



Importing Libraries

```
In [1]: import numpy as np, pandas as pd  
import matplotlib.pyplot as plt, seaborn as sns
```

```
In [2]: from IPython.core.interactiveshell import InteractiveShell  
InteractiveShell.ast_node_interactivity = "all"  
%matplotlib inline  
pd.set_option("display.max_columns", 300)  
pd.set_option("display.max_rows", 300)
```

Reading Application Dataset

```
In [3]: app_data=pd.read_csv(r'C:\Users\nikhi\Data Science\EDA Assignment\application_
```

```
In [4]: app_data.head()
```

```
Out[4]: SK_ID_CURR TARGET NAME_CONTRACT_TYPE CODE_GENDER FLAG_OWN_CAR  
0 100002 1 Cash loans M N  
1 100003 0 Cash loans F N  
2 100004 0 Revolving loans M Y  
3 100006 0 Cash loans F N  
4 100007 0 Cash loans M N
```

```
In [5]: app_data.shape
```

```
Out[5]: (307511, 122)
```

```
In [6]: missing_data=app_data.count()/len(app_data)  
missing_data=(1-missing_data)*100  
missing_data.sort_values(ascending=False)
```

```
Out[6]: COMMONAREA_MEDI           69.872297
COMMONAREA_AVG                    69.872297
COMMONAREA_MODE                   69.872297
NONLIVINGAPARTMENTS_MODE         69.432963
NONLIVINGAPARTMENTS_AVG          69.432963
NONLIVINGAPARTMENTS_MEDI         69.432963
FONDKAPREMONT_MODE              68.386172
LIVINGAPARTMENTS_MODE            68.354953
LIVINGAPARTMENTS_AVG             68.354953
LIVINGAPARTMENTS_MEDI            68.354953
FLOORSMIN_AVG                    67.848630
FLOORSMIN_MODE                   67.848630
FLOORSMIN_MEDI                  67.848630
YEARS_BUILD_MEDI                 66.497784
YEARS_BUILD_MODE                 66.497784
YEARS_BUILD_AVG                  66.497784
OWN_CAR_AGE                      65.990810
LANDAREA_MEDI                     59.376738
LANDAREA_MODE                     59.376738
LANDAREA_AVG                      59.376738
BASