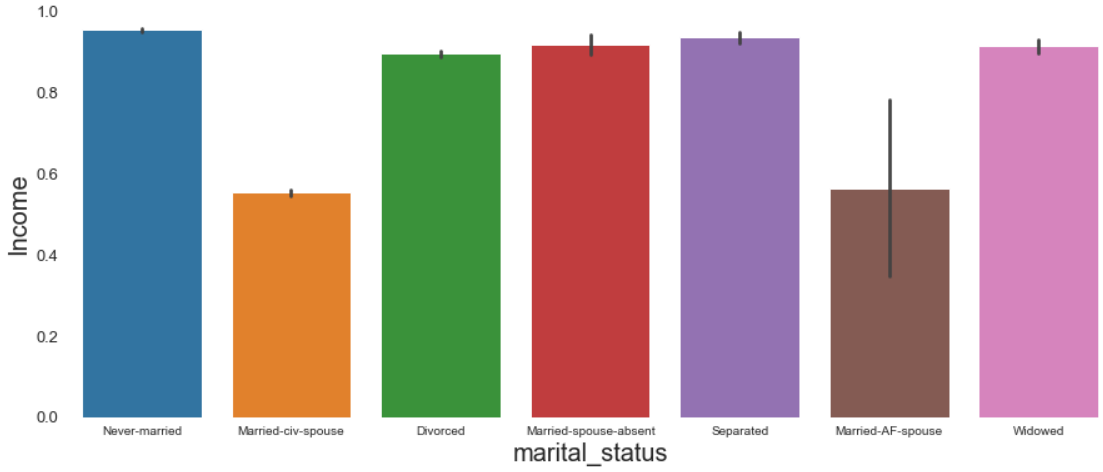

200
175
150
125
100
75
50
25
0
capital_loss
0.0
1.0
Income

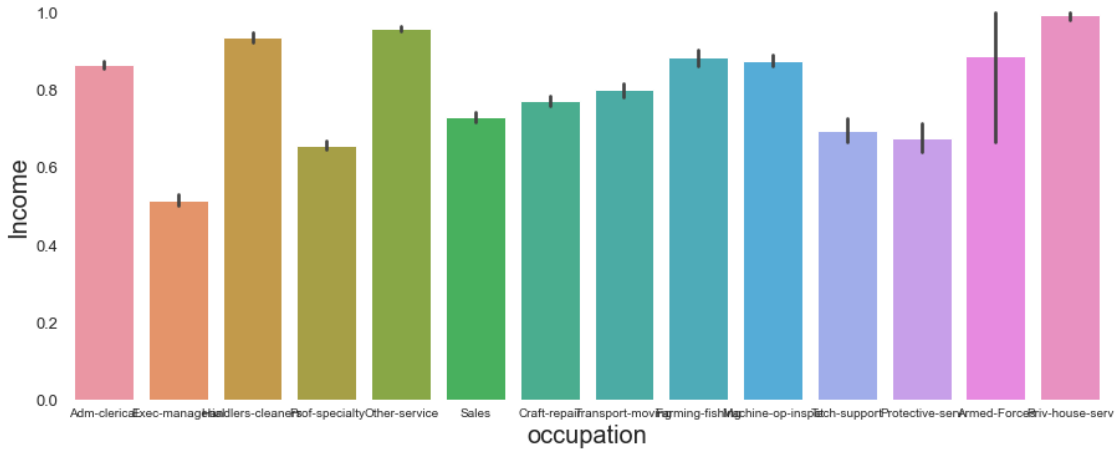


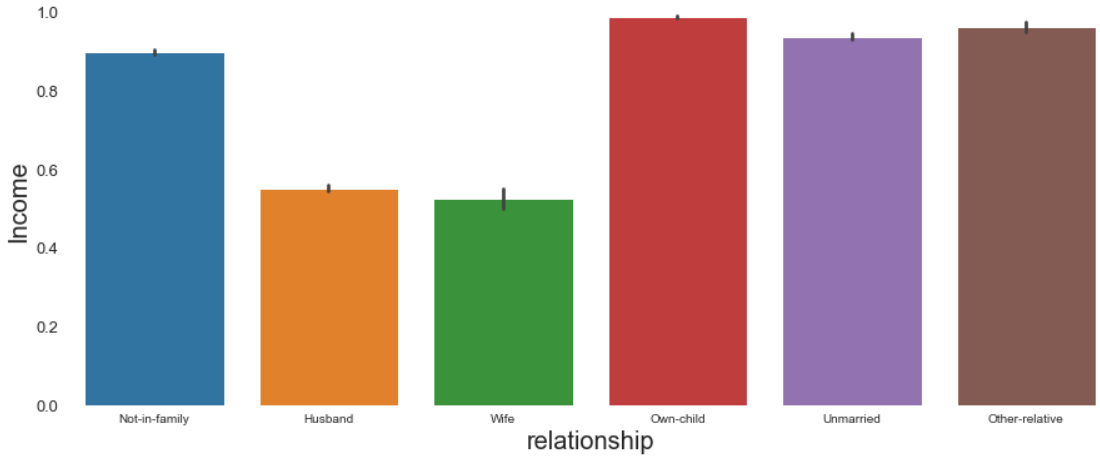
[40]:
```
for i in cat_variable:
    fig, ax = plt.subplots(1,1, figsize=(15, 6))
    sns.barplot(x=df[i],y=df['Income'],data=df)
    plt.xlabel(f'{i}', fontsize=20)
    plt.ylabel(f'Income',fontsize=20)
    plt.yticks(fontsize=13)
    plt.box(False)
```

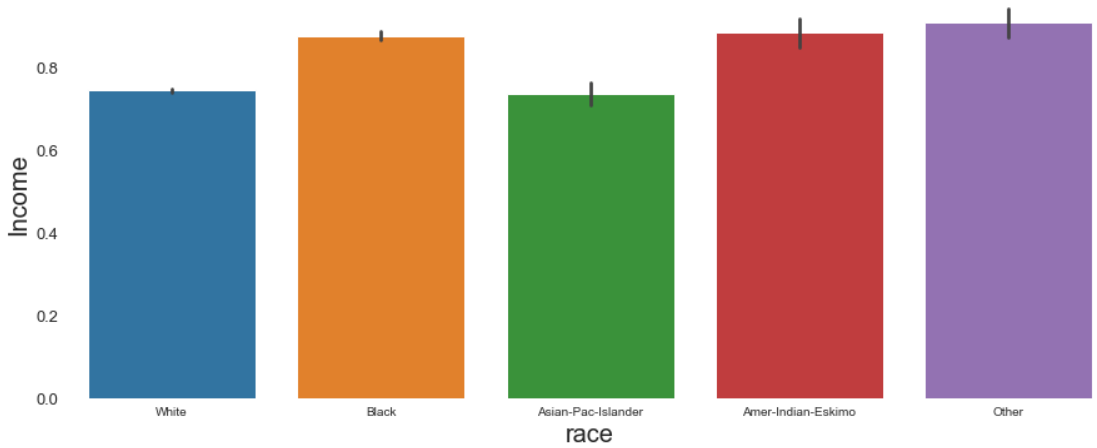


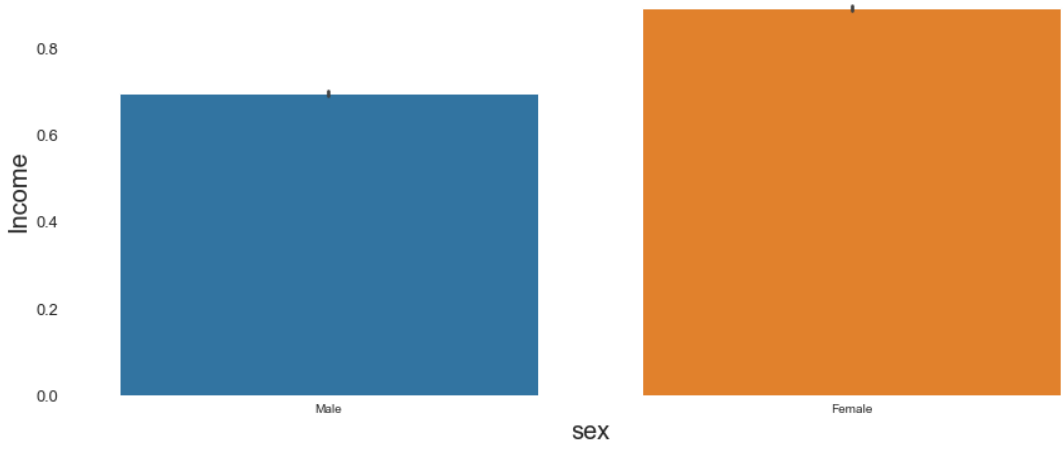

Income
1.0
0.8
0.6
0.4
0.2
0.0
Bachelors
HS-grad
11th
Masters
9th
Some-college
Assoc-acdm
Assoc-voc
7th-8th
Doctorate
Prof-school
5th-6th
10th
1st-4th
Preschool
12th
education

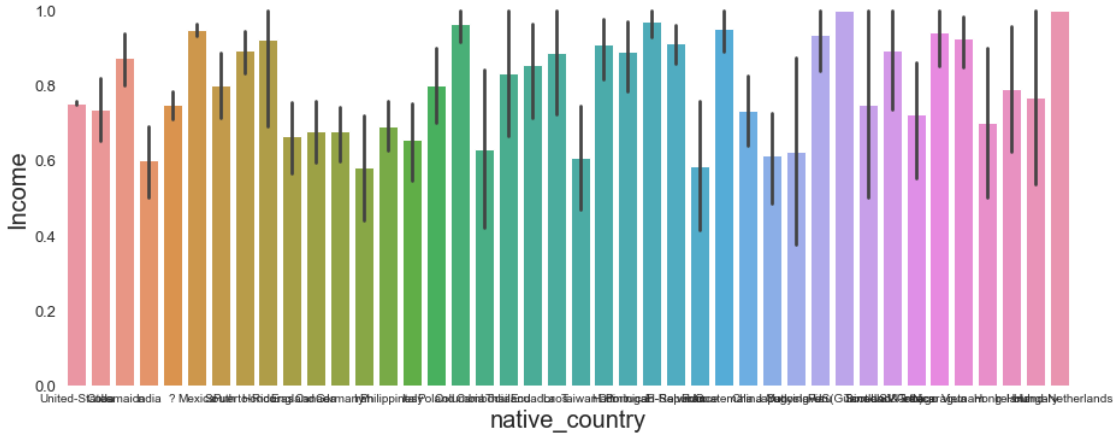

Income
1.0
0.8
0.6
0.4
0.2
0.0
Never-married
Married-civ-spouse
Divorced
Married-spouse-absent
Separated
Married-AF-spouse
Widowed
marital_status

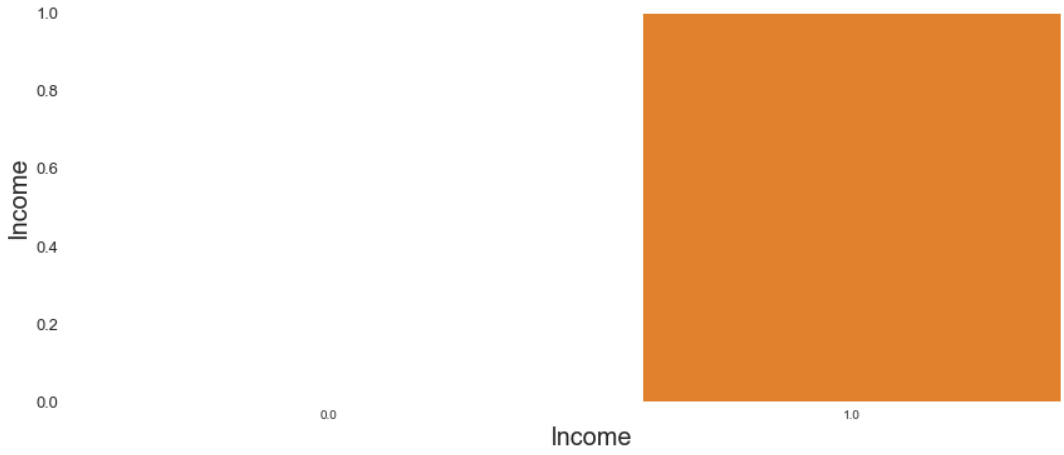

Income
1.0
0.8
0.6
0.4
0.2
0.0
Adm-clerical
Exec-managerial
Handlers-cleaners
Prof-specialty
Other-service
Sales
Craft-repair
Transport-moving
Farming-fishing
Machine-op-inspct
Tech-support
Protective-serv
Armed-Forces
Priv-house-serv
occupation


Income
1.0
0.8
0.6
0.4
0.2
0.0
Not-in-family
Husband
Wife
Own-child
Unmarried
Other-relative
relationship

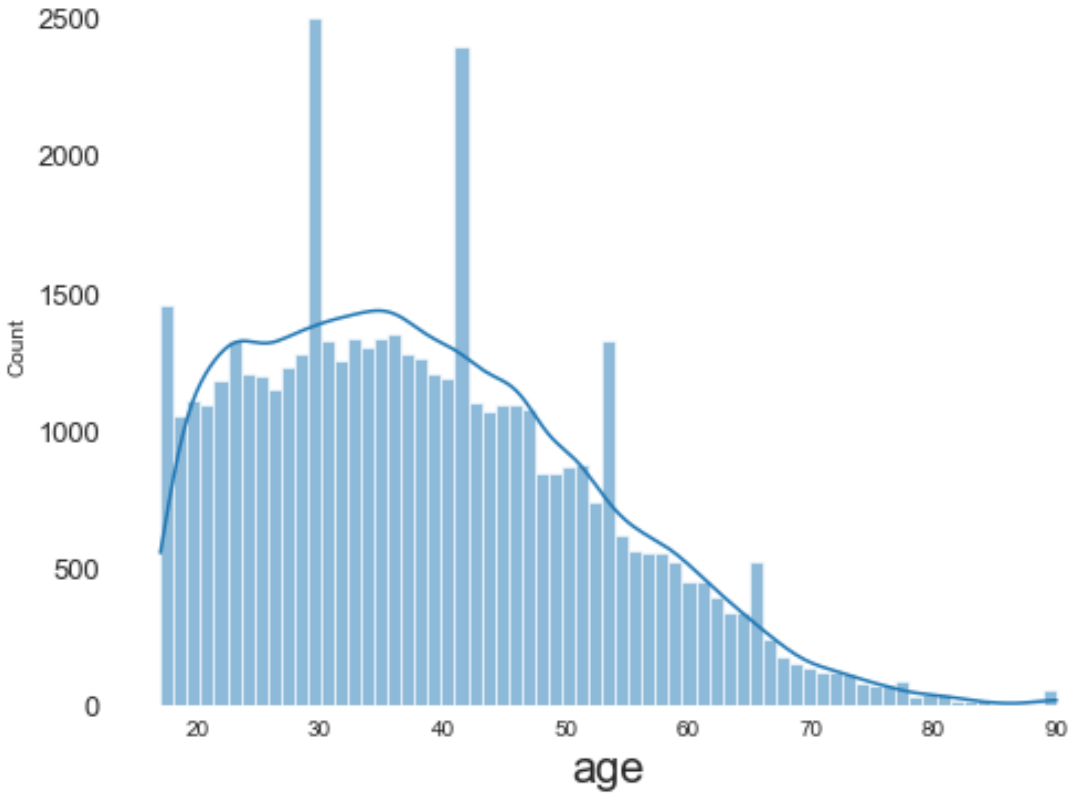

Income
0.8
0.6
0.4
0.2
0.0
White
Black
Asian-Pac-Islander
Amer-Indian-Eskimo
Other
race

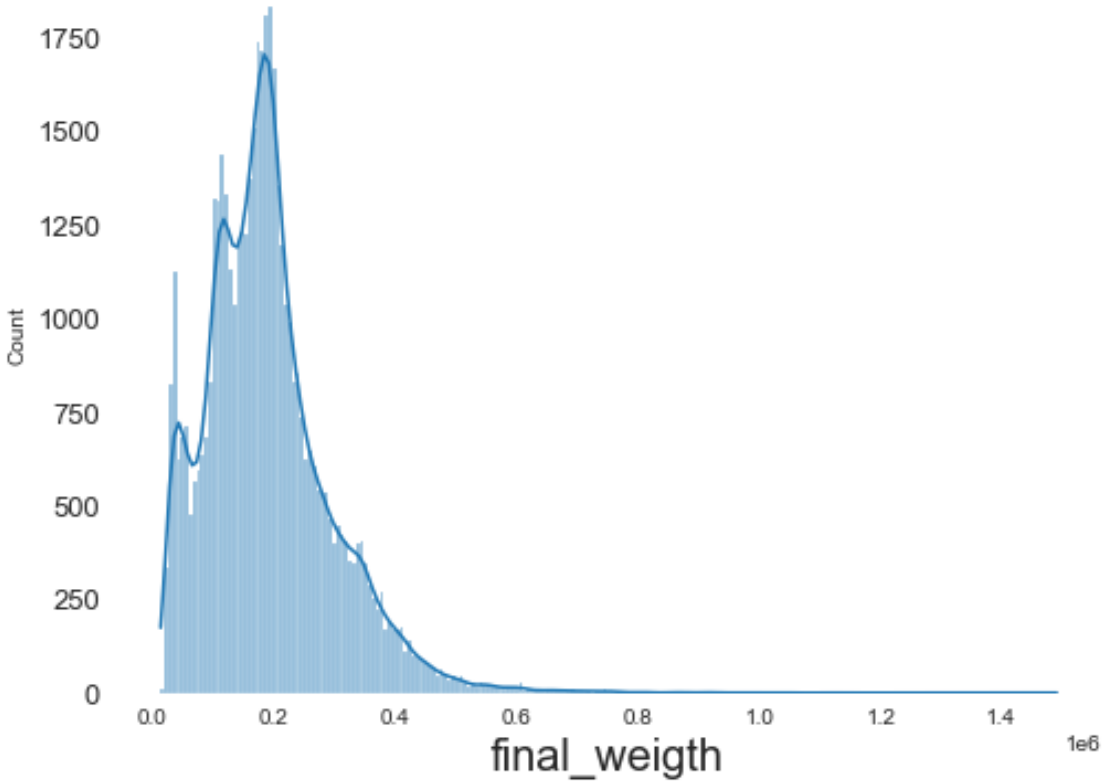

0.8
0.6
0.4
0.2
0.0
Income
Male
Female
sex

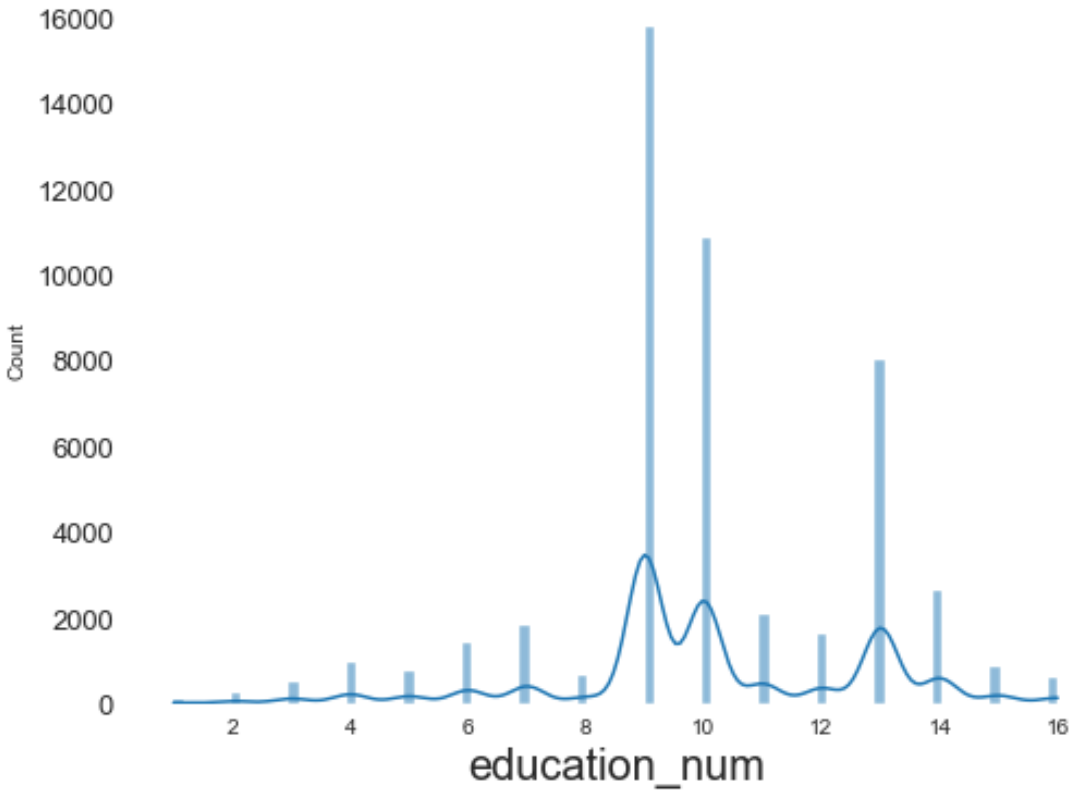

1.0
0.8
0.6
0.4
0.2
0.0
Income
native_country

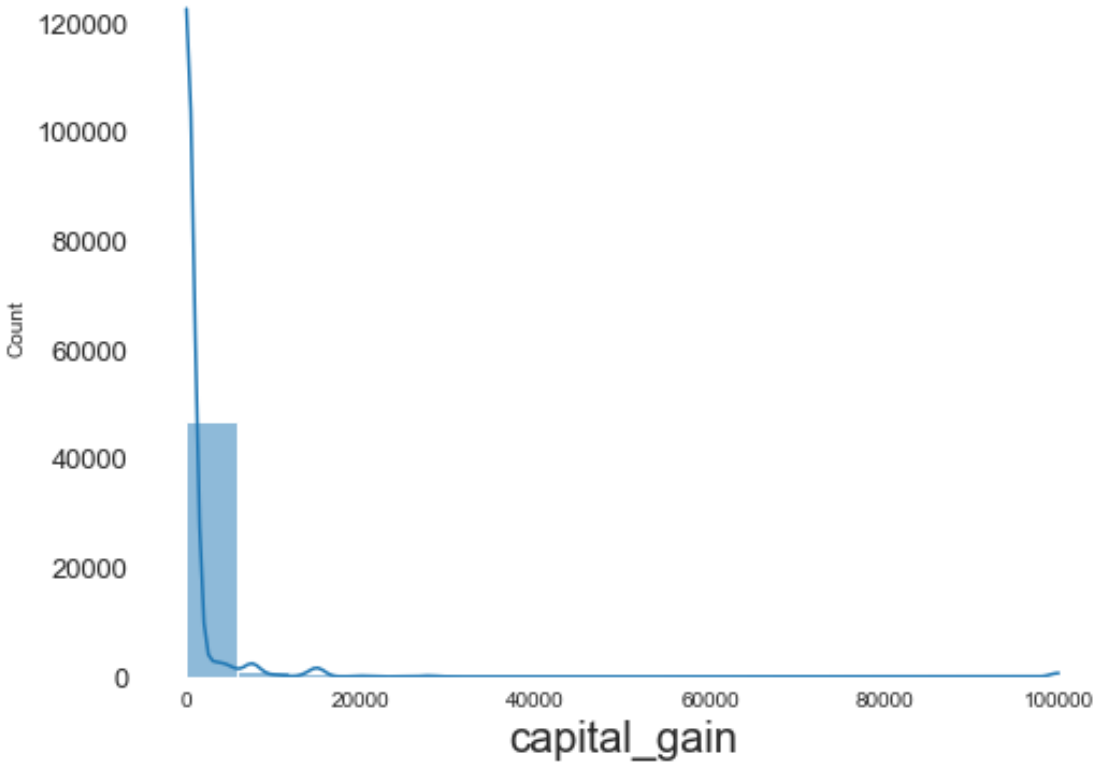

1.0
0.8
0.6
0.4
0.2
0.0
Income
0.0
1.0
Income

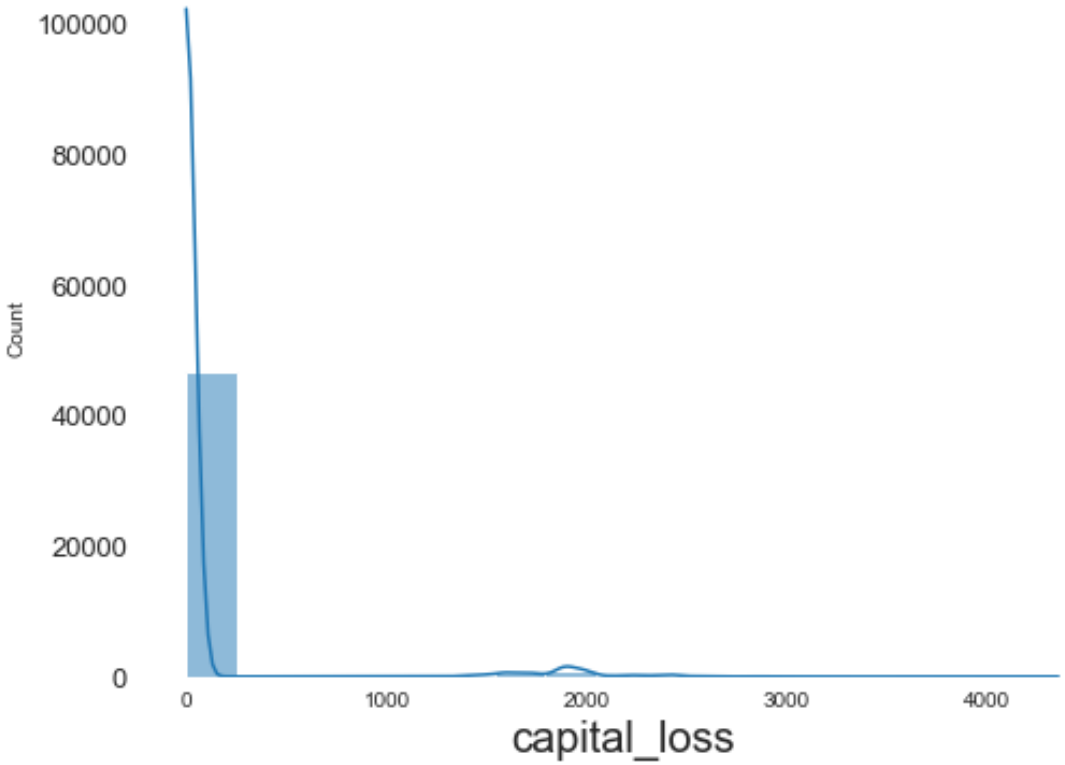
[41]:
```
for i in num_var:
    fig, ax = plt.subplots(1,1, figsize=(8, 6))
    sns.histplot(df[i],kde=True)
    plt.xlabel(f'{i}', fontsize=20)
    plt.yticks(fontsize=13)
    plt.box(False)
```

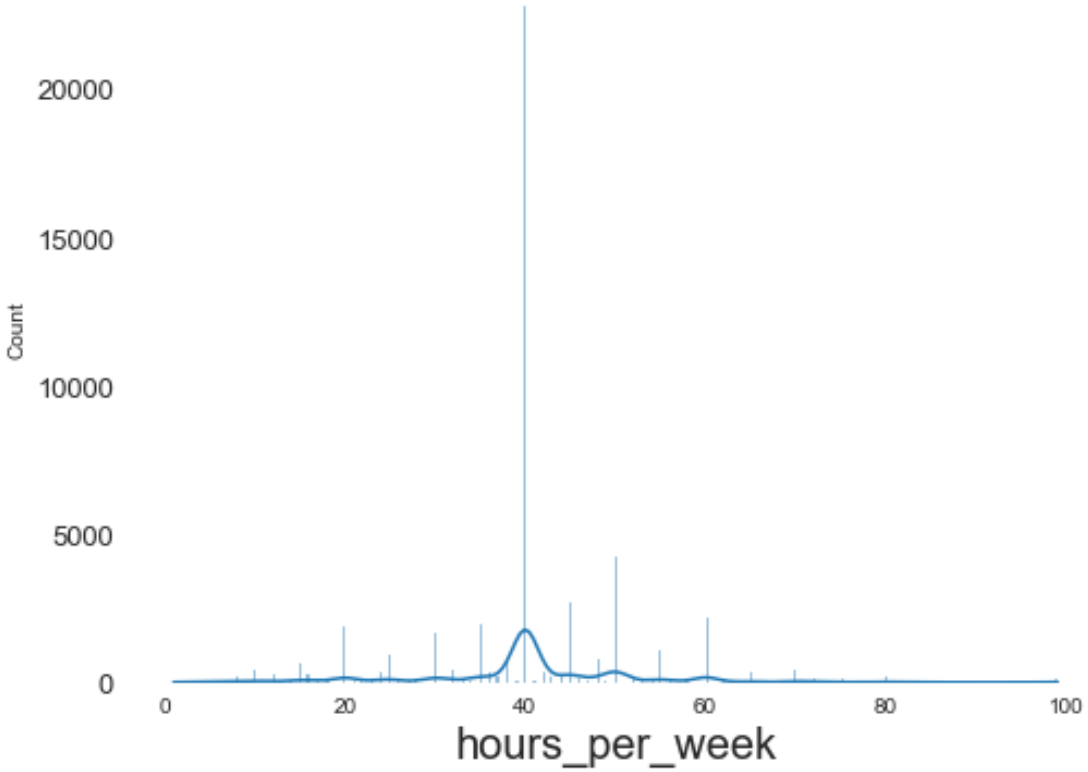



1750
1500
1250
1000
750
500
250
0
Count
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
1e6
final_weigth


16000
14000
12000
10000
8000
6000
4000
2000
0
Count
2
4
6
8
10
12
14
16
education_num


120000
100000
80000
60000
40000
20000
0
Count
0
20000
40000
60000
80000
100000
capital_gain


100000
80000
60000
40000
20000
0
Count
0
1000
2000
3000
4000
capital_loss



[42]:
```
def find_outliers(feature):
    q1,q3=df[feature].quantile([0.25,0.75])
    IQR=q3-q1
    lower_bound=q1-1.5*(IQR)
    upper_bound=q3+1.5*(IQR)
    return (lower_bound,upper_bound)
```

[43]:
```
q1,q3=find_outliers('age')
q1,q3
len(df[(df['age']<q1)|( df['age']>q3)])
```

[43]: 216

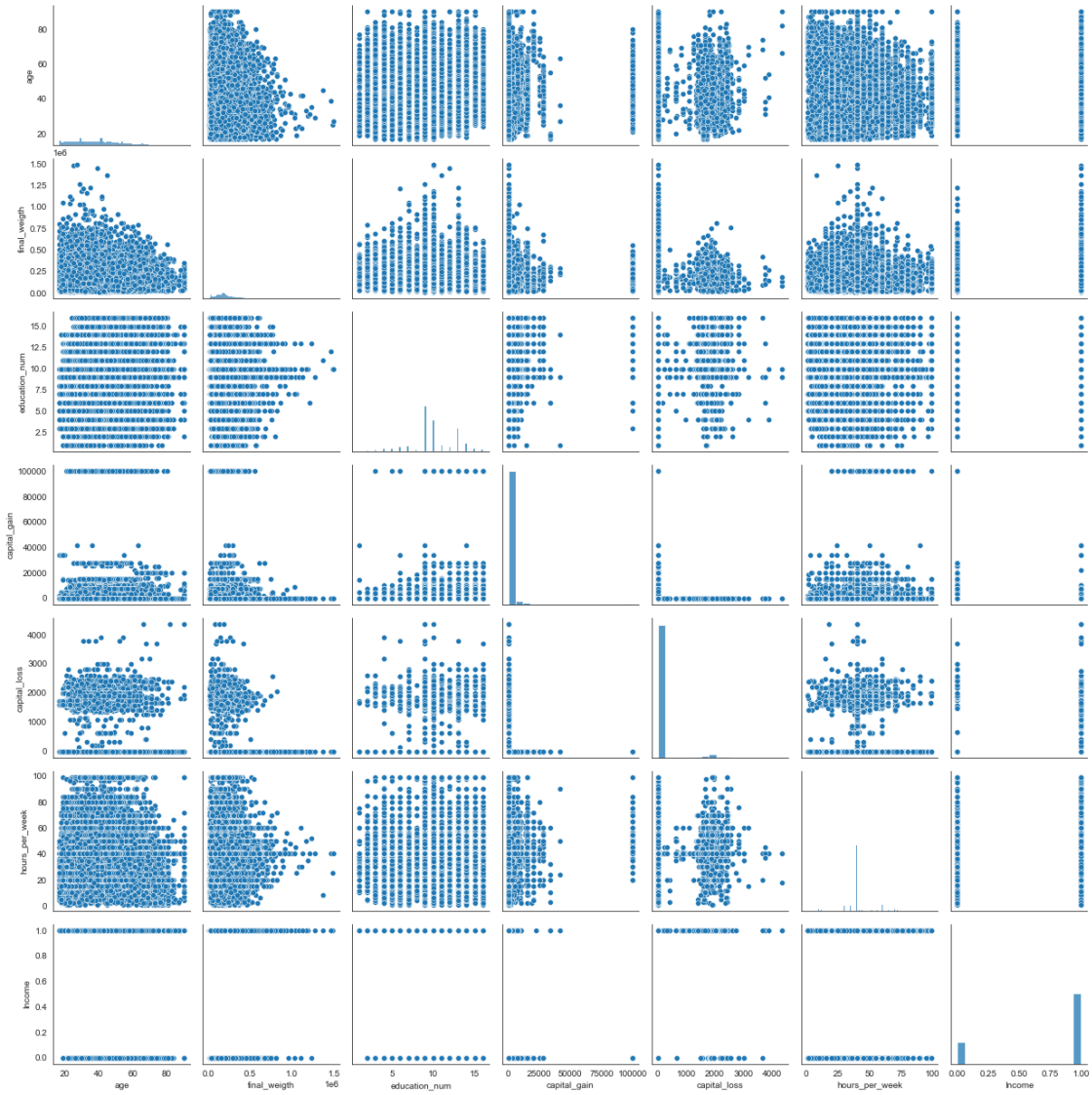
[44]:
```
for i in num_var:
    lower_bound,upper_bound=find_outliers(i)
    print(f"The Lower Bound is {lower_bound} and the Upper Bound is␣
 ↪{upper_bound}")
    length=len(df[(df[i]<q1)|( df[i]>q3)])
    print(f"The Total Outliers Lying in {i} Feature are {length}\n\n")
```

```
The Lower Bound is -2.0 and the Upper Bound is 78.0
```

```
The Total Outliers Lying in age Feature are 216


The Lower Bound is -62586.75 and the Upper Bound is 417779.25
The Total Outliers Lying in final_weigth Feature are 48842


The Lower Bound is 4.5 and the Upper Bound is 16.5
The Total Outliers Lying in education_num Feature are 0


The Lower Bound is 0.0 and the Upper Bound is 0.0
The Total Outliers Lying in capital_gain Feature are 4035


The Lower Bound is 0.0 and the Upper Bound is 0.0
The Total Outliers Lying in capital_loss Feature are 2282


The Lower Bound is 32.5 and the Upper Bound is 52.5
The Total Outliers Lying in hours_per_week Feature are 529
```

[45]:
```
sns.pairplot(df)
```

[45]:
```
<seaborn.axisgrid.PairGrid at 0x195203a2bb0>
```



[46]:
```
fig, ax = plt.subplots(1,1, figsize=(10, 8))
sns.heatmap(df.corr(),annot=True)
plt.show()
```

[47]:
```
data=df.copy()
```

[48]:
```
df=data
```

[49]:
```
df.head()
```

[49]:

	age	work_class	final_weigth	education	education_num	\
0	39	State-gov	77516	Bachelors	13	
1	50	Self-emp-not-inc	83311	Bachelors	13	
2	38	Private	215646	HS-grad	9	
3	53	Private	234721	11th	7	
4	28	Private	338409	Bachelors	13	

	marital_status	occupation	relationship	race	sex	\
0	Never-married	Adm-clerical	Not-in-family	White	Male	
1	Married-civ-spouse	Exec-managerial	Husband	White	Male	
2	Divorced	Handlers-cleaners	Not-in-family	White	Male	
3	Married-civ-spouse	Handlers-cleaners	Husband	Black	Male	
4	Married-civ-spouse	Prof-specialty	Wife	Black	Female	

```
   capital_gain  capital_loss  hours_per_week native_country  Income
0          2174             0              40  United-States     1.0
1             0             0              13  United-States     1.0
2             0             0              40  United-States     1.0
3             0             0              40  United-States     1.0
4             0             0              40           Cuba     1.0
```

[50]:
```
from sklearn.preprocessing import LabelEncoder
label_encoder=LabelEncoder()
label_encoder
```

[50]:
```
LabelEncoder()
```

[51]:
```
cat_variable=df.dtypes[df.dtypes=='object'].index
```

[52]:
```
cat_variable
```

[52]:
```
Index(['work_class', 'education', 'marital_status', 'occupation',
       'relationship', 'race', 'sex', 'native_country'],
      dtype='object')
```

[53]:
```
for i in cat_variable:
    df[i]=label_encoder.fit_transform(df[i])
```

[54]:
```
df.head()
```

[54]:
```
   age  work_class  final_weigth  education  education_num  marital_status  \
0   39           6         77516          9             13               4
1   50           5         83311          9             13               2
2   38           3        215646         11              9               0
3   53           3        234721          1              7               2
4   28           3        338409          9             13               2

   occupation  relationship  race  sex  capital_gain  capital_loss  \
0           0             1     4    1          2174             0
1           3             0     4    1             0             0
2           5             1     4    1             0             0
3           5             0     2    1             0             0
4           9             5     2    0             0             0

   hours_per_week  native_country  Income
0              40              39     1.0
1              13              39     1.0
2              40              39     1.0
3              40              39     1.0
4              40               5     1.0
```

[55]:
```
temp_df=df[~df['Income'].isna()]
```

[56]:
```
temp_df.isna().sum().sum()
```

[56]:
```
0
```

[57]:
```
X=temp_df.drop('Income',axis=1)
```

[58]:
```
y=temp_df['Income']
```

[59]:
```
from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.33,␣
↪random_state=42)
```

[60]:
```
print(f"The Shape of the Training Data is {X_train.shape} and the output␣
↪Feature is {y_train.shape}")
print(f"The Shape of the Testing Data is {X_test.shape} and the output Feature␣
↪is {y_test.shape}")
```

```
The Shape of the Training Data is (21815, 14) and the output Feature is (21815,)
The Shape of the Testing Data is (10746, 14) and the output Feature is (10746,)
```

[61]:
```
from sklearn.preprocessing import StandardScaler
scalar=StandardScaler()
scalar
```

[61]:
```
StandardScaler()
```

[62]:
```
scaled_X_train=scalar.fit_transform(X_train)
```

[63]:
```
scaled_X_test=scalar.transform(X_test)
```

1.0.4 Logistic Regression

[64]:
```
from sklearn.linear_model import LogisticRegression
regressor=LogisticRegression()
regressor
```

[64]:
```
LogisticRegression()
```

[65]:
```
regressor.fit(scaled_X_train,y_train)
```

[65]:
```
LogisticRegression()
```

[66]:
```
regressor.score(scaled_X_train,y_train)
```

[66]:
```
0.8238826495530598
```

[67]:
```
regressor.intercept_
```

[67]:
```
array([1.50602965])
```

[68]:
```
regressor.coef_
```

[68]:
```
array([[-0.47602821,  0.11068587, -0.05354157, -0.03295505, -0.8584081 ,
         0.34777154,  0.01931966,  0.17755024, -0.10123872, -0.43539399,
        -2.24177563, -0.26137456, -0.3825377 , -0.03950097]])
```

[69]:
```
y1_pred=regressor.predict(scaled_X_test)
```

[70]:
```
from sklearn.metrics import classification_report
```

[71]:
```
print(classification_report(y_test,y1_pred))
```

```
              precision    recall  f1-score   support

         0.0       0.71      0.45      0.55      2550
         1.0       0.85      0.94      0.89      8196

    accuracy                           0.83     10746
   macro avg       0.78      0.70      0.72     10746
weighted avg       0.81      0.83      0.81     10746
```

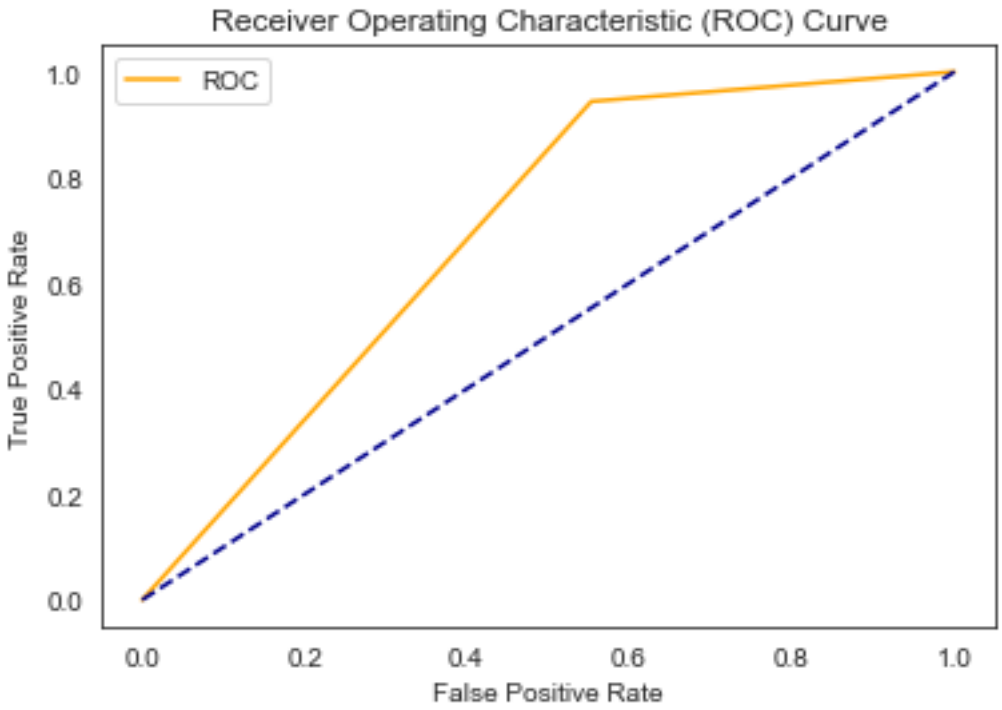
[72]:
```
from sklearn.metrics import roc_curve
from sklearn.metrics import roc_auc_score
```

[81]:
```
ytrain_pred = regressor.predict_proba(scaled_X_train)
print('Logistic train roc-auc: {}'.format(roc_auc_score(y_train, ytrain_pred[:
 ↪,1])))
ytest_pred = regressor.predict_proba(scaled_X_test)
print('Logistic test roc-auc: {}'.format(roc_auc_score(y_test, ytest_pred[:
 ↪,1])))
```

```
Logistic train roc-auc: 0.8551427358616901
Logistic test roc-auc: 0.8541286041014746
```

1.0.5 ROC Curve

[84]:
```
fpr, tpr, thresholds = roc_curve(y_test, y1_pred)
def plot_roc_curve(fpr, tpr):
    plt.plot(fpr, tpr, color='orange', label='ROC')
    plt.plot([0, 1], [0, 1], color='darkblue', linestyle='--')
    plt.xlabel('False Positive Rate')
    plt.ylabel('True Positive Rate')
    plt.title('Receiver Operating Characteristic (ROC) Curve')
    plt.legend()
    plt.show()

plot_roc_curve(fpr,tpr)
```



[80]:
```
from sklearn.metrics import accuracy_score
accuracy_score(y1_pred,y_test)
```

[80]: 0.8259817606551275

1.0.6 Support Vector Classifier

[74]:
```
from sklearn.svm import SVC
model=SVC(kernel='linear',random_state=42)
model
```

[74]: SVC(kernel='linear', random_state=42)

[75]:
```
model.fit(scaled_X_train,y_train)
```

[75]: SVC(kernel='linear', random_state=42)

[76]:
```
model.score(scaled_X_train,y_train)
```

[76]: 0.8136603254641301

[77]:
```
y_pred=model.predict(scaled_X_test)
```

[78]:
```
from sklearn.metrics import accuracy_score
accuracy_score(y_pred,y_test)
```

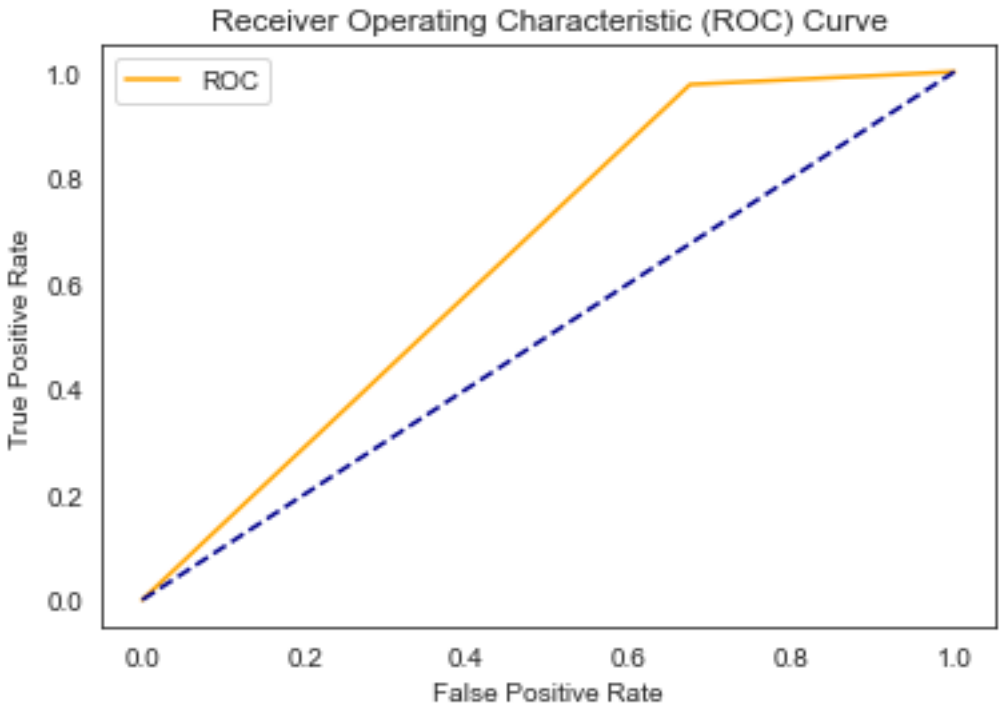
[78]:
```
0.821608040201005
```

[94]:
```
print(classification_report(y_test,y_pred))
```

```
              precision    recall  f1-score   support

         0.0       0.81      0.33      0.46      2550
         1.0       0.82      0.98      0.89      8196

    accuracy                           0.82     10746
   macro avg       0.82      0.65      0.68     10746
weighted avg       0.82      0.82      0.79     10746
```

1.0.7 ROC Curve

[85]:
```
fpr, tpr, thresholds = roc_curve(y_test, y_pred)
def plot_roc_curve(fpr, tpr):
    plt.plot(fpr, tpr, color='orange', label='ROC')
    plt.plot([0, 1], [0, 1], color='darkblue', linestyle='--')
    plt.xlabel('False Positive Rate')
    plt.ylabel('True Positive Rate')
    plt.title('Receiver Operating Characteristic (ROC) Curve')
    plt.legend()
    plt.show()

plot_roc_curve(fpr,tpr)
```



1.0.8 Decision Tree Classifier

[120]:
```
from sklearn.tree import DecisionTreeClassifier
model=DecisionTreeClassifier()
model.fit(X_train,y_train)
```

[120]:
```
DecisionTreeClassifier()
```

[121]:
```
model.score(X_train,y_train)
```

[121]:
```
1.0
```

[90]:
```
from sklearn import tree
import matplotlib.pyplot as plt
fig=plt.figure(figsize=(25,15))
tree.plot_tree(model,filled=True)
```

[90]:
```
[Text(0.7128216098235904, 0.9895833333333334, 'X[7] <= 0.5\ngini =
0.367\nsamples = 21815\nvalue = [5291, 16524]'),
 Text(0.488807188964246, 0.96875, 'X[4] <= 12.5\ngini = 0.495\nsamples =
8887\nvalue = [4002, 4885]'),
 Text(0.37291392777961574, 0.9479166666666666, 'X[10] <= 5095.5\ngini =
0.442\nsamples = 6177\nvalue = [2041, 4136]'),
 Text(0.28051118171175193, 0.9270833333333334, 'X[11] <= 1782.5\ngini =
```

```
0.417\nsamples = 5868\nvalue = [1739, 4129]'),
 Text(0.10172389464636207, 0.90625, 'X[4] <= 7.5\ngini = 0.402\nsamples =
5640\nvalue = [1574, 4066]'),
 Text(0.021185107677998673, 0.8854166666666666, 'X[3] <= 0.5\ngini =
0.179\nsamples = 874\nvalue = [87, 787]'),
 Text(0.010914835559385322, 0.8645833333333334, 'X[6] <= 11.5\ngini =
0.277\nsamples = 193\nvalue = [32, 161]'),
 Text(0.008261354232918214, 0.84375, 'X[2] <= 199643.5\ngini = 0.23\nsamples =
158\nvalue = [21, 137]'),
 Text(0.006018205070337772, 0.8229166666666666, 'X[2] <= 198099.5\ngini =
0.301\nsamples = 92\nvalue = [17, 75]'),
 Text(0.005580517428858661, 0.8020833333333334, 'X[6] <= 8.0\ngini =
0.278\nsamples = 90\nvalue = [15, 75]'),
 Text(0.00514282978737955, 0.78125, 'X[2] <= 160414.5\ngini = 0.317\nsamples =
76\nvalue = [15, 61]'),
 Text(0.0024072820281351086, 0.7604166666666666, 'X[12] <= 26.0\ngini =
0.233\nsamples = 52\nvalue = [7, 45]'),
 Text(0.0008753752829582214, 0.7395833333333334, 'X[0] <= 67.0\ngini =
0.5\nsamples = 2\nvalue = [1, 1]'),
 Text(0.0004376876414791107, 0.71875, 'gini = 0.0\nsamples = 1\nvalue = [1,
0]'),
 Text(0.001313062924437332, 0.71875, 'gini = 0.0\nsamples = 1\nvalue = [0, 1]'),
 Text(0.003939188773311996, 0.7395833333333334, 'X[6] <= 3.5\ngini =
0.211\nsamples = 50\nvalue = [6, 44]'),
 Text(0.0021884382073955534, 0.71875, 'X[0] <= 46.5\ngini = 0.311\nsamples =
26\nvalue = [5, 21]'),
 Text(0.0008753752829582214, 0.6979166666666666, 'X[12] <= 42.5\ngini =
0.133\nsamples = 14\nvalue = [1, 13]'),
 Text(0.0004376876414791107, 0.6770833333333334, 'gini = 0.0\nsamples = 9\nvalue
= [0, 9]'),
 Text(0.001313062924437332, 0.6770833333333334, 'X[12] <= 47.5\ngini =
0.32\nsamples = 5\nvalue = [1, 4]'),
 Text(0.0008753752829582214, 0.65625, 'X[6] <= 2.5\ngini = 0.5\nsamples =
2\nvalue = [1, 1]'),
 Text(0.0004376876414791107, 0.6354166666666666, 'gini = 0.0\nsamples = 1\nvalue
= [1, 0]'),
 Text(0.001313062924437332, 0.6354166666666666, 'gini = 0.0\nsamples = 1\nvalue
= [0, 1]'),
 Text(0.0017507505659164427, 0.65625, 'gini = 0.0\nsamples = 3\nvalue = [0,
3]'),
 Text(0.0035015011318328855, 0.6979166666666666, 'X[12] <= 49.0\ngini =
0.444\nsamples = 12\nvalue = [4, 8]'),
 Text(0.003063813490353775, 0.6770833333333334, 'X[6] <= 2.5\ngini =
0.397\nsamples = 11\nvalue = [3, 8]'),
 Text(0.002626125848874664, 0.65625, 'X[2] <= 125108.0\ngini = 0.32\nsamples =
10\nvalue = [2, 8]'),
 Text(0.0021884382073955534, 0.6354166666666666, 'X[2] <= 70240.5\ngini =
```

```
0.444\nsamples = 6\nvalue = [2, 4]'),
 Text(0.0017507505659164427, 0.6145833333333334, 'gini = 0.0\nsamples = 3\nvalue
= [0, 3]'),
 Text(0.002626125848874664, 0.6145833333333334, 'X[1] <= 3.5\ngini =
0.444\nsamples = 3\nvalue = [2, 1]'),
 Text(0.0021884382073955534, 0.59375, 'gini = 0.0\nsamples = 2\nvalue = [2,
0]'),
 Text(0.0030063813490353775, 0.59375, 'gini = 0.0\nsamples = 1\nvalue = [0, 1]'),
 Text(0.0030063813490353775, 0.6354166666666666, 'gini = 0.0\nsamples = 4\nvalue
= [0, 4]'),
 Text(0.0035015011318328855, 0.65625, 'gini = 0.0\nsamples = 1\nvalue = [1,
0]'),
 Text(0.003939188773311996, 0.6770833333333334, 'gini = 0.0\nsamples = 1\nvalue
= [1, 0]'),
 Text(0.0056899393392284385, 0.71875, 'X[0] <= 41.0\ngini = 0.08\nsamples =
24\nvalue = [1, 23]'),
 Text(0.005252251697749328, 0.6979166666666666, 'X[0] <= 39.5\ngini =
0.32\nsamples = 5\nvalue = [1, 4]'),
 Text(0.004814564056270217, 0.6770833333333334, 'gini = 0.0\nsamples = 4\nvalue
= [0, 4]'),
 Text(0.0056899393392284385, 0.6770833333333334, 'gini = 0.0\nsamples = 1\nvalue
= [1, 0]'),
 Text(0.00612762698070755, 0.6979166666666666, 'gini = 0.0\nsamples = 19\nvalue
= [0, 19]'),
 Text(0.007878377546623992, 0.7604166666666666, 'X[2] <= 165158.5\ngini =
0.444\nsamples = 24\nvalue = [8, 16]'),
 Text(0.007440689905144881, 0.7395833333333334, 'gini = 0.0\nsamples = 3\nvalue
= [3, 0]'),
 Text(0.008316065188103103, 0.7395833333333334, 'X[6] <= 5.5\ngini =
0.363\nsamples = 21\nvalue = [5, 16]'),
 Text(0.007440689905144881, 0.71875, 'X[0] <= 32.5\ngini = 0.153\nsamples =
12\nvalue = [1, 11]'),
 Text(0.007003002263665771, 0.6979166666666666, 'gini = 0.0\nsamples = 1\nvalue
= [1, 0]'),
 Text(0.007878377546623992, 0.6979166666666666, 'gini = 0.0\nsamples = 11\nvalue
= [0, 11]'),
 Text(0.009191440471061324, 0.71875, 'X[2] <= 171894.0\ngini = 0.494\nsamples =
9\nvalue = [4, 5]'),
 Text(0.008753752829582214, 0.6979166666666666, 'gini = 0.0\nsamples = 2\nvalue
= [2, 0]'),
 Text(0.009629128112540434, 0.6979166666666666, 'X[6] <= 6.5\ngini =
0.408\nsamples = 7\nvalue = [2, 5]'),
 Text(0.009191440471061324, 0.6770833333333334, 'X[12] <= 43.5\ngini =
0.278\nsamples = 6\nvalue = [1, 5]'),
 Text(0.008753752829582214, 0.65625, 'gini = 0.0\nsamples = 5\nvalue = [0, 5]'),
 Text(0.009629128112540434, 0.65625, 'gini = 0.0\nsamples = 1\nvalue = [1, 0]'),
 Text(0.010066815754019546, 0.6770833333333334, 'gini = 0.0\nsamples = 1\nvalue
```

```
= [1, 0]'),
 Text(0.006018205070337772, 0.78125, 'gini = 0.0\nsamples = 14\nvalue = [0,
14]'),
 Text(0.006455892711816883, 0.8020833333333334, 'gini = 0.0\nsamples = 2\nvalue
= [2, 0]'),
 Text(0.010504503395498656, 0.8229166666666666, 'X[2] <= 383987.5\ngini =
0.114\nsamples = 66\nvalue = [4, 62]'),
 Text(0.009629128112540434, 0.8020833333333334, 'X[12] <= 57.5\ngini =
0.034\nsamples = 57\nvalue = [1, 56]'),
 Text(0.009191440471061324, 0.78125, 'gini = 0.0\nsamples = 52\nvalue = [0,
52]'),
 Text(0.010066815754019546, 0.78125, 'X[12] <= 65.0\ngini = 0.32\nsamples =
5\nvalue = [1, 4]'),
 Text(0.009629128112540434, 0.7604166666666666, 'X[1] <= 2.0\ngini =
0.5\nsamples = 2\nvalue = [1, 1]'),
 Text(0.009191440471061324, 0.7395833333333334, 'gini = 0.0\nsamples = 1\nvalue
= [0, 1]'),
 Text(0.010066815754019546, 0.7395833333333334, 'gini = 0.0\nsamples = 1\nvalue
= [1, 0]'),
 Text(0.010504503395498656, 0.7604166666666666, 'gini = 0.0\nsamples = 3\nvalue
= [0, 3]'),
 Text(0.011379878678456877, 0.8020833333333334, 'X[2] <= 447239.5\ngini =
0.444\nsamples = 9\nvalue = [3, 6]'),
 Text(0.010942191036977767, 0.78125, 'gini = 0.0\nsamples = 2\nvalue = [2, 0]'),
 Text(0.011817566319935989, 0.78125, 'X[1] <= 5.0\ngini = 0.245\nsamples =
7\nvalue = [1, 6]'),
 Text(0.011379878678456877, 0.7604166666666666, 'gini = 0.0\nsamples = 6\nvalue
= [0, 6]'),
 Text(0.0122552539614151, 0.7604166666666666, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.01356831688585243, 0.84375, 'X[2] <= 59918.5\ngini = 0.431\nsamples =
35\nvalue = [11, 24]'),
 Text(0.01313062924437332, 0.8229166666666666, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.014006004527331542, 0.8229166666666666, 'X[6] <= 12.5\ngini =
0.415\nsamples = 34\nvalue = [10, 24]'),
 Text(0.01356831688585243, 0.8020833333333334, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.014443692168810652, 0.8020833333333334, 'X[0] <= 45.0\ngini =
0.397\nsamples = 33\nvalue = [9, 24]'),
 Text(0.01356831688585243, 0.78125, 'X[8] <= 3.0\ngini = 0.165\nsamples =
11\nvalue = [1, 10]'),
 Text(0.01313062924437332, 0.7604166666666666, 'X[12] <= 45.0\ngini =
0.5\nsamples = 2\nvalue = [1, 1]'),
 Text(0.01269294160289421, 0.7395833333333334, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.01356831688585243, 0.7395833333333334, 'gini = 0.0\nsamples = 1\nvalue =
```

```
[0, 1]'),
 Text(0.014006004527331542, 0.7604166666666666, 'gini = 0.0\nsamples = 9\nvalue
= [0, 9]'),
 Text(0.015319067451768873, 0.78125, 'X[12] <= 25.0\ngini = 0.463\nsamples =
22\nvalue = [8, 14]'),
 Text(0.014881379810289763, 0.7604166666666666, 'gini = 0.0\nsamples = 3\nvalue
= [0, 3]'),
 Text(0.015756755093247983, 0.7604166666666666, 'X[12] <= 43.5\ngini =
0.488\nsamples = 19\nvalue = [8, 11]'),
 Text(0.014881379810289763, 0.7395833333333334, 'X[2] <= 97944.0\ngini =
0.496\nsamples = 11\nvalue = [6, 5]'),
 Text(0.014443692168810652, 0.71875, 'gini = 0.0\nsamples = 1\nvalue = [0, 1]'),
 Text(0.015319067451768873, 0.71875, 'X[2] <= 433933.0\ngini = 0.48\nsamples =
10\nvalue = [6, 4]'),
 Text(0.014881379810289763, 0.6979166666666666, 'X[13] <= 36.0\ngini =
0.444\nsamples = 9\nvalue = [6, 3]'),
 Text(0.014443692168810652, 0.6770833333333334, 'gini = 0.0\nsamples = 1\nvalue
= [0, 1]'),
 Text(0.015319067451768873, 0.6770833333333334, 'X[0] <= 55.0\ngini =
0.375\nsamples = 8\nvalue = [6, 2]'),
 Text(0.014881379810289763, 0.65625, 'gini = 0.0\nsamples = 3\nvalue = [3, 0]'),
 Text(0.015756755093247983, 0.65625, 'X[0] <= 58.5\ngini = 0.48\nsamples =
5\nvalue = [3, 2]'),
 Text(0.015319067451768873, 0.6354166666666666, 'gini = 0.0\nsamples = 1\nvalue
= [0, 1]'),
 Text(0.016194442734727095, 0.6354166666666666, 'X[0] <= 60.0\ngini =
0.375\nsamples = 4\nvalue = [3, 1]'),
 Text(0.015756755093247983, 0.6145833333333334, 'gini = 0.0\nsamples = 2\nvalue
= [2, 0]'),
 Text(0.016632130376206207, 0.6145833333333334, 'X[2] <= 177866.0\ngini =
0.5\nsamples = 2\nvalue = [1, 1]'),
 Text(0.016194442734727095, 0.59375, 'gini = 0.0\nsamples = 1\nvalue = [0, 1]'),
 Text(0.017069818017685316, 0.59375, 'gini = 0.0\nsamples = 1\nvalue = [1, 0]'),
 Text(0.015756755093247983, 0.6979166666666666, 'gini = 0.0\nsamples = 1\nvalue
= [0, 1]'),
 Text(0.016632130376206207, 0.7395833333333334, 'X[0] <= 57.0\ngini =
0.375\nsamples = 8\nvalue = [2, 6]'),
 Text(0.016194442734727095, 0.71875, 'gini = 0.0\nsamples = 5\nvalue = [0, 5]'),
 Text(0.017069818017685316, 0.71875, 'X[0] <= 58.5\ngini = 0.444\nsamples =
3\nvalue = [2, 1]'),
 Text(0.016632130376206207, 0.6979166666666666, 'gini = 0.0\nsamples = 1\nvalue
= [1, 0]'),
 Text(0.017507505659164427, 0.6979166666666666, 'X[2] <= 204078.5\ngini =
0.5\nsamples = 2\nvalue = [1, 1]'),
 Text(0.017069818017685316, 0.6770833333333334, 'gini = 0.0\nsamples = 1\nvalue
= [1, 0]'),
 Text(0.017945193300643536, 0.6770833333333334, 'gini = 0.0\nsamples = 1\nvalue
```

```
= [0, 1]'),
 Text(0.031455379796612025, 0.8645833333333334, 'X[12] <= 43.5\ngini =
0.148\nsamples = 681\nvalue = [55, 626]'),
 Text(0.02154927747344809, 0.84375, 'X[2] <= 30970.0\ngini = 0.112\nsamples =
522\nvalue = [31, 491]'),
 Text(0.017671638524719094, 0.8229166666666666, 'X[0] <= 61.5\ngini =
0.49\nsamples = 7\nvalue = [3, 4]'),
 Text(0.017233950883239983, 0.8020833333333334, 'X[2] <= 29405.5\ngini =
0.375\nsamples = 4\nvalue = [3, 1]'),
 Text(0.01679626324176087, 0.78125, 'gini = 0.0\nsamples = 1\nvalue = [0, 1]'),
 Text(0.017671638524719094, 0.78125, 'gini = 0.0\nsamples = 3\nvalue = [3, 0]'),
 Text(0.018109326166198203, 0.8020833333333334, 'gini = 0.0\nsamples = 3\nvalue
= [0, 3]'),
 Text(0.025426916422177086, 0.8229166666666666, 'X[6] <= 3.5\ngini =
0.103\nsamples = 515\nvalue = [28, 487]'),
 Text(0.01964123291137509, 0.8020833333333334, 'X[8] <= 2.5\ngini =
0.183\nsamples = 147\nvalue = [15, 132]'),
 Text(0.018547013807677315, 0.78125, 'X[8] <= 1.5\ngini = 0.5\nsamples =
8\nvalue = [4, 4]'),
 Text(0.018109326166198203, 0.7604166666666666, 'gini = 0.0\nsamples = 4\nvalue
= [0, 4]'),
 Text(0.018984701449156427, 0.7604166666666666, 'gini = 0.0\nsamples = 4\nvalue
= [4, 0]'),
 Text(0.020735452015072868, 0.78125, 'X[2] <= 524278.5\ngini = 0.146\nsamples =
139\nvalue = [11, 128]'),
 Text(0.019860076732114647, 0.7604166666666666, 'X[12] <= 39.0\ngini =
0.135\nsamples = 137\nvalue = [10, 127]'),
 Text(0.019422389090635536, 0.7395833333333334, 'gini = 0.0\nsamples = 32\nvalue
= [0, 32]'),
 Text(0.020297764373593756, 0.7395833333333334, 'X[13] <= 1.0\ngini =
0.172\nsamples = 105\nvalue = [10, 95]'),
 Text(0.01882056858360176, 0.71875, 'X[0] <= 46.0\ngini = 0.5\nsamples =
2\nvalue = [1, 1]'),
 Text(0.018382880942122648, 0.6979166666666666, 'gini = 0.0\nsamples = 1\nvalue
= [1, 0]'),
 Text(0.01925825622508087, 0.6979166666666666, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.021774960163585756, 0.71875, 'X[6] <= 1.0\ngini = 0.159\nsamples =
103\nvalue = [9, 94]'),
 Text(0.020133631508039092, 0.6979166666666666, 'X[4] <= 6.0\ngini =
0.375\nsamples = 8\nvalue = [2, 6]'),
 Text(0.01969594386655998, 0.6770833333333334, 'gini = 0.0\nsamples = 5\nvalue =
[0, 5]'),
 Text(0.0205713191495182, 0.6770833333333334, 'X[0] <= 46.5\ngini =
0.444\nsamples = 3\nvalue = [2, 1]'),
 Text(0.020133631508039092, 0.65625, 'gini = 0.0\nsamples = 2\nvalue = [2, 0]'),
 Text(0.021009006790997313, 0.65625, 'gini = 0.0\nsamples = 1\nvalue = [0, 1]'),
```

```
 Text(0.023416288819132422, 0.6979166666666666, 'X[0] <= 51.5\ngini =
0.137\nsamples = 95\nvalue = [7, 88]'),
 Text(0.022322069715434645, 0.6770833333333334, 'X[2] <= 155998.0\ngini =
0.062\nsamples = 62\nvalue = [2, 60]'),
 Text(0.021884382073955534, 0.65625, 'X[2] <= 154487.5\ngini = 0.172\nsamples =
21\nvalue = [2, 19]'),
 Text(0.02144669443247642, 0.6354166666666666, 'X[3] <= 2.5\ngini =
0.095\nsamples = 20\nvalue = [1, 19]'),
 Text(0.021009006790997313, 0.6145833333333334, 'X[2] <= 116592.5\ngini =
0.219\nsamples = 8\nvalue = [1, 7]'),
 Text(0.0205713191495182, 0.59375, 'X[0] <= 36.5\ngini = 0.444\nsamples =
3\nvalue = [1, 2]'),
 Text(0.020133631508039092, 0.5729166666666666, 'gini = 0.0\nsamples = 2\nvalue
= [0, 2]'),
 Text(0.021009006790997313, 0.5729166666666666, 'gini = 0.0\nsamples = 1\nvalue
= [1, 0]'),
 Text(0.02144669443247642, 0.59375, 'gini = 0.0\nsamples = 5\nvalue = [0, 5]'),
 Text(0.021884382073955534, 0.6145833333333334, 'gini = 0.0\nsamples = 12\nvalue
= [0, 12]'),
 Text(0.022322069715434645, 0.6354166666666666, 'gini = 0.0\nsamples = 1\nvalue
= [1, 0]'),
 Text(0.022759757356913754, 0.65625, 'gini = 0.0\nsamples = 41\nvalue = [0,
41]'),
 Text(0.0245105079228302, 0.6770833333333334, 'X[0] <= 57.5\ngini =
0.257\nsamples = 33\nvalue = [5, 28]'),
 Text(0.023635132639871978, 0.65625, 'X[2] <= 36546.0\ngini = 0.444\nsamples =
12\nvalue = [4, 8]'),
 Text(0.023197444998392866, 0.6354166666666666, 'gini = 0.0\nsamples = 1\nvalue
= [1, 0]'),
 Text(0.024072820281351087, 0.6354166666666666, 'X[2] <= 137678.5\ngini =
0.397\nsamples = 11\nvalue = [3, 8]'),
 Text(0.023635132639871978, 0.6145833333333334, 'gini = 0.0\nsamples = 5\nvalue
= [0, 5]'),
 Text(0.0245105079228302, 0.6145833333333334, 'X[2] <= 200162.5\ngini =
0.5\nsamples = 6\nvalue = [3, 3]'),
 Text(0.024072820281351087, 0.59375, 'gini = 0.0\nsamples = 2\nvalue = [2, 0]'),
 Text(0.024948195564309307, 0.59375, 'X[6] <= 2.5\ngini = 0.375\nsamples =
4\nvalue = [1, 3]'),
 Text(0.0245105079228302, 0.5729166666666666, 'gini = 0.0\nsamples = 3\nvalue =
[0, 3]'),
 Text(0.02538588320578842, 0.5729166666666666, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.02538588320578842, 0.65625, 'X[0] <= 65.5\ngini = 0.091\nsamples =
21\nvalue = [1, 20]'),
 Text(0.024948195564309307, 0.6354166666666666, 'gini = 0.0\nsamples = 17\nvalue
= [0, 17]'),
 Text(0.02582357084726753, 0.6354166666666666, 'X[2] <= 126168.0\ngini =
```

```
0.375\nsamples = 4\nvalue = [1, 3]'),
 Text(0.02538588320578842, 0.6145833333333334, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.02626125848874664, 0.6145833333333334, 'gini = 0.0\nsamples = 3\nvalue =
[0, 3]'),
 Text(0.02161082729803109, 0.7604166666666666, 'X[13] <= 26.5\ngini =
0.5\nsamples = 2\nvalue = [1, 1]'),
 Text(0.02117313965655198, 0.7395833333333334, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.0220485149395102, 0.7395833333333334, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.03121259993297908, 0.8020833333333334, 'X[1] <= 3.5\ngini =
0.068\nsamples = 368\nvalue = [13, 355]'),
 Text(0.02850440765132708, 0.78125, 'X[2] <= 90590.0\ngini = 0.048\nsamples =
327\nvalue = [8, 319]'),
 Text(0.02626125848874664, 0.7604166666666666, 'X[2] <= 90062.5\ngini =
0.165\nsamples = 33\nvalue = [3, 30]'),
 Text(0.02582357084726753, 0.7395833333333334, 'X[3] <= 4.5\ngini =
0.117\nsamples = 32\nvalue = [2, 30]'),
 Text(0.02538588320578842, 0.71875, 'X[0] <= 52.5\ngini = 0.231\nsamples =
15\nvalue = [2, 13]'),
 Text(0.024948195564309307, 0.6979166666666666, 'gini = 0.0\nsamples = 9\nvalue
= [0, 9]'),
 Text(0.02582357084726753, 0.6979166666666666, 'X[0] <= 61.5\ngini =
0.444\nsamples = 6\nvalue = [2, 4]'),
 Text(0.02538588320578842, 0.6770833333333334, 'gini = 0.0\nsamples = 2\nvalue =
[2, 0]'),
 Text(0.02626125848874664, 0.6770833333333334, 'gini = 0.0\nsamples = 4\nvalue =
[0, 4]'),
 Text(0.02626125848874664, 0.71875, 'gini = 0.0\nsamples = 17\nvalue = [0,
17]'),
 Text(0.02669894613022575, 0.7395833333333334, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.030747556813907524, 0.7604166666666666, 'X[2] <= 439216.5\ngini =
0.033\nsamples = 294\nvalue = [5, 289]'),
 Text(0.02954391579983997, 0.7395833333333334, 'X[3] <= 3.5\ngini =
0.028\nsamples = 283\nvalue = [4, 279]'),
 Text(0.028449696696142193, 0.71875, 'X[6] <= 10.0\ngini = 0.054\nsamples =
109\nvalue = [3, 106]'),
 Text(0.027574321413183972, 0.6979166666666666, 'X[2] <= 237062.5\ngini =
0.024\nsamples = 81\nvalue = [1, 80]'),
 Text(0.02713663377170486, 0.6770833333333334, 'gini = 0.0\nsamples = 60\nvalue
= [0, 60]'),
 Text(0.028012009054663084, 0.6770833333333334, 'X[2] <= 239154.5\ngini =
0.091\nsamples = 21\nvalue = [1, 20]'),
 Text(0.027574321413183972, 0.65625, 'gini = 0.0\nsamples = 1\nvalue = [1, 0]'),
 Text(0.028449696696142193, 0.65625, 'gini = 0.0\nsamples = 20\nvalue = [0,
```

```
20]'),
 Text(0.029325071979100416, 0.6979166666666666, 'X[0] <= 47.5\ngini =
0.133\nsamples = 28\nvalue = [2, 26]'),
 Text(0.028887384337621305, 0.6770833333333334, 'gini = 0.0\nsamples = 13\nvalue
= [0, 13]'),
 Text(0.029762759620579525, 0.6770833333333334, 'X[0] <= 48.5\ngini =
0.231\nsamples = 15\nvalue = [2, 13]'),
 Text(0.029325071979100416, 0.65625, 'gini = 0.0\nsamples = 1\nvalue = [1, 0]'),
 Text(0.030200447262058637, 0.65625, 'X[2] <= 122612.5\ngini = 0.133\nsamples =
14\nvalue = [1, 13]'),
 Text(0.029762759620579525, 0.6354166666666666, 'X[2] <= 112890.0\ngini =
0.375\nsamples = 4\nvalue = [1, 3]'),
 Text(0.029325071979100416, 0.6145833333333334, 'gini = 0.0\nsamples = 3\nvalue
= [0, 3]'),
 Text(0.030200447262058637, 0.6145833333333334, 'gini = 0.0\nsamples = 1\nvalue
= [1, 0]'),
 Text(0.030638134903537746, 0.6354166666666666, 'gini = 0.0\nsamples = 10\nvalue
= [0, 10]'),
 Text(0.030638134903537746, 0.71875, 'X[2] <= 284216.5\ngini = 0.011\nsamples =
174\nvalue = [1, 173]'),
 Text(0.030200447262058637, 0.6979166666666666, 'gini = 0.0\nsamples =
144\nvalue = [0, 144]'),
 Text(0.031075822545016858, 0.6979166666666666, 'X[2] <= 287394.5\ngini =
0.064\nsamples = 30\nvalue = [1, 29]'),
 Text(0.030638134903537746, 0.6770833333333334, 'gini = 0.0\nsamples = 1\nvalue
= [1, 0]'),
 Text(0.031513510186495966, 0.6770833333333334, 'gini = 0.0\nsamples = 29\nvalue
= [0, 29]'),
 Text(0.03195119782797508, 0.7395833333333334, 'X[2] <= 445882.5\ngini =
0.165\nsamples = 11\nvalue = [1, 10]'),
 Text(0.031513510186495966, 0.71875, 'gini = 0.0\nsamples = 1\nvalue = [1, 0]'),
 Text(0.03238888546945419, 0.71875, 'gini = 0.0\nsamples = 10\nvalue = [0,
10]'),
 Text(0.033920792214631075, 0.78125, 'X[0] <= 51.5\ngini = 0.214\nsamples =
41\nvalue = [5, 36]'),
 Text(0.03348310457315197, 0.7604166666666666, 'gini = 0.0\nsamples = 15\nvalue
= [0, 15]'),
 Text(0.03435847985611019, 0.7604166666666666, 'X[0] <= 54.0\ngini =
0.311\nsamples = 26\nvalue = [5, 21]'),
 Text(0.033920792214631075, 0.7395833333333334, 'gini = 0.0\nsamples = 1\nvalue
= [1, 0]'),
 Text(0.0347961674975893, 0.7395833333333334, 'X[6] <= 10.0\ngini =
0.269\nsamples = 25\nvalue = [4, 21]'),
 Text(0.033264260752412414, 0.71875, 'X[2] <= 257786.0\ngini = 0.133\nsamples =
14\nvalue = [1, 13]'),
 Text(0.0328265731109333, 0.6979166666666666, 'gini = 0.0\nsamples = 12\nvalue =
[0, 12]'),
```

```
 Text(0.03370194839389152, 0.6979166666666666, 'X[0] <= 64.5\ngini =
0.5\nsamples = 2\nvalue = [1, 1]'),
 Text(0.033264260752412414, 0.6770833333333334, 'gini = 0.0\nsamples = 1\nvalue
= [0, 1]'),
 Text(0.03413963603537063, 0.6770833333333334, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.036328074242766184, 0.71875, 'X[2] <= 133348.5\ngini = 0.397\nsamples =
11\nvalue = [3, 8]'),
 Text(0.03545269895980797, 0.6979166666666666, 'X[1] <= 5.5\ngini =
0.444\nsamples = 3\nvalue = [2, 1]'),
 Text(0.035015011318328855, 0.6770833333333334, 'gini = 0.0\nsamples = 2\nvalue
= [2, 0]'),
 Text(0.03589038660128707, 0.6770833333333334, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.03720344952572441, 0.6979166666666666, 'X[2] <= 183592.0\ngini =
0.219\nsamples = 8\nvalue = [1, 7]'),
 Text(0.036765761884245296, 0.6770833333333334, 'gini = 0.0\nsamples = 4\nvalue
= [0, 4]'),
 Text(0.03764113716720352, 0.6770833333333334, 'X[2] <= 225343.5\ngini =
0.375\nsamples = 4\nvalue = [1, 3]'),
 Text(0.03720344952572441, 0.65625, 'gini = 0.0\nsamples = 1\nvalue = [1, 0]'),
 Text(0.03807882480868263, 0.65625, 'gini = 0.0\nsamples = 3\nvalue = [0, 3]'),
 Text(0.04136148211977596, 0.84375, 'X[0] <= 35.5\ngini = 0.256\nsamples =
159\nvalue = [24, 135]'),
 Text(0.03851651245016174, 0.8229166666666666, 'X[1] <= 3.5\ngini =
0.089\nsamples = 43\nvalue = [2, 41]'),
 Text(0.03807882480868263, 0.8020833333333334, 'gini = 0.0\nsamples = 34\nvalue
= [0, 34]'),
 Text(0.03895420009164085, 0.8020833333333334, 'X[1] <= 4.5\ngini =
0.346\nsamples = 9\nvalue = [2, 7]'),
 Text(0.03851651245016174, 0.78125, 'gini = 0.0\nsamples = 1\nvalue = [1, 0]'),
 Text(0.03939188773311996, 0.78125, 'X[2] <= 237936.5\ngini = 0.219\nsamples =
8\nvalue = [1, 7]'),
 Text(0.03895420009164085, 0.7604166666666666, 'gini = 0.0\nsamples = 7\nvalue =
[0, 7]'),
 Text(0.03982957537459907, 0.7604166666666666, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.04420645178939018, 0.8229166666666666, 'X[6] <= 10.5\ngini =
0.307\nsamples = 116\nvalue = [22, 94]'),
 Text(0.04245570122347374, 0.8020833333333334, 'X[0] <= 54.5\ngini =
0.236\nsamples = 88\nvalue = [12, 76]'),
 Text(0.042018013581994626, 0.78125, 'X[0] <= 50.5\ngini = 0.341\nsamples =
55\nvalue = [12, 43]'),
 Text(0.04070495065755729, 0.7604166666666666, 'X[13] <= 35.5\ngini =
0.25\nsamples = 41\nvalue = [6, 35]'),
 Text(0.040267263016078185, 0.7395833333333334, 'gini = 0.0\nsamples = 14\nvalue
= [0, 14]'),
```

```
 Text(0.0411426382990364, 0.7395833333333334, 'X[2] <= 332226.5\ngini =
0.346\nsamples = 27\nvalue = [6, 21]'),
 Text(0.04070495065755729, 0.71875, 'X[10] <= 1551.5\ngini = 0.311\nsamples =
26\nvalue = [5, 21]'),
 Text(0.040267263016078185, 0.6979166666666666, 'X[0] <= 46.5\ngini =
0.269\nsamples = 25\nvalue = [4, 21]'),
 Text(0.03982957537459907, 0.6770833333333334, 'X[0] <= 44.5\ngini =
0.375\nsamples = 16\nvalue = [4, 12]'),
 Text(0.03895420009164085, 0.65625, 'X[6] <= 5.5\ngini = 0.153\nsamples =
12\nvalue = [1, 11]'),
 Text(0.03851651245016174, 0.6354166666666666, 'gini = 0.0\nsamples = 10\nvalue
= [0, 10]'),
 Text(0.03939188773311996, 0.6354166666666666, 'X[0] <= 37.5\ngini =
0.5\nsamples = 2\nvalue = [1, 1]'),
 Text(0.03895420009164085, 0.6145833333333334, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.03982957537459907, 0.6145833333333334, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.04070495065755729, 0.65625, 'X[6] <= 2.5\ngini = 0.375\nsamples =
4\nvalue = [3, 1]'),
 Text(0.040267263016078185, 0.6354166666666666, 'gini = 0.0\nsamples = 1\nvalue
= [0, 1]'),
 Text(0.0411426382990364, 0.6354166666666666, 'gini = 0.0\nsamples = 3\nvalue =
[3, 0]'),
 Text(0.04070495065755729, 0.6770833333333334, 'gini = 0.0\nsamples = 9\nvalue =
[0, 9]'),
 Text(0.0411426382990364, 0.6979166666666666, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.041580325940515514, 0.71875, 'gini = 0.0\nsamples = 1\nvalue = [1, 0]'),
 Text(0.043331076506431955, 0.7604166666666666, 'X[2] <= 220719.5\ngini =
0.49\nsamples = 14\nvalue = [6, 8]'),
 Text(0.04289338886495284, 0.7395833333333334, 'X[3] <= 3.5\ngini =
0.48\nsamples = 10\nvalue = [6, 4]'),
 Text(0.04245570122347374, 0.71875, 'gini = 0.0\nsamples = 4\nvalue = [4, 0]'),
 Text(0.043331076506431955, 0.71875, 'X[6] <= 3.5\ngini = 0.444\nsamples =
6\nvalue = [2, 4]'),
 Text(0.04289338886495284, 0.6979166666666666, 'X[3] <= 5.0\ngini =
0.444\nsamples = 3\nvalue = [2, 1]'),
 Text(0.04245570122347374, 0.6770833333333334, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.043331076506431955, 0.6770833333333334, 'gini = 0.0\nsamples = 2\nvalue
= [2, 0]'),
 Text(0.04376876414791107, 0.6979166666666666, 'gini = 0.0\nsamples = 3\nvalue =
[0, 3]'),
 Text(0.04376876414791107, 0.7395833333333334, 'gini = 0.0\nsamples = 4\nvalue =
[0, 4]'),
 Text(0.04289338886495284, 0.78125, 'gini = 0.0\nsamples = 33\nvalue = [0,
```

```
33]'),
 Text(0.04595720235530662, 0.8020833333333334, 'X[2] <= 265004.0\ngini =
0.459\nsamples = 28\nvalue = [10, 18]'),
 Text(0.04551951471382751, 0.78125, 'X[2] <= 159629.0\ngini = 0.403\nsamples =
25\nvalue = [7, 18]'),
 Text(0.045081827072348396, 0.7604166666666666, 'X[3] <= 5.5\ngini =
0.5\nsamples = 14\nvalue = [7, 7]'),
 Text(0.04464413943086929, 0.7395833333333334, 'X[2] <= 112563.5\ngini =
0.463\nsamples = 11\nvalue = [4, 7]'),
 Text(0.04420645178939018, 0.71875, 'gini = 0.0\nsamples = 5\nvalue = [0, 5]'),
 Text(0.045081827072348396, 0.71875, 'X[3] <= 3.5\ngini = 0.444\nsamples =
6\nvalue = [4, 2]'),
 Text(0.04464413943086929, 0.6979166666666666, 'X[12] <= 65.0\ngini =
0.444\nsamples = 3\nvalue = [1, 2]'),
 Text(0.04420645178939018, 0.6770833333333334, 'gini = 0.0\nsamples = 2\nvalue =
[0, 2]'),
 Text(0.045081827072348396, 0.6770833333333334, 'gini = 0.0\nsamples = 1\nvalue
= [1, 0]'),
 Text(0.04551951471382751, 0.6979166666666666, 'gini = 0.0\nsamples = 3\nvalue =
[3, 0]'),
 Text(0.04551951471382751, 0.7395833333333334, 'gini = 0.0\nsamples = 3\nvalue =
[3, 0]'),
 Text(0.04595720235530662, 0.7604166666666666, 'gini = 0.0\nsamples = 11\nvalue
= [0, 11]'),
 Text(0.04639488999678573, 0.78125, 'gini = 0.0\nsamples = 3\nvalue = [3, 0]'),
 Text(0.18226268161472545, 0.8854166666666666, 'X[0] <= 36.5\ngini =
0.429\nsamples = 4766\nvalue = [1487, 3279]'),
 Text(0.09514551671505167, 0.8645833333333334, 'X[4] <= 9.5\ngini =
0.322\nsamples = 1698\nvalue = [343, 1355]'),
 Text(0.07212477367616586, 0.84375, 'X[12] <= 51.5\ngini = 0.245\nsamples =
952\nvalue = [136, 816]'),
 Text(0.054523580849114024, 0.8229166666666666, 'X[0] <= 28.5\ngini =
0.21\nsamples = 820\nvalue = [98, 722]'),
 Text(0.05144197561259173, 0.8020833333333334, 'X[5] <= 1.5\ngini =
0.117\nsamples = 273\nvalue = [17, 256]'),
 Text(0.051004287971112616, 0.78125, 'gini = 0.0\nsamples = 1\nvalue = [1, 0]'),
 Text(0.05187966325407084, 0.78125, 'X[10] <= 4225.0\ngini = 0.111\nsamples =
272\nvalue = [16, 256]'),
 Text(0.05144197561259173, 0.7604166666666666, 'X[1] <= 5.5\ngini =
0.105\nsamples = 271\nvalue = [15, 256]'),
 Text(0.049321926099177286, 0.7395833333333334, 'X[12] <= 44.5\ngini =
0.093\nsamples = 267\nvalue = [13, 254]'),
 Text(0.047270265279743956, 0.71875, 'X[10] <= 2994.0\ngini = 0.058\nsamples =
200\nvalue = [6, 194]'),
 Text(0.04639488999678573, 0.6979166666666666, 'X[6] <= 2.5\ngini =
0.05\nsamples = 196\nvalue = [5, 191]'),
 Text(0.04595720235530662, 0.6770833333333334, 'X[2] <= 178071.0\ngini =
```

```
0.123\nsamples = 76\nvalue = [5, 71]'),
 Text(0.04551951471382751, 0.65625, 'X[2] <= 176559.0\ngini = 0.245\nsamples =
35\nvalue = [5, 30]'),
 Text(0.045081827072348396, 0.6354166666666666, 'X[4] <= 8.5\ngini =
0.208\nsamples = 34\nvalue = [4, 30]'),
 Text(0.04420645178939018, 0.6145833333333334, 'X[0] <= 23.0\ngini =
0.5\nsamples = 2\nvalue = [1, 1]'),
 Text(0.04376876414791107, 0.59375, 'gini = 0.0\nsamples = 1\nvalue = [0, 1]'),
 Text(0.04464413943086929, 0.59375, 'gini = 0.0\nsamples = 1\nvalue = [1, 0]'),
 Text(0.04595720235530662, 0.6145833333333334, 'X[2] <= 118073.5\ngini =
0.17\nsamples = 32\nvalue = [3, 29]'),
 Text(0.04551951471382751, 0.59375, 'X[2] <= 114842.0\ngini = 0.278\nsamples =
18\nvalue = [3, 15]'),
 Text(0.045081827072348396, 0.5729166666666666, 'X[2] <= 64728.5\ngini =
0.208\nsamples = 17\nvalue = [2, 15]'),
 Text(0.04464413943086929, 0.5520833333333334, 'gini = 0.0\nsamples = 8\nvalue =
[0, 8]'),
 Text(0.04551951471382751, 0.5520833333333334, 'X[2] <= 83258.5\ngini =
0.346\nsamples = 9\nvalue = [2, 7]'),
 Text(0.045081827072348396, 0.53125, 'X[0] <= 26.5\ngini = 0.444\nsamples =
3\nvalue = [2, 1]'),
 Text(0.04464413943086929, 0.5104166666666666, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.04551951471382751, 0.5104166666666666, 'gini = 0.0\nsamples = 2\nvalue =
[2, 0]'),
 Text(0.04595720235530662, 0.53125, 'gini = 0.0\nsamples = 6\nvalue = [0, 6]'),
 Text(0.04595720235530662, 0.5729166666666666, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.04639488999678573, 0.59375, 'gini = 0.0\nsamples = 14\nvalue = [0,
14]'),
 Text(0.04595720235530662, 0.6354166666666666, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.04639488999678573, 0.65625, 'gini = 0.0\nsamples = 41\nvalue = [0,
41]'),
 Text(0.046832577638264844, 0.6770833333333334, 'gini = 0.0\nsamples =
120\nvalue = [0, 120]'),
 Text(0.04814564056270217, 0.6979166666666666, 'X[10] <= 3257.0\ngini =
0.375\nsamples = 4\nvalue = [1, 3]'),
 Text(0.04770795292122306, 0.6770833333333334, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.048583328204181285, 0.6770833333333334, 'gini = 0.0\nsamples = 3\nvalue
= [0, 3]'),
 Text(0.05137358691861062, 0.71875, 'X[2] <= 32224.5\ngini = 0.187\nsamples =
67\nvalue = [7, 60]'),
 Text(0.050935899277131505, 0.6979166666666666, 'gini = 0.0\nsamples = 1\nvalue
= [1, 0]'),
 Text(0.05181127456008973, 0.6979166666666666, 'X[0] <= 26.5\ngini =
```

```
0.165\nsamples = 66\nvalue = [6, 60]'),
 Text(0.04945870348713951, 0.6770833333333334, 'X[6] <= 12.0\ngini =
0.05\nsamples = 39\nvalue = [1, 38]'),
 Text(0.0490210158456604, 0.65625, 'gini = 0.0\nsamples = 37\nvalue = [0, 37]'),
 Text(0.049896391128618614, 0.65625, 'X[2] <= 246405.0\ngini = 0.5\nsamples =
2\nvalue = [1, 1]'),
 Text(0.04945870348713951, 0.6354166666666666, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.050334078770097726, 0.6354166666666666, 'gini = 0.0\nsamples = 1\nvalue
= [1, 0]'),
 Text(0.05416384563303995, 0.6770833333333334, 'X[1] <= 4.5\ngini =
0.302\nsamples = 27\nvalue = [5, 22]'),
 Text(0.052741360798232835, 0.65625, 'X[12] <= 47.5\ngini = 0.269\nsamples =
25\nvalue = [4, 21]'),
 Text(0.05120945405305595, 0.6354166666666666, 'X[2] <= 234015.0\ngini =
0.397\nsamples = 11\nvalue = [3, 8]'),
 Text(0.050334078770097726, 0.6145833333333334, 'X[6] <= 11.0\ngini =
0.245\nsamples = 7\nvalue = [1, 6]'),
 Text(0.049896391128618614, 0.59375, 'gini = 0.0\nsamples = 6\nvalue = [0, 6]'),
 Text(0.05077176641157684, 0.59375, 'gini = 0.0\nsamples = 1\nvalue = [1, 0]'),
 Text(0.05208482933601417, 0.6145833333333334, 'X[2] <= 286828.0\ngini =
0.5\nsamples = 4\nvalue = [2, 2]'),
 Text(0.05164714169453506, 0.59375, 'gini = 0.0\nsamples = 2\nvalue = [2, 0]'),
 Text(0.05252251697749328, 0.59375, 'gini = 0.0\nsamples = 2\nvalue = [0, 2]'),
 Text(0.05427326754340972, 0.6354166666666666, 'X[0] <= 27.5\ngini =
0.133\nsamples = 14\nvalue = [1, 13]'),
 Text(0.053835579901930615, 0.6145833333333334, 'X[6] <= 4.0\ngini =
0.375\nsamples = 4\nvalue = [1, 3]'),
 Text(0.0533978922604515, 0.59375, 'gini = 0.0\nsamples = 2\nvalue = [0, 2]'),
 Text(0.05427326754340972, 0.59375, 'X[2] <= 248248.5\ngini = 0.5\nsamples =
2\nvalue = [1, 1]'),
 Text(0.053835579901930615, 0.5729166666666666, 'gini = 0.0\nsamples = 1\nvalue
= [1, 0]'),
 Text(0.05471095518488883, 0.5729166666666666, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.05471095518488883, 0.6145833333333334, 'gini = 0.0\nsamples = 10\nvalue
= [0, 10]'),
 Text(0.055586330467847056, 0.65625, 'X[6] <= 6.5\ngini = 0.5\nsamples =
2\nvalue = [1, 1]'),
 Text(0.055148642826367944, 0.6354166666666666, 'gini = 0.0\nsamples = 1\nvalue
= [0, 1]'),
 Text(0.05602401810932617, 0.6354166666666666, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.05356202512600617, 0.7395833333333334, 'X[6] <= 12.5\ngini =
0.5\nsamples = 4\nvalue = [2, 2]'),
 Text(0.05312433748452706, 0.71875, 'X[8] <= 2.5\ngini = 0.444\nsamples =
3\nvalue = [1, 2]'),
```

```
 Text(0.052686649843047946, 0.6979166666666666, 'gini = 0.0\nsamples = 1\nvalue
= [1, 0]'),
 Text(0.05356202512600617, 0.6979166666666666, 'gini = 0.0\nsamples = 2\nvalue =
[0, 2]'),
 Text(0.05399971276748528, 0.71875, 'gini = 0.0\nsamples = 1\nvalue = [1, 0]'),
 Text(0.052317350895549945, 0.7604166666666666, 'gini = 0.0\nsamples = 1\nvalue
= [1, 0]'),
 Text(0.05760518608563632, 0.8020833333333334, 'X[8] <= 3.5\ngini =
0.252\nsamples = 547\nvalue = [81, 466]'),
 Text(0.05531277569192261, 0.78125, 'X[2] <= 76158.5\ngini = 0.034\nsamples =
58\nvalue = [1, 57]'),
 Text(0.0548750880504435, 0.7604166666666666, 'X[2] <= 68167.0\ngini =
0.278\nsamples = 6\nvalue = [1, 5]'),
 Text(0.05443740040896439, 0.7395833333333334, 'gini = 0.0\nsamples = 5\nvalue =
[0, 5]'),
 Text(0.05531277569192261, 0.7395833333333334, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.05575046333340172, 0.7604166666666666, 'gini = 0.0\nsamples = 52\nvalue
= [0, 52]'),
 Text(0.059897596479350036, 0.78125, 'X[2] <= 88409.0\ngini = 0.274\nsamples =
489\nvalue = [80, 409]'),
 Text(0.05706352625783905, 0.7604166666666666, 'X[0] <= 35.5\ngini =
0.114\nsamples = 66\nvalue = [4, 62]'),
 Text(0.056188150974880835, 0.7395833333333334, 'X[6] <= 10.0\ngini =
0.062\nsamples = 62\nvalue = [2, 60]'),
 Text(0.05575046333340172, 0.71875, 'gini = 0.0\nsamples = 52\nvalue = [0,
52]'),
 Text(0.05662583861635994, 0.71875, 'X[12] <= 47.0\ngini = 0.32\nsamples =
10\nvalue = [2, 8]'),
 Text(0.056188150974880835, 0.6979166666666666, 'gini = 0.0\nsamples = 7\nvalue
= [0, 7]'),
 Text(0.05706352625783905, 0.6979166666666666, 'X[1] <= 4.0\ngini =
0.444\nsamples = 3\nvalue = [2, 1]'),
 Text(0.05662583861635994, 0.6770833333333334, 'gini = 0.0\nsamples = 2\nvalue =
[2, 0]'),
 Text(0.057501213899318164, 0.6770833333333334, 'gini = 0.0\nsamples = 1\nvalue
= [0, 1]'),
 Text(0.057938901540797276, 0.7395833333333334, 'X[4] <= 8.5\ngini =
0.5\nsamples = 4\nvalue = [2, 2]'),
 Text(0.057501213899318164, 0.71875, 'gini = 0.0\nsamples = 1\nvalue = [0, 1]'),
 Text(0.05837658918227639, 0.71875, 'X[2] <= 52075.5\ngini = 0.444\nsamples =
3\nvalue = [2, 1]'),
 Text(0.057938901540797276, 0.6979166666666666, 'gini = 0.0\nsamples = 2\nvalue
= [2, 0]'),
 Text(0.0588142768237555, 0.6979166666666666, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.06273166670086101, 0.7604166666666666, 'X[2] <= 99477.5\ngini =
```

```
0.295\nsamples = 423\nvalue = [76, 347]'),
 Text(0.06056502738967194, 0.7395833333333334, 'X[12] <= 42.5\ngini =
0.49\nsamples = 14\nvalue = [6, 8]'),
 Text(0.06012733974819283, 0.71875, 'X[2] <= 98834.5\ngini = 0.32\nsamples =
10\nvalue = [2, 8]'),
 Text(0.05968965210671372, 0.6979166666666666, 'X[6] <= 12.0\ngini =
0.198\nsamples = 9\nvalue = [1, 8]'),
 Text(0.059251964465234605, 0.6770833333333334, 'gini = 0.0\nsamples = 8\nvalue
= [0, 8]'),
 Text(0.06012733974819283, 0.6770833333333334, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.06056502738967194, 0.6979166666666666, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.06100271503115105, 0.71875, 'gini = 0.0\nsamples = 4\nvalue = [4, 0]'),
 Text(0.06489830601205009, 0.7395833333333334, 'X[1] <= 0.5\ngini =
0.284\nsamples = 409\nvalue = [70, 339]'),
 Text(0.06402293072909186, 0.71875, 'X[0] <= 31.5\ngini = 0.444\nsamples =
3\nvalue = [2, 1]'),
 Text(0.06358524308761275, 0.6979166666666666, 'gini = 0.0\nsamples = 2\nvalue =
[2, 0]'),
 Text(0.06446061837057097, 0.6979166666666666, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.06577368129500831, 0.71875, 'X[13] <= 40.0\ngini = 0.279\nsamples =
406\nvalue = [68, 338]'),
 Text(0.0653359936535292, 0.6979166666666666, 'X[2] <= 125520.5\ngini =
0.276\nsamples = 405\nvalue = [67, 338]'),
 Text(0.06100271503115105, 0.6770833333333334, 'X[1] <= 4.5\ngini =
0.142\nsamples = 65\nvalue = [5, 60]'),
 Text(0.06012733974819283, 0.65625, 'X[0] <= 31.5\ngini = 0.07\nsamples =
55\nvalue = [2, 53]'),
 Text(0.05968965210671372, 0.6354166666666666, 'X[2] <= 113601.0\ngini =
0.18\nsamples = 20\nvalue = [2, 18]'),
 Text(0.059251964465234605, 0.6145833333333334, 'X[2] <= 112873.5\ngini =
0.26\nsamples = 13\nvalue = [2, 11]'),
 Text(0.0588142768237555, 0.59375, 'X[0] <= 29.5\ngini = 0.153\nsamples =
12\nvalue = [1, 11]'),
 Text(0.05837658918227639, 0.5729166666666666, 'X[2] <= 105558.0\ngini =
0.32\nsamples = 5\nvalue = [1, 4]'),
 Text(0.057938901540797276, 0.5520833333333334, 'gini = 0.0\nsamples = 1\nvalue
= [1, 0]'),
 Text(0.0588142768237555, 0.5520833333333334, 'gini = 0.0\nsamples = 4\nvalue =
[0, 4]'),
 Text(0.059251964465234605, 0.5729166666666666, 'gini = 0.0\nsamples = 7\nvalue
= [0, 7]'),
 Text(0.05968965210671372, 0.59375, 'gini = 0.0\nsamples = 1\nvalue = [1, 0]'),
 Text(0.06012733974819283, 0.6145833333333334, 'gini = 0.0\nsamples = 7\nvalue =
[0, 7]'),
```

```
 Text(0.06056502738967194, 0.6354166666666666, 'gini = 0.0\nsamples = 35\nvalue
= [0, 35]'),
 Text(0.06187809031410927, 0.65625, 'X[0] <= 31.5\ngini = 0.42\nsamples =
10\nvalue = [3, 7]'),
 Text(0.06144040267263016, 0.6354166666666666, 'gini = 0.0\nsamples = 4\nvalue =
[0, 4]'),
 Text(0.06231577795558838, 0.6354166666666666, 'X[12] <= 49.5\ngini =
0.5\nsamples = 6\nvalue = [3, 3]'),
 Text(0.06187809031410927, 0.6145833333333334, 'X[0] <= 32.5\ngini =
0.375\nsamples = 4\nvalue = [3, 1]'),
 Text(0.06144040267263016, 0.59375, 'gini = 0.0\nsamples = 2\nvalue = [2, 0]'),
 Text(0.06231577795558838, 0.59375, 'X[2] <= 112488.5\ngini = 0.5\nsamples =
2\nvalue = [1, 1]'),
 Text(0.06187809031410927, 0.5729166666666666, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.06275346559706749, 0.5729166666666666, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.06275346559706749, 0.6145833333333334, 'gini = 0.0\nsamples = 2\nvalue =
[0, 2]'),
 Text(0.06966927227590734, 0.6770833333333334, 'X[2] <= 126622.0\ngini =
0.298\nsamples = 340\nvalue = [62, 278]'),
 Text(0.06613870594913249, 0.65625, 'X[6] <= 5.5\ngini = 0.444\nsamples =
3\nvalue = [2, 1]'),
 Text(0.06570101830765337, 0.6354166666666666, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.0665763935906116, 0.6354166666666666, 'gini = 0.0\nsamples = 2\nvalue =
[2, 0]'),
 Text(0.0731998386026822, 0.65625, 'X[2] <= 304735.0\ngini = 0.293\nsamples =
337\nvalue = [60, 277]'),
 Text(0.06745176887356982, 0.6354166666666666, 'X[2] <= 155409.0\ngini =
0.271\nsamples = 279\nvalue = [45, 234]'),
 Text(0.06450421616298394, 0.6145833333333334, 'X[2] <= 154157.0\ngini =
0.37\nsamples = 49\nvalue = [12, 37]'),
 Text(0.06406652852150482, 0.59375, 'X[1] <= 2.0\ngini = 0.335\nsamples =
47\nvalue = [10, 37]'),
 Text(0.06362884088002571, 0.5729166666666666, 'gini = 0.0\nsamples = 2\nvalue =
[2, 0]'),
 Text(0.06450421616298394, 0.5729166666666666, 'X[12] <= 49.0\ngini =
0.292\nsamples = 45\nvalue = [8, 37]'),
 Text(0.06297230941780704, 0.5520833333333334, 'X[0] <= 31.5\ngini =
0.245\nsamples = 42\nvalue = [6, 36]'),
 Text(0.0612215588518906, 0.53125, 'X[6] <= 12.0\ngini = 0.408\nsamples =
14\nvalue = [4, 10]'),
 Text(0.060346183568932385, 0.5104166666666666, 'X[2] <= 151229.0\ngini =
0.298\nsamples = 11\nvalue = [2, 9]'),
 Text(0.05990849592745327, 0.4895833333333333, 'X[12] <= 46.5\ngini =
0.18\nsamples = 10\nvalue = [1, 9]'),
```

```
 Text(0.05947080828597416, 0.46875, 'gini = 0.0\nsamples = 9\nvalue = [0, 9]'),
 Text(0.060346183568932385, 0.46875, 'gini = 0.0\nsamples = 1\nvalue = [1, 0]'),
 Text(0.0607838712104115, 0.4895833333333333, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.062096934134848826, 0.5104166666666666, 'X[2] <= 137341.0\ngini =
0.444\nsamples = 3\nvalue = [2, 1]'),
 Text(0.061659246493369714, 0.4895833333333333, 'gini = 0.0\nsamples = 2\nvalue
= [2, 0]'),
 Text(0.06253462177632793, 0.4895833333333333, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.06472305998372349, 0.53125, 'X[6] <= 10.5\ngini = 0.133\nsamples =
28\nvalue = [2, 26]'),
 Text(0.06384768470076527, 0.5104166666666666, 'X[1] <= 3.5\ngini =
0.08\nsamples = 24\nvalue = [1, 23]'),
 Text(0.06340999705928616, 0.4895833333333333, 'gini = 0.0\nsamples = 19\nvalue
= [0, 19]'),
 Text(0.06428537234224438, 0.4895833333333333, 'X[1] <= 4.5\ngini =
0.32\nsamples = 5\nvalue = [1, 4]'),
 Text(0.06384768470076527, 0.46875, 'gini = 0.0\nsamples = 1\nvalue = [1, 0]'),
 Text(0.06472305998372349, 0.46875, 'gini = 0.0\nsamples = 4\nvalue = [0, 4]'),
 Text(0.06559843526668171, 0.5104166666666666, 'X[6] <= 12.0\ngini =
0.375\nsamples = 4\nvalue = [1, 3]'),
 Text(0.0651607476252026, 0.4895833333333333, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.06603612290816083, 0.4895833333333333, 'gini = 0.0\nsamples = 3\nvalue =
[0, 3]'),
 Text(0.06603612290816083, 0.5520833333333334, 'X[2] <= 153762.5\ngini =
0.444\nsamples = 3\nvalue = [2, 1]'),
 Text(0.06559843526668171, 0.53125, 'gini = 0.0\nsamples = 2\nvalue = [2, 0]'),
 Text(0.06647381054963994, 0.53125, 'gini = 0.0\nsamples = 1\nvalue = [0, 1]'),
 Text(0.06494190380446305, 0.59375, 'gini = 0.0\nsamples = 2\nvalue = [2, 0]'),
 Text(0.07039932158415571, 0.6145833333333334, 'X[2] <= 169781.0\ngini =
0.246\nsamples = 230\nvalue = [33, 197]'),
 Text(0.06915464735369949, 0.59375, 'X[6] <= 1.0\ngini = 0.054\nsamples =
36\nvalue = [1, 35]'),
 Text(0.06871695971222037, 0.5729166666666666, 'X[12] <= 42.5\ngini =
0.444\nsamples = 3\nvalue = [1, 2]'),
 Text(0.06827927207074126, 0.5520833333333334, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.06915464735369949, 0.5520833333333334, 'X[1] <= 3.5\ngini = 0.5\nsamples
= 2\nvalue = [1, 1]'),
 Text(0.06871695971222037, 0.53125, 'gini = 0.0\nsamples = 1\nvalue = [1, 0]'),
 Text(0.0695923349951786, 0.53125, 'gini = 0.0\nsamples = 1\nvalue = [0, 1]'),
 Text(0.0695923349951786, 0.5729166666666666, 'gini = 0.0\nsamples = 33\nvalue =
[0, 33]'),
 Text(0.07164399581461192, 0.59375, 'X[2] <= 171209.0\ngini = 0.275\nsamples =
194\nvalue = [32, 162]'),
```

```
 Text(0.07076862053165371, 0.5729166666666666, 'X[12] <= 42.5\ngini =
0.48\nsamples = 5\nvalue = [3, 2]'),
 Text(0.0703309328901746, 0.5520833333333334, 'gini = 0.0\nsamples = 2\nvalue =
[0, 2]'),
 Text(0.07120630817313282, 0.5520833333333334, 'gini = 0.0\nsamples = 3\nvalue =
[3, 0]'),
 Text(0.07251937109757015, 0.5729166666666666, 'X[1] <= 2.0\ngini =
0.26\nsamples = 189\nvalue = [29, 160]'),
 Text(0.07208168345609103, 0.5520833333333334, 'gini = 0.0\nsamples = 11\nvalue
= [0, 11]'),
 Text(0.07295705873904926, 0.5520833333333334, 'X[2] <= 266736.5\ngini =
0.273\nsamples = 178\nvalue = [29, 149]'),
 Text(0.07046771027813682, 0.53125, 'X[2] <= 263714.0\ngini = 0.295\nsamples =
150\nvalue = [27, 123]'),
 Text(0.07003002263665771, 0.5104166666666666, 'X[2] <= 208434.0\ngini =
0.288\nsamples = 149\nvalue = [26, 123]'),
 Text(0.0671303420118586, 0.4895833333333333, 'X[0] <= 31.5\ngini =
0.225\nsamples = 85\nvalue = [11, 74]'),
 Text(0.06559843526668171, 0.46875, 'X[0] <= 29.5\ngini = 0.334\nsamples =
33\nvalue = [7, 26]'),
 Text(0.06472305998372349, 0.4479166666666667, 'X[2] <= 177828.5\ngini =
0.133\nsamples = 14\nvalue = [1, 13]'),
 Text(0.06428537234224438, 0.4270833333333333, 'X[12] <= 45.0\ngini =
0.5\nsamples = 2\nvalue = [1, 1]'),
 Text(0.06384768470076527, 0.40625, 'gini = 0.0\nsamples = 1\nvalue = [1, 0]'),
 Text(0.06472305998372349, 0.40625, 'gini = 0.0\nsamples = 1\nvalue = [0, 1]'),
 Text(0.0651607476252026, 0.4270833333333333, 'gini = 0.0\nsamples = 12\nvalue =
[0, 12]'),
 Text(0.06647381054963994, 0.4479166666666667, 'X[6] <= 12.5\ngini =
0.432\nsamples = 19\nvalue = [6, 13]'),
 Text(0.06603612290816083, 0.4270833333333333, 'X[6] <= 11.5\ngini =
0.48\nsamples = 15\nvalue = [6, 9]'),
 Text(0.06559843526668171, 0.40625, 'X[1] <= 4.0\ngini = 0.459\nsamples =
14\nvalue = [5, 9]'),
 Text(0.0651607476252026, 0.3854166666666667, 'X[2] <= 205051.5\ngini =
0.426\nsamples = 13\nvalue = [4, 9]'),
 Text(0.06472305998372349, 0.3645833333333333, 'X[12] <= 47.5\ngini =
0.375\nsamples = 12\nvalue = [3, 9]'),
 Text(0.06384768470076527, 0.34375, 'X[6] <= 5.5\ngini = 0.198\nsamples =
9\nvalue = [1, 8]'),
 Text(0.06340999705928616, 0.3229166666666667, 'gini = 0.0\nsamples = 6\nvalue =
[0, 6]'),
 Text(0.06428537234224438, 0.3229166666666667, 'X[6] <= 6.5\ngini =
0.444\nsamples = 3\nvalue = [1, 2]'),
 Text(0.06384768470076527, 0.3020833333333333, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.06472305998372349, 0.3020833333333333, 'gini = 0.0\nsamples = 2\nvalue =
```

```
[0, 2]'),
 Text(0.06559843526668171, 0.34375, 'X[0] <= 30.5\ngini = 0.444\nsamples =
3\nvalue = [2, 1]'),
 Text(0.0651607476252026, 0.3229166666666667, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.06603612290816083, 0.3229166666666667, 'gini = 0.0\nsamples = 2\nvalue =
[2, 0]'),
 Text(0.06559843526668171, 0.3645833333333333, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.06603612290816083, 0.3854166666666667, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.06647381054963994, 0.40625, 'gini = 0.0\nsamples = 1\nvalue = [1, 0]'),
 Text(0.06691149819111905, 0.4270833333333333, 'gini = 0.0\nsamples = 4\nvalue =
[0, 4]'),
 Text(0.06866224875703549, 0.46875, 'X[2] <= 190592.5\ngini = 0.142\nsamples =
52\nvalue = [4, 48]'),
 Text(0.06822456111555637, 0.4479166666666667, 'X[2] <= 183445.5\ngini =
0.26\nsamples = 26\nvalue = [4, 22]'),
 Text(0.06778687347407726, 0.4270833333333333, 'gini = 0.0\nsamples = 12\nvalue
= [0, 12]'),
 Text(0.06866224875703549, 0.4270833333333333, 'X[2] <= 188122.5\ngini =
0.408\nsamples = 14\nvalue = [4, 10]'),
 Text(0.06778687347407726, 0.40625, 'X[2] <= 185332.0\ngini = 0.5\nsamples =
6\nvalue = [3, 3]'),
 Text(0.06734918583259815, 0.3854166666666667, 'X[2] <= 183695.0\ngini =
0.375\nsamples = 4\nvalue = [1, 3]'),
 Text(0.06691149819111905, 0.3645833333333333, 'X[6] <= 9.5\ngini = 0.5\nsamples
= 2\nvalue = [1, 1]'),
 Text(0.06647381054963994, 0.34375, 'gini = 0.0\nsamples = 1\nvalue = [1, 0]'),
 Text(0.06734918583259815, 0.34375, 'gini = 0.0\nsamples = 1\nvalue = [0, 1]'),
 Text(0.06778687347407726, 0.3645833333333333, 'gini = 0.0\nsamples = 2\nvalue =
[0, 2]'),
 Text(0.06822456111555637, 0.3854166666666667, 'gini = 0.0\nsamples = 2\nvalue =
[2, 0]'),
 Text(0.06953762403999371, 0.40625, 'X[2] <= 190066.5\ngini = 0.219\nsamples =
8\nvalue = [1, 7]'),
 Text(0.0690999363985146, 0.3854166666666667, 'gini = 0.0\nsamples = 6\nvalue =
[0, 6]'),
 Text(0.06997531168147282, 0.3854166666666667, 'X[0] <= 33.0\ngini =
0.5\nsamples = 2\nvalue = [1, 1]'),
 Text(0.06953762403999371, 0.3645833333333333, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.07041299932295193, 0.3645833333333333, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.0690999363985146, 0.4479166666666667, 'gini = 0.0\nsamples = 26\nvalue =
[0, 26]'),
 Text(0.07292970326145681, 0.4895833333333333, 'X[2] <= 209210.0\ngini =
```

```
0.359\nsamples = 64\nvalue = [15, 49]'),
 Text(0.0724920156199777, 0.46875, 'gini = 0.0\nsamples = 2\nvalue = [2, 0]'),
 Text(0.07336739090293593, 0.46875, 'X[0] <= 30.5\ngini = 0.331\nsamples =
62\nvalue = [13, 49]'),
 Text(0.07128837460591016, 0.4479166666666667, 'X[1] <= 5.5\ngini =
0.188\nsamples = 19\nvalue = [2, 17]'),
 Text(0.07085068696443104, 0.4270833333333333, 'X[2] <= 240215.0\ngini =
0.105\nsamples = 18\nvalue = [1, 17]'),
 Text(0.07041299932295193, 0.40625, 'gini = 0.0\nsamples = 14\nvalue = [0,
14]'),
 Text(0.07128837460591016, 0.40625, 'X[2] <= 243662.5\ngini = 0.375\nsamples =
4\nvalue = [1, 3]'),
 Text(0.07085068696443104, 0.3854166666666667, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.07172606224738926, 0.3854166666666667, 'gini = 0.0\nsamples = 3\nvalue =
[0, 3]'),
 Text(0.07172606224738926, 0.4270833333333333, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.07544640719996171, 0.4479166666666667, 'X[0] <= 32.5\ngini =
0.381\nsamples = 43\nvalue = [11, 32]'),
 Text(0.0734768128133057, 0.4270833333333333, 'X[2] <= 249902.0\ngini =
0.486\nsamples = 12\nvalue = [5, 7]'),
 Text(0.07303912517182659, 0.40625, 'X[6] <= 4.0\ngini = 0.42\nsamples =
10\nvalue = [3, 7]'),
 Text(0.07260143753034748, 0.3854166666666667, 'gini = 0.0\nsamples = 4\nvalue =
[0, 4]'),
 Text(0.0734768128133057, 0.3854166666666667, 'X[6] <= 10.0\ngini = 0.5\nsamples
= 6\nvalue = [3, 3]'),
 Text(0.07303912517182659, 0.3645833333333333, 'X[2] <= 215035.5\ngini =
0.48\nsamples = 5\nvalue = [2, 3]'),
 Text(0.07260143753034748, 0.34375, 'gini = 0.0\nsamples = 1\nvalue = [0, 1]'),
 Text(0.0734768128133057, 0.34375, 'X[2] <= 221360.5\ngini = 0.5\nsamples =
4\nvalue = [2, 2]'),
 Text(0.07303912517182659, 0.3229166666666667, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.07391450045478482, 0.3229166666666667, 'X[0] <= 31.5\ngini =
0.444\nsamples = 3\nvalue = [1, 2]'),
 Text(0.0734768128133057, 0.3020833333333333, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.07435218809626393, 0.3020833333333333, 'X[2] <= 235178.5\ngini =
0.5\nsamples = 2\nvalue = [1, 1]'),
 Text(0.07391450045478482, 0.28125, 'gini = 0.0\nsamples = 1\nvalue = [1, 0]'),
 Text(0.07478987573774304, 0.28125, 'gini = 0.0\nsamples = 1\nvalue = [0, 1]'),
 Text(0.07391450045478482, 0.3645833333333333, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.07391450045478482, 0.40625, 'gini = 0.0\nsamples = 2\nvalue = [2, 0]'),
 Text(0.0774160015866177, 0.4270833333333333, 'X[6] <= 2.5\ngini =
```

```
0.312\nsamples = 31\nvalue = [6, 25]'),
 Text(0.07654062630365947, 0.40625, 'X[2] <= 243544.0\ngini = 0.459\nsamples =
14\nvalue = [5, 9]'),
 Text(0.07610293866218037, 0.3854166666666667, 'X[1] <= 4.0\ngini = 0.5\nsamples
= 10\nvalue = [5, 5]'),
 Text(0.07566525102070126, 0.3645833333333333, 'X[6] <= 1.0\ngini =
0.469\nsamples = 8\nvalue = [5, 3]'),
 Text(0.07522756337922215, 0.34375, 'gini = 0.0\nsamples = 2\nvalue = [0, 2]'),
 Text(0.07610293866218037, 0.34375, 'X[2] <= 216428.0\ngini = 0.278\nsamples =
6\nvalue = [5, 1]'),
 Text(0.07566525102070126, 0.3229166666666667, 'X[2] <= 214305.0\ngini =
0.5\nsamples = 2\nvalue = [1, 1]'),
 Text(0.07522756337922215, 0.3020833333333333, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.07610293866218037, 0.3020833333333333, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.07654062630365947, 0.3229166666666667, 'gini = 0.0\nsamples = 4\nvalue =
[4, 0]'),
 Text(0.07654062630365947, 0.3645833333333333, 'gini = 0.0\nsamples = 2\nvalue =
[0, 2]'),
 Text(0.07697831394513859, 0.3854166666666667, 'gini = 0.0\nsamples = 4\nvalue =
[0, 4]'),
 Text(0.07829137686957592, 0.40625, 'X[2] <= 232549.0\ngini = 0.111\nsamples =
17\nvalue = [1, 16]'),
 Text(0.07785368922809681, 0.3854166666666667, 'X[2] <= 225218.5\ngini =
0.245\nsamples = 7\nvalue = [1, 6]'),
 Text(0.0774160015866177, 0.3645833333333333, 'gini = 0.0\nsamples = 6\nvalue =
[0, 6]'),
 Text(0.07829137686957592, 0.3645833333333333, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.07872906451105503, 0.3854166666666667, 'gini = 0.0\nsamples = 10\nvalue
= [0, 10]'),
 Text(0.07090539791961593, 0.5104166666666666, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.07544640719996171, 0.53125, 'X[2] <= 286067.5\ngini = 0.133\nsamples =
28\nvalue = [2, 26]'),
 Text(0.0750087195584826, 0.5104166666666666, 'gini = 0.0\nsamples = 15\nvalue =
[0, 15]'),
 Text(0.07588409484144082, 0.5104166666666666, 'X[2] <= 286395.0\ngini =
0.26\nsamples = 13\nvalue = [2, 11]'),
 Text(0.07544640719996171, 0.4895833333333333, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.07632178248291992, 0.4895833333333333, 'X[6] <= 9.0\ngini =
0.153\nsamples = 12\nvalue = [1, 11]'),
 Text(0.07588409484144082, 0.46875, 'gini = 0.0\nsamples = 9\nvalue = [0, 9]'),
 Text(0.07675947012439903, 0.46875, 'X[6] <= 12.0\ngini = 0.444\nsamples =
3\nvalue = [1, 2]'),
```

```
 Text(0.07632178248291992, 0.4479166666666667, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.07719715776587814, 0.4479166666666667, 'gini = 0.0\nsamples = 2\nvalue =
[0, 2]'),
 Text(0.07894790833179459, 0.6354166666666666, 'X[1] <= 2.0\ngini =
0.383\nsamples = 58\nvalue = [15, 43]'),
 Text(0.07807253304883637, 0.6145833333333334, 'X[2] <= 327931.5\ngini =
0.375\nsamples = 4\nvalue = [3, 1]'),
 Text(0.07763484540735725, 0.59375, 'gini = 0.0\nsamples = 3\nvalue = [3, 0]'),
 Text(0.07851022069031548, 0.59375, 'gini = 0.0\nsamples = 1\nvalue = [0, 1]'),
 Text(0.07982328361475281, 0.6145833333333334, 'X[2] <= 306148.5\ngini =
0.346\nsamples = 54\nvalue = [12, 42]'),
 Text(0.0793855959732737, 0.59375, 'gini = 0.0\nsamples = 1\nvalue = [1, 0]'),
 Text(0.08026097125623192, 0.59375, 'X[6] <= 10.5\ngini = 0.329\nsamples =
53\nvalue = [11, 42]'),
 Text(0.07894790833179459, 0.5729166666666666, 'X[2] <= 344920.5\ngini =
0.273\nsamples = 43\nvalue = [7, 36]'),
 Text(0.07763484540735725, 0.5520833333333334, 'X[2] <= 325570.0\ngini =
0.391\nsamples = 15\nvalue = [4, 11]'),
 Text(0.07719715776587814, 0.53125, 'gini = 0.0\nsamples = 7\nvalue = [0, 7]'),
 Text(0.07807253304883637, 0.53125, 'X[0] <= 31.0\ngini = 0.5\nsamples =
8\nvalue = [4, 4]'),
 Text(0.07763484540735725, 0.5104166666666666, 'gini = 0.0\nsamples = 2\nvalue =
[0, 2]'),
 Text(0.07851022069031548, 0.5104166666666666, 'X[0] <= 34.5\ngini =
0.444\nsamples = 6\nvalue = [4, 2]'),
 Text(0.07807253304883637, 0.4895833333333333, 'X[2] <= 334220.0\ngini =
0.5\nsamples = 4\nvalue = [2, 2]'),
 Text(0.07763484540735725, 0.46875, 'gini = 0.0\nsamples = 1\nvalue = [1, 0]'),
 Text(0.07851022069031548, 0.46875, 'X[6] <= 4.5\ngini = 0.444\nsamples =
3\nvalue = [1, 2]'),
 Text(0.07807253304883637, 0.4479166666666667, 'gini = 0.0\nsamples = 2\nvalue =
[0, 2]'),
 Text(0.07894790833179459, 0.4479166666666667, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.07894790833179459, 0.4895833333333333, 'gini = 0.0\nsamples = 2\nvalue =
[2, 0]'),
 Text(0.08026097125623192, 0.5520833333333334, 'X[2] <= 380475.5\ngini =
0.191\nsamples = 28\nvalue = [3, 25]'),
 Text(0.07982328361475281, 0.53125, 'gini = 0.0\nsamples = 11\nvalue = [0,
11]'),
 Text(0.08069865889771104, 0.53125, 'X[2] <= 389310.0\ngini = 0.291\nsamples =
17\nvalue = [3, 14]'),
 Text(0.08026097125623192, 0.5104166666666666, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.08113634653919014, 0.5104166666666666, 'X[12] <= 47.5\ngini =
0.219\nsamples = 16\nvalue = [2, 14]'),
```

```
 Text(0.08069865889771104, 0.4895833333333333, 'X[0] <= 30.5\ngini =
0.124\nsamples = 15\nvalue = [1, 14]'),
 Text(0.08026097125623192, 0.46875, 'X[2] <= 514449.0\ngini = 0.5\nsamples =
2\nvalue = [1, 1]'),
 Text(0.07982328361475281, 0.4479166666666667, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.08069865889771104, 0.4479166666666667, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.08113634653919014, 0.46875, 'gini = 0.0\nsamples = 13\nvalue = [0,
13]'),
 Text(0.08157403418066925, 0.4895833333333333, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.08157403418066925, 0.5729166666666666, 'X[6] <= 11.5\ngini =
0.48\nsamples = 10\nvalue = [4, 6]'),
 Text(0.08113634653919014, 0.5520833333333334, 'gini = 0.0\nsamples = 4\nvalue =
[4, 0]'),
 Text(0.08201172182214836, 0.5520833333333334, 'gini = 0.0\nsamples = 6\nvalue =
[0, 6]'),
 Text(0.06621136893648742, 0.6979166666666666, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.0897259665032177, 0.8229166666666666, 'X[1] <= 2.0\ngini = 0.41\nsamples
= 132\nvalue = [38, 94]'),
 Text(0.08928827886173858, 0.8020833333333334, 'gini = 0.0\nsamples = 4\nvalue =
[4, 0]'),
 Text(0.09016365414469679, 0.8020833333333334, 'X[0] <= 31.5\ngini =
0.39\nsamples = 128\nvalue = [34, 94]'),
 Text(0.08616975441619991, 0.78125, 'X[13] <= 4.0\ngini = 0.293\nsamples =
73\nvalue = [13, 60]'),
 Text(0.0857320667747208, 0.7604166666666666, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.08660744205767902, 0.7604166666666666, 'X[2] <= 34977.5\ngini =
0.278\nsamples = 72\nvalue = [12, 60]'),
 Text(0.08616975441619991, 0.7395833333333334, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.08704512969915813, 0.7395833333333334, 'X[2] <= 420719.5\ngini =
0.262\nsamples = 71\nvalue = [11, 60]'),
 Text(0.08660744205767902, 0.71875, 'X[2] <= 213707.0\ngini = 0.245\nsamples =
70\nvalue = [10, 60]'),
 Text(0.08463784767102303, 0.6979166666666666, 'X[2] <= 198382.0\ngini =
0.315\nsamples = 46\nvalue = [9, 37]'),
 Text(0.0833247847465857, 0.6770833333333334, 'X[2] <= 150135.0\ngini =
0.188\nsamples = 38\nvalue = [4, 34]'),
 Text(0.08288709710510658, 0.65625, 'X[2] <= 141327.0\ngini = 0.269\nsamples =
25\nvalue = [4, 21]'),
 Text(0.08244940946362747, 0.6354166666666666, 'X[12] <= 53.0\ngini =
0.219\nsamples = 24\nvalue = [3, 21]'),
 Text(0.08157403418066925, 0.6145833333333334, 'X[6] <= 6.5\ngini = 0.5\nsamples
```

```
= 2\nvalue = [1, 1]'),
 Text(0.08113634653919014, 0.59375, 'gini = 0.0\nsamples = 1\nvalue = [0, 1]'),
 Text(0.08201172182214836, 0.59375, 'gini = 0.0\nsamples = 1\nvalue = [1, 0]'),
 Text(0.0833247847465857, 0.6145833333333334, 'X[2] <= 111765.0\ngini =
0.165\nsamples = 22\nvalue = [2, 20]'),
 Text(0.08288709710510658, 0.59375, 'gini = 0.0\nsamples = 15\nvalue = [0,
15]'),
 Text(0.08376247238806481, 0.59375, 'X[2] <= 117272.0\ngini = 0.408\nsamples =
7\nvalue = [2, 5]'),
 Text(0.0833247847465857, 0.5729166666666666, 'gini = 0.0\nsamples = 2\nvalue =
[2, 0]'),
 Text(0.08420016002954392, 0.5729166666666666, 'gini = 0.0\nsamples = 5\nvalue =
[0, 5]'),
 Text(0.0833247847465857, 0.6354166666666666, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.08376247238806481, 0.65625, 'gini = 0.0\nsamples = 13\nvalue = [0,
13]'),
 Text(0.08595091059546035, 0.6770833333333334, 'X[12] <= 61.0\ngini =
0.469\nsamples = 8\nvalue = [5, 3]'),
 Text(0.08507553531250214, 0.65625, 'X[0] <= 25.5\ngini = 0.32\nsamples =
5\nvalue = [4, 1]'),
 Text(0.08463784767102303, 0.6354166666666666, 'X[13] <= 32.5\ngini =
0.5\nsamples = 2\nvalue = [1, 1]'),
 Text(0.08420016002954392, 0.6145833333333334, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.08507553531250214, 0.6145833333333334, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.08551322295398124, 0.6354166666666666, 'gini = 0.0\nsamples = 3\nvalue =
[3, 0]'),
 Text(0.08682628587841858, 0.65625, 'X[2] <= 205349.5\ngini = 0.444\nsamples =
3\nvalue = [1, 2]'),
 Text(0.08638859823693947, 0.6354166666666666, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.08726397351989769, 0.6354166666666666, 'gini = 0.0\nsamples = 2\nvalue =
[0, 2]'),
 Text(0.08857703644433502, 0.6979166666666666, 'X[8] <= 2.5\ngini =
0.08\nsamples = 24\nvalue = [1, 23]'),
 Text(0.08813934880285591, 0.6770833333333334, 'X[6] <= 4.0\ngini = 0.5\nsamples
= 2\nvalue = [1, 1]'),
 Text(0.0877016611613768, 0.65625, 'gini = 0.0\nsamples = 1\nvalue = [0, 1]'),
 Text(0.08857703644433502, 0.65625, 'gini = 0.0\nsamples = 1\nvalue = [1, 0]'),
 Text(0.08901472408581414, 0.6770833333333334, 'gini = 0.0\nsamples = 22\nvalue
= [0, 22]'),
 Text(0.08748281734063725, 0.71875, 'gini = 0.0\nsamples = 1\nvalue = [1, 0]'),
 Text(0.09415755387319369, 0.78125, 'X[13] <= 37.5\ngini = 0.472\nsamples =
55\nvalue = [21, 34]'),
 Text(0.09371986623171458, 0.7604166666666666, 'gini = 0.0\nsamples = 3\nvalue =
```

```
[3, 0]'),
 Text(0.0945952415146728, 0.7604166666666666, 'X[2] <= 232152.0\ngini =
0.453\nsamples = 52\nvalue = [18, 34]'),
 Text(0.09317275667986569, 0.7395833333333334, 'X[12] <= 65.5\ngini =
0.485\nsamples = 41\nvalue = [17, 24]'),
 Text(0.0916408499346888, 0.71875, 'X[6] <= 12.5\ngini = 0.499\nsamples =
31\nvalue = [16, 15]'),
 Text(0.09076547465173057, 0.6979166666666666, 'X[10] <= 1954.0\ngini =
0.472\nsamples = 21\nvalue = [13, 8]'),
 Text(0.09032778701025146, 0.6770833333333334, 'X[11] <= 836.0\ngini =
0.455\nsamples = 20\nvalue = [13, 7]'),
 Text(0.08989009936877236, 0.65625, 'X[2] <= 168585.5\ngini = 0.432\nsamples =
19\nvalue = [13, 6]'),
 Text(0.08813934880285591, 0.6354166666666666, 'X[1] <= 4.5\ngini =
0.219\nsamples = 8\nvalue = [7, 1]'),
 Text(0.0877016611613768, 0.6145833333333334, 'gini = 0.0\nsamples = 6\nvalue =
[6, 0]'),
 Text(0.08857703644433502, 0.6145833333333334, 'X[12] <= 57.5\ngini =
0.5\nsamples = 2\nvalue = [1, 1]'),
 Text(0.08813934880285591, 0.59375, 'gini = 0.0\nsamples = 1\nvalue = [0, 1]'),
 Text(0.08901472408581414, 0.59375, 'gini = 0.0\nsamples = 1\nvalue = [1, 0]'),
 Text(0.0916408499346888, 0.6354166666666666, 'X[2] <= 190854.0\ngini =
0.496\nsamples = 11\nvalue = [6, 5]'),
 Text(0.09032778701025146, 0.6145833333333334, 'X[1] <= 3.5\ngini =
0.375\nsamples = 4\nvalue = [1, 3]'),
 Text(0.08989009936877236, 0.59375, 'gini = 0.0\nsamples = 2\nvalue = [0, 2]'),
 Text(0.09076547465173057, 0.59375, 'X[12] <= 63.5\ngini = 0.5\nsamples =
2\nvalue = [1, 1]'),
 Text(0.09032778701025146, 0.5729166666666666, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.09120316229320968, 0.5729166666666666, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.09295391285912613, 0.6145833333333334, 'X[6] <= 2.5\ngini =
0.408\nsamples = 7\nvalue = [5, 2]'),
 Text(0.09251622521764702, 0.59375, 'X[2] <= 202969.0\ngini = 0.5\nsamples =
4\nvalue = [2, 2]'),
 Text(0.09207853757616791, 0.5729166666666666, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.09295391285912613, 0.5729166666666666, 'X[0] <= 35.0\ngini =
0.444\nsamples = 3\nvalue = [1, 2]'),
 Text(0.09251622521764702, 0.5520833333333334, 'gini = 0.0\nsamples = 2\nvalue =
[0, 2]'),
 Text(0.09339160050060524, 0.5520833333333334, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.09339160050060524, 0.59375, 'gini = 0.0\nsamples = 3\nvalue = [3, 0]'),
 Text(0.09076547465173057, 0.65625, 'gini = 0.0\nsamples = 1\nvalue = [0, 1]'),
 Text(0.09120316229320968, 0.6770833333333334, 'gini = 0.0\nsamples = 1\nvalue =
```

```
[0, 1]'),
 Text(0.09251622521764702, 0.6979166666666666, 'X[0] <= 32.5\ngini =
0.42\nsamples = 10\nvalue = [3, 7]'),
 Text(0.09207853757616791, 0.6770833333333334, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.09295391285912613, 0.6770833333333334, 'X[2] <= 144761.0\ngini =
0.346\nsamples = 9\nvalue = [2, 7]'),
 Text(0.09251622521764702, 0.65625, 'gini = 0.0\nsamples = 5\nvalue = [0, 5]'),
 Text(0.09339160050060524, 0.65625, 'X[2] <= 194781.0\ngini = 0.5\nsamples =
4\nvalue = [2, 2]'),
 Text(0.09295391285912613, 0.6354166666666666, 'gini = 0.0\nsamples = 2\nvalue =
[2, 0]'),
 Text(0.09382928814208435, 0.6354166666666666, 'gini = 0.0\nsamples = 2\nvalue =
[0, 2]'),
 Text(0.09470466342504257, 0.71875, 'X[0] <= 33.5\ngini = 0.18\nsamples =
10\nvalue = [1, 9]'),
 Text(0.09426697578356347, 0.6979166666666666, 'X[2] <= 127429.5\ngini =
0.5\nsamples = 2\nvalue = [1, 1]'),
 Text(0.09382928814208435, 0.6770833333333334, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.09470466342504257, 0.6770833333333334, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.09514235106652168, 0.6979166666666666, 'gini = 0.0\nsamples = 8\nvalue =
[0, 8]'),
 Text(0.0960177263494799, 0.7395833333333334, 'X[2] <= 426548.0\ngini =
0.165\nsamples = 11\nvalue = [1, 10]'),
 Text(0.09558003870800079, 0.71875, 'gini = 0.0\nsamples = 10\nvalue = [0,
10]'),
 Text(0.09645541399095901, 0.71875, 'gini = 0.0\nsamples = 1\nvalue = [1, 0]'),
 Text(0.11816625975393748, 0.84375, 'X[2] <= 149922.0\ngini = 0.401\nsamples =
746\nvalue = [207, 539]'),
 Text(0.1004493137194559, 0.8229166666666666, 'X[2] <= 27275.0\ngini =
0.313\nsamples = 263\nvalue = [51, 212]'),
 Text(0.09689310163243813, 0.8020833333333334, 'X[8] <= 1.5\ngini =
0.375\nsamples = 8\nvalue = [6, 2]'),
 Text(0.09645541399095901, 0.78125, 'gini = 0.0\nsamples = 2\nvalue = [0, 2]'),
 Text(0.09733078927391724, 0.78125, 'gini = 0.0\nsamples = 6\nvalue = [6, 0]'),
 Text(0.10400552580647367, 0.8020833333333334, 'X[6] <= 1.0\ngini =
0.291\nsamples = 255\nvalue = [45, 210]'),
 Text(0.09820616455687546, 0.78125, 'X[3] <= 11.5\ngini = 0.469\nsamples =
16\nvalue = [6, 10]'),
 Text(0.09776847691539635, 0.7604166666666666, 'gini = 0.0\nsamples = 2\nvalue =
[2, 0]'),
 Text(0.09864385219835457, 0.7604166666666666, 'X[2] <= 107825.0\ngini =
0.408\nsamples = 14\nvalue = [4, 10]'),
 Text(0.09776847691539635, 0.7395833333333334, 'X[1] <= 0.5\ngini =
0.198\nsamples = 9\nvalue = [1, 8]'),
```

```
 Text(0.09733078927391724, 0.71875, 'X[12] <= 37.5\ngini = 0.5\nsamples =
2\nvalue = [1, 1]'),
 Text(0.09689310163243813, 0.6979166666666666, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.09776847691539635, 0.6979166666666666, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.09820616455687546, 0.71875, 'gini = 0.0\nsamples = 7\nvalue = [0, 7]'),
 Text(0.09951922748131278, 0.7395833333333334, 'X[0] <= 28.0\ngini =
0.48\nsamples = 5\nvalue = [3, 2]'),
 Text(0.09908153983983368, 0.71875, 'X[2] <= 113184.5\ngini = 0.444\nsamples =
3\nvalue = [1, 2]'),
 Text(0.09864385219835457, 0.6979166666666666, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.09951922748131278, 0.6979166666666666, 'gini = 0.0\nsamples = 2\nvalue =
[0, 2]'),
 Text(0.0999569151227919, 0.71875, 'gini = 0.0\nsamples = 2\nvalue = [2, 0]'),
 Text(0.10980488705607189, 0.78125, 'X[6] <= 12.5\ngini = 0.273\nsamples =
239\nvalue = [39, 200]'),
 Text(0.10520916682054123, 0.7604166666666666, 'X[2] <= 77066.5\ngini =
0.257\nsamples = 225\nvalue = [34, 191]'),
 Text(0.1019265095094479, 0.7395833333333334, 'X[13] <= 18.5\ngini =
0.172\nsamples = 84\nvalue = [8, 76]'),
 Text(0.10083229040575012, 0.71875, 'X[13] <= 10.0\ngini = 0.5\nsamples =
2\nvalue = [1, 1]'),
 Text(0.10039460276427101, 0.6979166666666666, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.10126997804722923, 0.6979166666666666, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.10302072861314568, 0.71875, 'X[2] <= 35856.5\ngini = 0.156\nsamples =
82\nvalue = [7, 75]'),
 Text(0.10214535333018745, 0.6979166666666666, 'X[2] <= 34027.5\ngini =
0.298\nsamples = 22\nvalue = [4, 18]'),
 Text(0.10170766568870834, 0.6770833333333334, 'gini = 0.0\nsamples = 13\nvalue
= [0, 13]'),
 Text(0.10258304097166657, 0.6770833333333334, 'X[12] <= 42.5\ngini =
0.494\nsamples = 9\nvalue = [4, 5]'),
 Text(0.10214535333018745, 0.65625, 'X[4] <= 11.5\ngini = 0.444\nsamples =
6\nvalue = [4, 2]'),
 Text(0.10170766568870834, 0.6354166666666666, 'X[2] <= 34303.5\ngini =
0.32\nsamples = 5\nvalue = [4, 1]'),
 Text(0.10126997804722923, 0.6145833333333334, 'X[1] <= 2.0\ngini = 0.5\nsamples
= 2\nvalue = [1, 1]'),
 Text(0.10083229040575012, 0.59375, 'gini = 0.0\nsamples = 1\nvalue = [1, 0]'),
 Text(0.10170766568870834, 0.59375, 'gini = 0.0\nsamples = 1\nvalue = [0, 1]'),
 Text(0.10214535333018745, 0.6145833333333334, 'gini = 0.0\nsamples = 3\nvalue =
[3, 0]'),
 Text(0.10258304097166657, 0.6354166666666666, 'gini = 0.0\nsamples = 1\nvalue =
```

```
[0, 1]'),
 Text(0.10302072861314568, 0.65625, 'gini = 0.0\nsamples = 3\nvalue = [0, 3]'),
 Text(0.1038961038961039, 0.6979166666666666, 'X[4] <= 10.5\ngini =
0.095\nsamples = 60\nvalue = [3, 57]'),
 Text(0.10345841625462479, 0.6770833333333334, 'gini = 0.0\nsamples = 41\nvalue
= [0, 41]'),
 Text(0.104333791537583, 0.6770833333333334, 'X[0] <= 35.5\ngini =
0.266\nsamples = 19\nvalue = [3, 16]'),
 Text(0.1038961038961039, 0.65625, 'X[6] <= 3.5\ngini = 0.198\nsamples =
18\nvalue = [2, 16]'),
 Text(0.10345841625462479, 0.6354166666666666, 'X[0] <= 30.5\ngini =
0.444\nsamples = 6\nvalue = [2, 4]'),
 Text(0.10302072861314568, 0.6145833333333334, 'gini = 0.0\nsamples = 2\nvalue =
[2, 0]'),
 Text(0.1038961038961039, 0.6145833333333334, 'gini = 0.0\nsamples = 4\nvalue =
[0, 4]'),
 Text(0.104333791537583, 0.6354166666666666, 'gini = 0.0\nsamples = 12\nvalue =
[0, 12]'),
 Text(0.10477147917906211, 0.65625, 'gini = 0.0\nsamples = 1\nvalue = [1, 0]'),
 Text(0.10849182413163456, 0.7395833333333334, 'X[2] <= 90072.0\ngini =
0.301\nsamples = 141\nvalue = [26, 115]'),
 Text(0.10695991738645767, 0.71875, 'X[2] <= 89363.5\ngini = 0.475\nsamples =
18\nvalue = [7, 11]'),
 Text(0.10652222974497856, 0.6979166666666666, 'X[2] <= 87328.5\ngini =
0.43\nsamples = 16\nvalue = [5, 11]'),
 Text(0.10608454210349945, 0.6770833333333334, 'X[2] <= 83257.5\ngini =
0.486\nsamples = 12\nvalue = [5, 7]'),
 Text(0.10564685446202034, 0.65625, 'X[0] <= 35.5\ngini = 0.346\nsamples =
9\nvalue = [2, 7]'),
 Text(0.10520916682054123, 0.6354166666666666, 'X[3] <= 11.5\ngini =
0.219\nsamples = 8\nvalue = [1, 7]'),
 Text(0.10477147917906211, 0.6145833333333334, 'X[6] <= 5.0\ngini =
0.444\nsamples = 3\nvalue = [1, 2]'),
 Text(0.104333791537583, 0.59375, 'gini = 0.0\nsamples = 2\nvalue = [0, 2]'),
 Text(0.10520916682054123, 0.59375, 'gini = 0.0\nsamples = 1\nvalue = [1, 0]'),
 Text(0.10564685446202034, 0.6145833333333334, 'gini = 0.0\nsamples = 5\nvalue =
[0, 5]'),
 Text(0.10608454210349945, 0.6354166666666666, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.10652222974497856, 0.65625, 'gini = 0.0\nsamples = 3\nvalue = [3, 0]'),
 Text(0.10695991738645767, 0.6770833333333334, 'gini = 0.0\nsamples = 4\nvalue =
[0, 4]'),
 Text(0.10739760502793678, 0.6979166666666666, 'gini = 0.0\nsamples = 2\nvalue =
[2, 0]'),
 Text(0.11002373087681144, 0.71875, 'X[2] <= 112568.5\ngini = 0.261\nsamples =
123\nvalue = [19, 104]'),
 Text(0.10827298031089501, 0.6979166666666666, 'X[13] <= 34.5\ngini =
```

```
0.097\nsamples = 39\nvalue = [2, 37]'),
 Text(0.1078352926694159, 0.6770833333333334, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.10871066795237411, 0.6770833333333334, 'X[6] <= 11.5\ngini =
0.051\nsamples = 38\nvalue = [1, 37]'),
 Text(0.10827298031089501, 0.65625, 'gini = 0.0\nsamples = 34\nvalue = [0,
34]'),
 Text(0.10914835559385322, 0.65625, 'X[0] <= 34.5\ngini = 0.375\nsamples =
4\nvalue = [1, 3]'),
 Text(0.10871066795237411, 0.6354166666666666, 'gini = 0.0\nsamples = 3\nvalue =
[0, 3]'),
 Text(0.10958604323533233, 0.6354166666666666, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.11177448144272789, 0.6979166666666666, 'X[2] <= 137367.5\ngini =
0.323\nsamples = 84\nvalue = [17, 67]'),
 Text(0.11046141851829056, 0.6770833333333334, 'X[12] <= 39.0\ngini =
0.388\nsamples = 57\nvalue = [15, 42]'),
 Text(0.11002373087681144, 0.65625, 'gini = 0.0\nsamples = 1\nvalue = [1, 0]'),
 Text(0.11089910615976967, 0.65625, 'X[2] <= 113038.0\ngini = 0.375\nsamples =
56\nvalue = [14, 42]'),
 Text(0.11046141851829056, 0.6354166666666666, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.11133679380124878, 0.6354166666666666, 'X[0] <= 27.5\ngini =
0.361\nsamples = 55\nvalue = [13, 42]'),
 Text(0.11089910615976967, 0.6145833333333334, 'gini = 0.0\nsamples = 10\nvalue
= [0, 10]'),
 Text(0.11177448144272789, 0.6145833333333334, 'X[6] <= 2.5\ngini =
0.411\nsamples = 45\nvalue = [13, 32]'),
 Text(0.110242574697551, 0.59375, 'X[12] <= 52.5\ngini = 0.153\nsamples =
12\nvalue = [1, 11]'),
 Text(0.10980488705607189, 0.5729166666666666, 'gini = 0.0\nsamples = 9\nvalue =
[0, 9]'),
 Text(0.11068026233903011, 0.5729166666666666, 'X[12] <= 57.5\ngini =
0.444\nsamples = 3\nvalue = [1, 2]'),
 Text(0.110242574697551, 0.5520833333333334, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.11111794998050922, 0.5520833333333334, 'gini = 0.0\nsamples = 2\nvalue =
[0, 2]'),
 Text(0.11330638818790477, 0.59375, 'X[2] <= 132643.5\ngini = 0.463\nsamples =
33\nvalue = [12, 21]'),
 Text(0.11243101290494656, 0.5729166666666666, 'X[2] <= 125116.0\ngini =
0.428\nsamples = 29\nvalue = [9, 20]'),
 Text(0.11199332526346745, 0.5520833333333334, 'X[2] <= 124756.0\ngini =
0.483\nsamples = 22\nvalue = [9, 13]'),
 Text(0.11155563762198833, 0.53125, 'X[12] <= 52.0\ngini = 0.432\nsamples =
19\nvalue = [6, 13]'),
 Text(0.11111794998050922, 0.5104166666666666, 'X[0] <= 34.5\ngini =
```

```
0.49\nsamples = 14\nvalue = [6, 8]'),
 Text(0.11068026233903011, 0.4895833333333333, 'X[6] <= 4.5\ngini =
0.496\nsamples = 11\nvalue = [6, 5]'),
 Text(0.110242574697551, 0.46875, 'gini = 0.0\nsamples = 3\nvalue = [3, 0]'),
 Text(0.11111794998050922, 0.46875, 'X[6] <= 6.0\ngini = 0.469\nsamples =
8\nvalue = [3, 5]'),
 Text(0.11068026233903011, 0.4479166666666667, 'gini = 0.0\nsamples = 3\nvalue =
[0, 3]'),
 Text(0.11155563762198833, 0.4479166666666667, 'X[2] <= 116815.5\ngini =
0.48\nsamples = 5\nvalue = [3, 2]'),
 Text(0.11111794998050922, 0.4270833333333333, 'gini = 0.0\nsamples = 2\nvalue =
[2, 0]'),
 Text(0.11199332526346745, 0.4270833333333333, 'X[6] <= 11.5\ngini =
0.444\nsamples = 3\nvalue = [1, 2]'),
 Text(0.11155563762198833, 0.40625, 'gini = 0.0\nsamples = 2\nvalue = [0, 2]'),
 Text(0.11243101290494656, 0.40625, 'gini = 0.0\nsamples = 1\nvalue = [1, 0]'),
 Text(0.11155563762198833, 0.4895833333333333, 'gini = 0.0\nsamples = 3\nvalue =
[0, 3]'),
 Text(0.11199332526346745, 0.5104166666666666, 'gini = 0.0\nsamples = 5\nvalue =
[0, 5]'),
 Text(0.11243101290494656, 0.53125, 'gini = 0.0\nsamples = 3\nvalue = [3, 0]'),
 Text(0.11286870054642567, 0.5520833333333334, 'gini = 0.0\nsamples = 7\nvalue =
[0, 7]'),
 Text(0.11418176347086299, 0.5729166666666666, 'X[6] <= 10.5\ngini =
0.375\nsamples = 4\nvalue = [3, 1]'),
 Text(0.11374407582938388, 0.5520833333333334, 'gini = 0.0\nsamples = 3\nvalue =
[3, 0]'),
 Text(0.1146194511123421, 0.5520833333333334, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.11308754436716523, 0.6770833333333334, 'X[2] <= 147981.0\ngini =
0.137\nsamples = 27\nvalue = [2, 25]'),
 Text(0.11264985672568611, 0.65625, 'gini = 0.0\nsamples = 20\nvalue = [0,
20]'),
 Text(0.11352523200864433, 0.65625, 'X[0] <= 33.5\ngini = 0.408\nsamples =
7\nvalue = [2, 5]'),
 Text(0.11308754436716523, 0.6354166666666666, 'X[11] <= 742.5\ngini =
0.278\nsamples = 6\nvalue = [1, 5]'),
 Text(0.11264985672568611, 0.6145833333333334, 'gini = 0.0\nsamples = 5\nvalue =
[0, 5]'),
 Text(0.11352523200864433, 0.6145833333333334, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.11396291965012344, 0.6354166666666666, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.11440060729160255, 0.7604166666666666, 'X[12] <= 42.5\ngini =
0.459\nsamples = 14\nvalue = [5, 9]'),
 Text(0.11352523200864433, 0.7395833333333334, 'X[12] <= 35.0\ngini =
0.245\nsamples = 7\nvalue = [1, 6]'),
```

```
 Text(0.11308754436716523, 0.71875, 'gini = 0.0\nsamples = 1\nvalue = [1, 0]'),
 Text(0.11396291965012344, 0.71875, 'gini = 0.0\nsamples = 6\nvalue = [0, 6]'),
 Text(0.11527598257456077, 0.7395833333333334, 'X[2] <= 87111.0\ngini =
0.49\nsamples = 7\nvalue = [4, 3]'),
 Text(0.11483829493308166, 0.71875, 'X[2] <= 44501.0\ngini = 0.48\nsamples =
5\nvalue = [2, 3]'),
 Text(0.11440060729160255, 0.6979166666666666, 'X[4] <= 10.5\ngini =
0.444\nsamples = 3\nvalue = [2, 1]'),
 Text(0.11396291965012344, 0.6770833333333334, 'gini = 0.0\nsamples = 2\nvalue =
[2, 0]'),
 Text(0.11483829493308166, 0.6770833333333334, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.11527598257456077, 0.6979166666666666, 'gini = 0.0\nsamples = 2\nvalue =
[0, 2]'),
 Text(0.11571367021603989, 0.71875, 'gini = 0.0\nsamples = 2\nvalue = [2, 0]'),
 Text(0.13588320578841906, 0.8229166666666666, 'X[0] <= 30.5\ngini =
0.437\nsamples = 483\nvalue = [156, 327]'),
 Text(0.12457684495599188, 0.8020833333333334, 'X[6] <= 11.5\ngini =
0.359\nsamples = 222\nvalue = [52, 170]'),
 Text(0.12071972261545721, 0.78125, 'X[0] <= 24.5\ngini = 0.324\nsamples =
197\nvalue = [40, 157]'),
 Text(0.1188048391839861, 0.7604166666666666, 'X[12] <= 68.5\ngini =
0.105\nsamples = 36\nvalue = [2, 34]'),
 Text(0.11836715154250699, 0.7395833333333334, 'X[6] <= 1.0\ngini =
0.056\nsamples = 35\nvalue = [1, 34]'),
 Text(0.11792946390102788, 0.71875, 'X[2] <= 339707.5\ngini = 0.444\nsamples =
3\nvalue = [1, 2]'),
 Text(0.11749177625954878, 0.6979166666666666, 'gini = 0.0\nsamples = 2\nvalue =
[0, 2]'),
 Text(0.11836715154250699, 0.6979166666666666, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.1188048391839861, 0.71875, 'gini = 0.0\nsamples = 32\nvalue = [0, 32]'),
 Text(0.11924252682546521, 0.7395833333333334, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.12263460604692832, 0.7604166666666666, 'X[8] <= 1.5\ngini =
0.361\nsamples = 161\nvalue = [38, 123]'),
 Text(0.12011790210842344, 0.7395833333333334, 'X[13] <= 23.5\ngini =
0.49\nsamples = 7\nvalue = [4, 3]'),
 Text(0.11968021446694432, 0.71875, 'gini = 0.0\nsamples = 2\nvalue = [2, 0]'),
 Text(0.12055558974990255, 0.71875, 'X[1] <= 4.0\ngini = 0.48\nsamples =
5\nvalue = [2, 3]'),
 Text(0.12011790210842344, 0.6979166666666666, 'gini = 0.0\nsamples = 2\nvalue =
[0, 2]'),
 Text(0.12099327739138166, 0.6979166666666666, 'X[2] <= 314988.5\ngini =
0.444\nsamples = 3\nvalue = [2, 1]'),
 Text(0.12055558974990255, 0.6770833333333334, 'gini = 0.0\nsamples = 2\nvalue =
[2, 0]'),
```

```
 Text(0.12143096503286077, 0.6770833333333334, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.1251513099854332, 0.7395833333333334, 'X[2] <= 158843.5\ngini =
0.344\nsamples = 154\nvalue = [34, 120]'),
 Text(0.12361940324025632, 0.71875, 'X[0] <= 26.5\ngini = 0.496\nsamples =
11\nvalue = [5, 6]'),
 Text(0.1227440279572981, 0.6979166666666666, 'X[12] <= 45.0\ngini =
0.32\nsamples = 5\nvalue = [1, 4]'),
 Text(0.122306340315819, 0.6770833333333334, 'gini = 0.0\nsamples = 4\nvalue =
[0, 4]'),
 Text(0.1231817155987772, 0.6770833333333334, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.12449477852321454, 0.6979166666666666, 'X[2] <= 153644.0\ngini =
0.444\nsamples = 6\nvalue = [4, 2]'),
 Text(0.12405709088173543, 0.6770833333333334, 'X[12] <= 42.5\ngini =
0.444\nsamples = 3\nvalue = [1, 2]'),
 Text(0.12361940324025632, 0.65625, 'gini = 0.0\nsamples = 1\nvalue = [1, 0]'),
 Text(0.12449477852321454, 0.65625, 'gini = 0.0\nsamples = 2\nvalue = [0, 2]'),
 Text(0.12493246616469365, 0.6770833333333334, 'gini = 0.0\nsamples = 3\nvalue =
[3, 0]'),
 Text(0.1266832167306101, 0.71875, 'X[2] <= 577805.5\ngini = 0.323\nsamples =
143\nvalue = [29, 114]'),
 Text(0.12624552908913098, 0.6979166666666666, 'X[5] <= 1.5\ngini =
0.317\nsamples = 142\nvalue = [28, 114]'),
 Text(0.12580784144765186, 0.6770833333333334, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.1266832167306101, 0.6770833333333334, 'X[2] <= 393185.0\ngini =
0.31\nsamples = 141\nvalue = [27, 114]'),
 Text(0.12624552908913098, 0.65625, 'X[2] <= 382449.0\ngini = 0.333\nsamples =
128\nvalue = [27, 101]'),
 Text(0.12320907107636965, 0.6354166666666666, 'X[0] <= 28.5\ngini =
0.32\nsamples = 125\nvalue = [25, 100]'),
 Text(0.11844921797528432, 0.6145833333333334, 'X[2] <= 191756.5\ngini =
0.372\nsamples = 77\nvalue = [19, 58]'),
 Text(0.11549482639530033, 0.59375, 'X[3] <= 7.5\ngini = 0.188\nsamples =
19\nvalue = [2, 17]'),
 Text(0.11505713875382122, 0.5729166666666666, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.11593251403677944, 0.5729166666666666, 'X[2] <= 174627.0\ngini =
0.105\nsamples = 18\nvalue = [1, 17]'),
 Text(0.11549482639530033, 0.5520833333333334, 'X[2] <= 166950.5\ngini =
0.5\nsamples = 2\nvalue = [1, 1]'),
 Text(0.11505713875382122, 0.53125, 'gini = 0.0\nsamples = 1\nvalue = [0, 1]'),
 Text(0.11593251403677944, 0.53125, 'gini = 0.0\nsamples = 1\nvalue = [1, 0]'),
 Text(0.11637020167825855, 0.5520833333333334, 'gini = 0.0\nsamples = 16\nvalue
= [0, 16]'),
 Text(0.12140360955526833, 0.59375, 'X[4] <= 10.5\ngini = 0.414\nsamples =
```

```
58\nvalue = [17, 41]'),
 Text(0.11899632752713321, 0.5729166666666666, 'X[2] <= 197888.0\ngini =
0.454\nsamples = 46\nvalue = [16, 30]'),
 Text(0.1185586398856541, 0.5520833333333334, 'gini = 0.0\nsamples = 2\nvalue =
[2, 0]'),
 Text(0.11943401516861232, 0.5520833333333334, 'X[6] <= 5.5\ngini =
0.434\nsamples = 44\nvalue = [14, 30]'),
 Text(0.11680788931973766, 0.53125, 'X[2] <= 207733.0\ngini = 0.491\nsamples =
23\nvalue = [10, 13]'),
 Text(0.11637020167825855, 0.5104166666666666, 'gini = 0.0\nsamples = 2\nvalue =
[0, 2]'),
 Text(0.11724557696121678, 0.5104166666666666, 'X[2] <= 267687.0\ngini =
0.499\nsamples = 21\nvalue = [10, 11]'),
 Text(0.11637020167825855, 0.4895833333333333, 'X[12] <= 42.5\ngini =
0.49\nsamples = 14\nvalue = [8, 6]'),
 Text(0.11593251403677944, 0.46875, 'X[0] <= 25.5\ngini = 0.496\nsamples =
11\nvalue = [5, 6]'),
 Text(0.11549482639530033, 0.4479166666666667, 'gini = 0.0\nsamples = 3\nvalue =
[0, 3]'),
 Text(0.11637020167825855, 0.4479166666666667, 'X[10] <= 1414.5\ngini =
0.469\nsamples = 8\nvalue = [5, 3]'),
 Text(0.11593251403677944, 0.4270833333333333, 'X[2] <= 224056.0\ngini =
0.408\nsamples = 7\nvalue = [5, 2]'),
 Text(0.11549482639530033, 0.40625, 'gini = 0.0\nsamples = 4\nvalue = [4, 0]'),
 Text(0.11637020167825855, 0.40625, 'X[6] <= 1.0\ngini = 0.444\nsamples =
3\nvalue = [1, 2]'),
 Text(0.11593251403677944, 0.3854166666666667, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.11680788931973766, 0.3854166666666667, 'gini = 0.0\nsamples = 2\nvalue =
[0, 2]'),
 Text(0.11680788931973766, 0.4270833333333333, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.11680788931973766, 0.46875, 'gini = 0.0\nsamples = 3\nvalue = [3, 0]'),
 Text(0.11812095224417499, 0.4895833333333333, 'X[2] <= 320390.0\ngini =
0.408\nsamples = 7\nvalue = [2, 5]'),
 Text(0.11768326460269589, 0.46875, 'gini = 0.0\nsamples = 4\nvalue = [0, 4]'),
 Text(0.1185586398856541, 0.46875, 'X[2] <= 339164.0\ngini = 0.444\nsamples =
3\nvalue = [2, 1]'),
 Text(0.11812095224417499, 0.4479166666666667, 'gini = 0.0\nsamples = 2\nvalue =
[2, 0]'),
 Text(0.11899632752713321, 0.4479166666666667, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.122060141017487, 0.53125, 'X[2] <= 293717.5\ngini = 0.308\nsamples =
21\nvalue = [4, 17]'),
 Text(0.12118476573452877, 0.5104166666666666, 'X[2] <= 215188.5\ngini =
0.219\nsamples = 16\nvalue = [2, 14]'),
 Text(0.12074707809304966, 0.4895833333333333, 'X[2] <= 212454.5\ngini =
```

```
0.408\nsamples = 7\nvalue = [2, 5]'),
 Text(0.12030939045157055, 0.46875, 'X[2] <= 201193.5\ngini = 0.278\nsamples =
6\nvalue = [1, 5]'),
 Text(0.11987170281009144, 0.4479166666666667, 'X[12] <= 37.5\ngini =
0.5\nsamples = 2\nvalue = [1, 1]'),
 Text(0.11943401516861232, 0.4270833333333333, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.12030939045157055, 0.4270833333333333, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.12074707809304966, 0.4479166666666667, 'gini = 0.0\nsamples = 4\nvalue =
[0, 4]'),
 Text(0.12118476573452877, 0.46875, 'gini = 0.0\nsamples = 1\nvalue = [1, 0]'),
 Text(0.12162245337600788, 0.4895833333333333, 'gini = 0.0\nsamples = 9\nvalue =
[0, 9]'),
 Text(0.1229355163004452, 0.5104166666666666, 'X[2] <= 306013.0\ngini =
0.48\nsamples = 5\nvalue = [2, 3]'),
 Text(0.1224978286589661, 0.4895833333333333, 'gini = 0.0\nsamples = 2\nvalue =
[2, 0]'),
 Text(0.12337320394192432, 0.4895833333333333, 'gini = 0.0\nsamples = 3\nvalue =
[0, 3]'),
 Text(0.12381089158340343, 0.5729166666666666, 'X[6] <= 2.5\ngini =
0.153\nsamples = 12\nvalue = [1, 11]'),
 Text(0.12337320394192432, 0.5520833333333334, 'X[2] <= 230000.0\ngini =
0.444\nsamples = 3\nvalue = [1, 2]'),
 Text(0.1229355163004452, 0.53125, 'gini = 0.0\nsamples = 2\nvalue = [0, 2]'),
 Text(0.12381089158340343, 0.53125, 'gini = 0.0\nsamples = 1\nvalue = [1, 0]'),
 Text(0.12424857922488254, 0.5520833333333334, 'gini = 0.0\nsamples = 9\nvalue =
[0, 9]'),
 Text(0.12796892417745498, 0.6145833333333334, 'X[2] <= 231483.5\ngini =
0.219\nsamples = 48\nvalue = [6, 42]'),
 Text(0.12753123653597587, 0.59375, 'X[2] <= 231228.0\ngini = 0.32\nsamples =
30\nvalue = [6, 24]'),
 Text(0.12709354889449676, 0.5729166666666666, 'X[6] <= 10.5\ngini =
0.285\nsamples = 29\nvalue = [5, 24]'),
 Text(0.12556164214931986, 0.5520833333333334, 'X[13] <= 31.0\ngini =
0.219\nsamples = 24\nvalue = [3, 21]'),
 Text(0.12468626686636165, 0.53125, 'X[2] <= 197964.0\ngini = 0.5\nsamples =
2\nvalue = [1, 1]'),
 Text(0.12424857922488254, 0.5104166666666666, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.12512395450784075, 0.5104166666666666, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.1264370174322781, 0.53125, 'X[0] <= 29.5\ngini = 0.165\nsamples =
22\nvalue = [2, 20]'),
 Text(0.12599932979079898, 0.5104166666666666, 'gini = 0.0\nsamples = 10\nvalue
= [0, 10]'),
 Text(0.1268747050737572, 0.5104166666666666, 'X[2] <= 195606.0\ngini =
```

```
0.278\nsamples = 12\nvalue = [2, 10]'),
 Text(0.1264370174322781, 0.4895833333333333, 'gini = 0.0\nsamples = 6\nvalue =
[0, 6]'),
 Text(0.1273123927152363, 0.4895833333333333, 'X[2] <= 206295.5\ngini =
0.444\nsamples = 6\nvalue = [2, 4]'),
 Text(0.1268747050737572, 0.46875, 'X[6] <= 7.5\ngini = 0.444\nsamples =
3\nvalue = [2, 1]'),
 Text(0.1264370174322781, 0.4479166666666667, 'gini = 0.0\nsamples = 2\nvalue =
[2, 0]'),
 Text(0.1273123927152363, 0.4479166666666667, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.12775008035671542, 0.46875, 'gini = 0.0\nsamples = 3\nvalue = [0, 3]'),
 Text(0.12862545563967365, 0.5520833333333334, 'X[1] <= 4.0\ngini =
0.48\nsamples = 5\nvalue = [2, 3]'),
 Text(0.12818776799819454, 0.53125, 'X[4] <= 10.5\ngini = 0.444\nsamples =
3\nvalue = [2, 1]'),
 Text(0.12775008035671542, 0.5104166666666666, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.12862545563967365, 0.5104166666666666, 'gini = 0.0\nsamples = 2\nvalue =
[2, 0]'),
 Text(0.12906314328115276, 0.53125, 'gini = 0.0\nsamples = 2\nvalue = [0, 2]'),
 Text(0.12796892417745498, 0.5729166666666666, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.1284066118189341, 0.59375, 'gini = 0.0\nsamples = 18\nvalue = [0, 18]'),
 Text(0.12928198710189232, 0.6354166666666666, 'X[13] <= 32.5\ngini =
0.444\nsamples = 3\nvalue = [2, 1]'),
 Text(0.1288442994604132, 0.6145833333333334, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.12971967474337143, 0.6145833333333334, 'gini = 0.0\nsamples = 2\nvalue =
[2, 0]'),
 Text(0.1271209043720892, 0.65625, 'gini = 0.0\nsamples = 13\nvalue = [0, 13]'),
 Text(0.1271209043720892, 0.6979166666666666, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.12843396729652654, 0.78125, 'X[12] <= 43.5\ngini = 0.499\nsamples =
25\nvalue = [12, 13]'),
 Text(0.1275585920135683, 0.7604166666666666, 'X[1] <= 0.5\ngini = 0.43\nsamples
= 16\nvalue = [5, 11]'),
 Text(0.1271209043720892, 0.7395833333333334, 'gini = 0.0\nsamples = 2\nvalue =
[2, 0]'),
 Text(0.12799627965504742, 0.7395833333333334, 'X[2] <= 169068.0\ngini =
0.337\nsamples = 14\nvalue = [3, 11]'),
 Text(0.1275585920135683, 0.71875, 'gini = 0.0\nsamples = 1\nvalue = [1, 0]'),
 Text(0.12843396729652654, 0.71875, 'X[0] <= 26.5\ngini = 0.26\nsamples =
13\nvalue = [2, 11]'),
 Text(0.12799627965504742, 0.6979166666666666, 'gini = 0.0\nsamples = 5\nvalue =
[0, 5]'),
 Text(0.12887165493800565, 0.6979166666666666, 'X[0] <= 28.5\ngini =
```

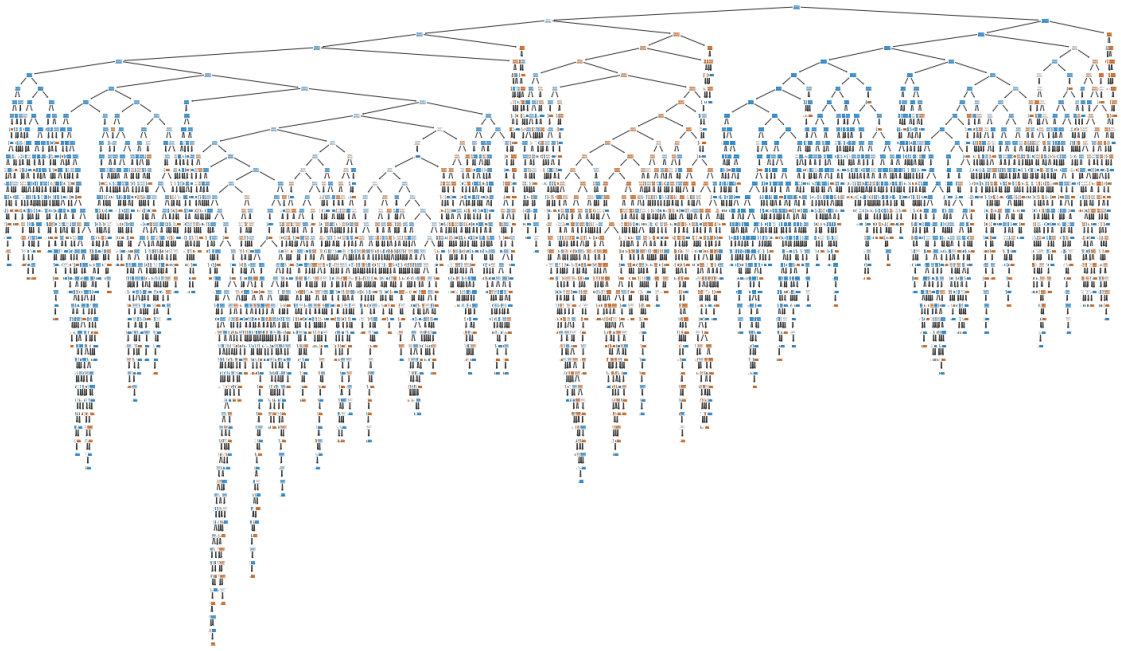
```
0.375\nsamples = 8\nvalue = [2, 6]'),
 Text(0.12843396729652654, 0.6770833333333334, 'X[10] <= 1705.5\ngini =
0.5\nsamples = 4\nvalue = [2, 2]'),
 Text(0.12799627965504742, 0.65625, 'gini = 0.0\nsamples = 2\nvalue = [2, 0]'),
 Text(0.12887165493800565, 0.65625, 'gini = 0.0\nsamples = 2\nvalue = [0, 2]'),
 Text(0.12930934257948476, 0.6770833333333334, 'gini = 0.0\nsamples = 4\nvalue =
[0, 4]'),
 Text(0.12930934257948476, 0.7604166666666666, 'X[12] <= 56.0\ngini =
0.346\nsamples = 9\nvalue = [7, 2]'),
 Text(0.12887165493800565, 0.7395833333333334, 'gini = 0.0\nsamples = 6\nvalue =
[6, 0]'),
 Text(0.12974703022096387, 0.7395833333333334, 'X[0] <= 27.0\ngini =
0.444\nsamples = 3\nvalue = [1, 2]'),
 Text(0.12930934257948476, 0.71875, 'gini = 0.0\nsamples = 2\nvalue = [0, 2]'),
 Text(0.13018471786244298, 0.71875, 'gini = 0.0\nsamples = 1\nvalue = [1, 0]'),
 Text(0.14718956662084623, 0.8020833333333334, 'X[12] <= 39.5\ngini =
0.479\nsamples = 261\nvalue = [104, 157]'),
 Text(0.1444779549044952, 0.78125, 'X[0] <= 31.5\ngini = 0.111\nsamples =
17\nvalue = [1, 16]'),
 Text(0.14404026726301608, 0.7604166666666666, 'X[12] <= 36.5\ngini =
0.5\nsamples = 2\nvalue = [1, 1]'),
 Text(0.14360257962153697, 0.7395833333333334, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.1444779549044952, 0.7395833333333334, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.1449156425459743, 0.7604166666666666, 'gini = 0.0\nsamples = 15\nvalue =
[0, 15]'),
 Text(0.1499011783371973, 0.78125, 'X[12] <= 52.5\ngini = 0.488\nsamples =
244\nvalue = [103, 141]'),
 Text(0.14579101782893253, 0.7604166666666666, 'X[10] <= 4147.0\ngini =
0.474\nsamples = 197\nvalue = [76, 121]'),
 Text(0.14535333018745342, 0.7395833333333334, 'X[2] <= 350508.5\ngini =
0.466\nsamples = 192\nvalue = [71, 121]'),
 Text(0.13975913501979853, 0.71875, 'X[2] <= 208041.0\ngini = 0.45\nsamples =
170\nvalue = [58, 112]'),
 Text(0.13371357447186832, 0.6979166666666666, 'X[2] <= 205105.5\ngini =
0.49\nsamples = 84\nvalue = [36, 48]'),
 Text(0.1332758868303892, 0.6770833333333334, 'X[2] <= 153504.0\ngini =
0.483\nsamples = 81\nvalue = [33, 48]'),
 Text(0.13147042530928787, 0.65625, 'X[4] <= 10.5\ngini = 0.32\nsamples =
5\nvalue = [4, 1]'),
 Text(0.13103273766780876, 0.6354166666666666, 'gini = 0.0\nsamples = 4\nvalue =
[4, 0]'),
 Text(0.131908112950767, 0.6354166666666666, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.13508134835149052, 0.65625, 'X[2] <= 177657.0\ngini = 0.472\nsamples =
76\nvalue = [29, 47]'),
```

```
 Text(0.1327834882337252, 0.6354166666666666, 'X[6] <= 2.5\ngini =
0.375\nsamples = 32\nvalue = [8, 24]'),
 Text(0.13125158148854832, 0.6145833333333334, 'X[2] <= 163740.0\ngini =
0.497\nsamples = 13\nvalue = [6, 7]'),
 Text(0.1303762062055901, 0.59375, 'X[2] <= 158238.0\ngini = 0.32\nsamples =
5\nvalue = [4, 1]'),
 Text(0.12993851856411098, 0.5729166666666666, 'X[2] <= 156220.0\ngini =
0.5\nsamples = 2\nvalue = [1, 1]'),
 Text(0.12950083092263187, 0.5520833333333334, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.1303762062055901, 0.5520833333333334, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.1308138938470692, 0.5729166666666666, 'gini = 0.0\nsamples = 3\nvalue =
[3, 0]'),
 Text(0.13212695677150654, 0.59375, 'X[12] <= 44.0\ngini = 0.375\nsamples =
8\nvalue = [2, 6]'),
 Text(0.13168926913002743, 0.5729166666666666, 'gini = 0.0\nsamples = 6\nvalue =
[0, 6]'),
 Text(0.13256464441298565, 0.5729166666666666, 'gini = 0.0\nsamples = 2\nvalue =
[2, 0]'),
 Text(0.13431539497890208, 0.6145833333333334, 'X[0] <= 32.5\ngini =
0.188\nsamples = 19\nvalue = [2, 17]'),
 Text(0.133877707337423, 0.59375, 'X[6] <= 4.0\ngini = 0.48\nsamples = 5\nvalue
= [2, 3]'),
 Text(0.13344001969594388, 0.5729166666666666, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.13431539497890208, 0.5729166666666666, 'X[4] <= 10.5\ngini =
0.5\nsamples = 4\nvalue = [2, 2]'),
 Text(0.133877707337423, 0.5520833333333334, 'X[6] <= 7.0\ngini = 0.444\nsamples
= 3\nvalue = [1, 2]'),
 Text(0.13344001969594388, 0.53125, 'gini = 0.0\nsamples = 1\nvalue = [0, 1]'),
 Text(0.13431539497890208, 0.53125, 'X[2] <= 166656.0\ngini = 0.5\nsamples =
2\nvalue = [1, 1]'),
 Text(0.133877707337423, 0.5104166666666666, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.1347530826203812, 0.5104166666666666, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.1347530826203812, 0.5520833333333334, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.1347530826203812, 0.59375, 'gini = 0.0\nsamples = 14\nvalue = [0, 14]'),
 Text(0.13737920846925586, 0.6354166666666666, 'X[2] <= 184373.5\ngini =
0.499\nsamples = 44\nvalue = [21, 23]'),
 Text(0.13606614554481852, 0.6145833333333334, 'X[0] <= 34.5\ngini =
0.375\nsamples = 12\nvalue = [9, 3]'),
 Text(0.1356284579033394, 0.59375, 'X[2] <= 180080.5\ngini = 0.5\nsamples =
6\nvalue = [3, 3]'),
 Text(0.1351907702618603, 0.5729166666666666, 'gini = 0.0\nsamples = 1\nvalue =
```

```
[1, 0]'),
 Text(0.13606614554481852, 0.5729166666666666, 'X[6] <= 4.5\ngini =
0.48\nsamples = 5\nvalue = [2, 3]'),
 Text(0.1356284579033394, 0.5520833333333334, 'X[6] <= 2.5\ngini =
0.444\nsamples = 3\nvalue = [2, 1]'),
 Text(0.1351907702618603, 0.53125, 'gini = 0.0\nsamples = 1\nvalue = [0, 1]'),
 Text(0.13606614554481852, 0.53125, 'gini = 0.0\nsamples = 2\nvalue = [2, 0]'),
 Text(0.13650383318629764, 0.5520833333333334, 'gini = 0.0\nsamples = 2\nvalue =
[0, 2]'),
 Text(0.13650383318629764, 0.59375, 'gini = 0.0\nsamples = 6\nvalue = [6, 0]'),
 Text(0.1386922713936932, 0.6145833333333334, 'X[2] <= 203485.0\ngini =
0.469\nsamples = 32\nvalue = [12, 20]'),
 Text(0.13825458375221408, 0.59375, 'X[2] <= 198062.5\ngini = 0.49\nsamples =
28\nvalue = [12, 16]'),
 Text(0.13781689611073497, 0.5729166666666666, 'X[2] <= 196165.5\ngini =
0.444\nsamples = 24\nvalue = [8, 16]'),
 Text(0.13737920846925586, 0.5520833333333334, 'X[2] <= 195746.5\ngini =
0.472\nsamples = 21\nvalue = [8, 13]'),
 Text(0.13694152082777675, 0.53125, 'X[2] <= 191731.0\ngini = 0.432\nsamples =
19\nvalue = [6, 13]'),
 Text(0.13606614554481852, 0.5104166666666666, 'X[2] <= 189386.0\ngini =
0.355\nsamples = 13\nvalue = [3, 10]'),
 Text(0.1356284579033394, 0.4895833333333333, 'X[2] <= 186515.5\ngini =
0.469\nsamples = 8\nvalue = [3, 5]'),
 Text(0.1351907702618603, 0.46875, 'X[0] <= 34.5\ngini = 0.278\nsamples =
6\nvalue = [1, 5]'),
 Text(0.1347530826203812, 0.4479166666666667, 'X[12] <= 42.5\ngini =
0.5\nsamples = 2\nvalue = [1, 1]'),
 Text(0.13431539497890208, 0.4270833333333333, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.1351907702618603, 0.4270833333333333, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.1356284579033394, 0.4479166666666667, 'gini = 0.0\nsamples = 4\nvalue =
[0, 4]'),
 Text(0.13606614554481852, 0.46875, 'gini = 0.0\nsamples = 2\nvalue = [2, 0]'),
 Text(0.13650383318629764, 0.4895833333333333, 'gini = 0.0\nsamples = 5\nvalue =
[0, 5]'),
 Text(0.13781689611073497, 0.5104166666666666, 'X[2] <= 193745.5\ngini =
0.5\nsamples = 6\nvalue = [3, 3]'),
 Text(0.13737920846925586, 0.4895833333333333, 'X[4] <= 10.5\ngini =
0.375\nsamples = 4\nvalue = [3, 1]'),
 Text(0.13694152082777675, 0.46875, 'gini = 0.0\nsamples = 3\nvalue = [3, 0]'),
 Text(0.13781689611073497, 0.46875, 'gini = 0.0\nsamples = 1\nvalue = [0, 1]'),
 Text(0.13825458375221408, 0.4895833333333333, 'gini = 0.0\nsamples = 2\nvalue =
[0, 2]'),
 Text(0.13781689611073497, 0.53125, 'gini = 0.0\nsamples = 2\nvalue = [2, 0]'),
 Text(0.13825458375221408, 0.5520833333333334, 'gini = 0.0\nsamples = 3\nvalue =
```

```
[0, 3]'),
 Text(0.1386922713936932, 0.5729166666666666, 'gini = 0.0\nsamples = 4\nvalue =
[4, 0]'),
 Text(0.1391299590351723, 0.59375, 'gini = 0.0\nsamples = 4\nvalue = [0, 4]'),
 Text(0.1341512621133474, 0.6770833333333334, 'gini = 0.0\nsamples = 3\nvalue =
[3, 0]'),
 Text(0.14580469556772874, 0.6979166666666666, 'X[12] <= 44.0\ngini =
0.381\nsamples = 86\nvalue = [22, 64]'),
 Text(0.14328799162922384, 0.6770833333333334, 'X[6] <= 8.0\ngini =
0.444\nsamples = 54\nvalue = [18, 36]'),
 Text(0.14088070960108875, 0.65625, 'X[13] <= 14.5\ngini = 0.32\nsamples =
35\nvalue = [7, 28]'),
 Text(0.14000533431813053, 0.6354166666666666, 'X[3] <= 11.5\ngini =
0.444\nsamples = 3\nvalue = [2, 1]'),
 Text(0.13956764667665142, 0.6145833333333334, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.14044302195960964, 0.6145833333333334, 'gini = 0.0\nsamples = 2\nvalue =
[2, 0]'),
 Text(0.14175608488404698, 0.6354166666666666, 'X[2] <= 337738.5\ngini =
0.264\nsamples = 32\nvalue = [5, 27]'),
 Text(0.14131839724256787, 0.6145833333333334, 'X[2] <= 213447.5\ngini =
0.225\nsamples = 31\nvalue = [4, 27]'),
 Text(0.14044302195960964, 0.59375, 'X[2] <= 211926.0\ngini = 0.444\nsamples =
6\nvalue = [2, 4]'),
 Text(0.14000533431813053, 0.5729166666666666, 'X[6] <= 1.0\ngini =
0.32\nsamples = 5\nvalue = [1, 4]'),
 Text(0.13956764667665142, 0.5520833333333334, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.14044302195960964, 0.5520833333333334, 'gini = 0.0\nsamples = 4\nvalue =
[0, 4]'),
 Text(0.14088070960108875, 0.5729166666666666, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.1421937725255261, 0.59375, 'X[2] <= 301705.0\ngini = 0.147\nsamples =
25\nvalue = [2, 23]'),
 Text(0.14175608488404698, 0.5729166666666666, 'gini = 0.0\nsamples = 17\nvalue
= [0, 17]'),
 Text(0.1426314601670052, 0.5729166666666666, 'X[2] <= 323092.5\ngini =
0.375\nsamples = 8\nvalue = [2, 6]'),
 Text(0.1421937725255261, 0.5520833333333334, 'X[4] <= 11.0\ngini =
0.444\nsamples = 3\nvalue = [2, 1]'),
 Text(0.14175608488404698, 0.53125, 'gini = 0.0\nsamples = 2\nvalue = [2, 0]'),
 Text(0.1426314601670052, 0.53125, 'gini = 0.0\nsamples = 1\nvalue = [0, 1]'),
 Text(0.14306914780848431, 0.5520833333333334, 'gini = 0.0\nsamples = 5\nvalue =
[0, 5]'),
 Text(0.1421937725255261, 0.6145833333333334, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.14569527365735896, 0.65625, 'X[6] <= 12.5\ngini = 0.488\nsamples =
```

```
19\nvalue = [11, 8]'),
 Text(0.14525758601587985, 0.6354166666666666, 'X[8] <= 3.0\ngini =
0.457\nsamples = 17\nvalue = [11, 6]'),
 Text(0.14481989837440074, 0.6145833333333334, 'gini = 0.0\nsamples = 3\nvalue =
[3, 0]'),
 Text(0.14569527365735896, 0.6145833333333334, 'X[6] <= 11.5\ngini =
0.49\nsamples = 14\nvalue = [8, 6]'),
 Text(0.14525758601587985, 0.59375, 'X[0] <= 33.5\ngini = 0.5\nsamples =
12\nvalue = [6, 6]'),
 Text(0.14438221073292162, 0.5729166666666666, 'X[2] <= 221863.5\ngini =
0.375\nsamples = 4\nvalue = [1, 3]'),
 Text(0.1439445230914425, 0.5520833333333334, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.14481989837440074, 0.5520833333333334, 'gini = 0.0\nsamples = 3\nvalue =
[0, 3]'),
 Text(0.14613296129883807, 0.5729166666666666, 'X[2] <= 217880.0\ngini =
0.469\nsamples = 8\nvalue = [5, 3]'),
 Text(0.14569527365735896, 0.5520833333333334, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.14657064894031718, 0.5520833333333334, 'X[2] <= 266028.5\ngini =
0.408\nsamples = 7\nvalue = [5, 2]'),
 Text(0.14613296129883807, 0.53125, 'gini = 0.0\nsamples = 4\nvalue = [4, 0]'),
 Text(0.1470083365817963, 0.53125, 'X[6] <= 9.5\ngini = 0.444\nsamples =
3\nvalue = [1, 2]'),
 Text(0.14657064894031718, 0.5104166666666666, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.1474460242232754, 0.5104166666666666, 'gini = 0.0\nsamples = 2\nvalue =
[0, 2]'),
 …]
```



[92]:
```
y_predict=model.predict(X_test)
```

[93]:
```
accuracy_score(y_test,y_predict)
```

[93]: 0.8073701842546064

[95]:
```
print(classification_report(y_test,y_predict))
```

```
              precision    recall  f1-score   support

         0.0       0.59      0.62      0.61      2550
         1.0       0.88      0.86      0.87      8196

    accuracy                           0.81     10746
   macro avg       0.73      0.74      0.74     10746
weighted avg       0.81      0.81      0.81     10746
```

1.0.9 Random Forest Classifier

[96]:
```
from sklearn.ensemble import RandomForestClassifier
Rf_model=RandomForestClassifier()
```

[97]:
```
Rf_model.fit(X_train,y_train)
```

[97]: RandomForestClassifier()

[98]:
```
y_pred_rf=Rf_model.predict(X_test)
```

[99]:
```
accuracy_score(y_test,y_pred_rf)
```

[99]:
```
0.8555741671319561
```

[100]:
```
print(classification_report(y_test,y_pred_rf))
```

```
              precision    recall  f1-score   support

         0.0       0.73      0.63      0.67      2550
         1.0       0.89      0.93      0.91      8196

    accuracy                           0.86     10746
   macro avg       0.81      0.78      0.79     10746
weighted avg       0.85      0.86      0.85     10746
```

1.0.10 Bagging Classifier

[102]:
```
from sklearn.svm import SVC
from sklearn.ensemble import BaggingClassifier
from sklearn.datasets import make_classification

model_bagging_svc = BaggingClassifier(base_estimator=SVC(),n_estimators=50,␣
 ↪random_state=0).fit(X_train,y_train)
```

[103]:
```
y_predict_bagging=model_bagging_svc.predict(X_test)
```

[104]:
```
accuracy_score(y_test,y_predict_bagging)
```

[104]:
```
0.7968546435883119
```

[105]:
```
print(classification_report(y_test,y_predict_bagging))
```

```
              precision    recall  f1-score   support

         0.0       0.99      0.15      0.25      2550
         1.0       0.79      1.00      0.88      8196

    accuracy                           0.80     10746
   macro avg       0.89      0.57      0.57     10746
weighted avg       0.84      0.80      0.73     10746
```

1.0.11 Extra Trees Classifier

[106]:
```
from sklearn.ensemble import ExtraTreesClassifier
clf = ExtraTreesClassifier(n_estimators=100, random_state=0)
```

[122]:
```
clf.fit(X_train,y_train)
```

[122]:
```
ExtraTreesClassifier(random_state=0)
```

[123]:
```
clf.score(X_train,y_train)
```

[123]:
```
1.0
```

[108]:
```
y_predict_clf=clf.predict(X_test)
```

[109]:
```
accuracy_score(y_test,y_predict_clf)
```

[109]:
```
0.8458030895216825
```

[110]:
```
print(classification_report(y_test,y_predict_bagging))
```

```
              precision    recall  f1-score   support

         0.0       0.99      0.15      0.25      2550
         1.0       0.79      1.00      0.88      8196

    accuracy                           0.80     10746
   macro avg       0.89      0.57      0.57     10746
weighted avg       0.84      0.80      0.73     10746
```

1.0.12 Voting Classifier

[111]:
```
from sklearn.linear_model import LogisticRegression
from sklearn.naive_bayes import GaussianNB
from sklearn.ensemble import RandomForestClassifier, VotingClassifier
clf1 = LogisticRegression(multi_class='multinomial', random_state=1)
clf2 = RandomForestClassifier(n_estimators=50, random_state=1)
clf3 = GaussianNB()
eclf1 = VotingClassifier(estimators=[ ('lr', clf1), ('rf', clf2), ('gnb',␣
 ↪clf3)], voting='hard')
eclf1 = eclf1.fit(X_train, y_train)
```

[112]:
```
y_predict_eclf1=eclf1.predict(X_test)
```

[113]:
```
accuracy_score(y_test,y_predict_eclf1)
```

[113]:
```
0.8179787828029034
```

[114]:
```
print(classification_report(y_test,y_predict_eclf1))
```

```
              precision    recall  f1-score   support

         0.0       0.79      0.32      0.45      2550
         1.0       0.82      0.97      0.89      8196

    accuracy                           0.82     10746
   macro avg       0.81      0.65      0.67     10746
weighted avg       0.81      0.82      0.79     10746
```

From the Model Building Stage We can Conclude the Model with Random Forest Classifier got the maximum score of 85% so we can use this model for productionizing our application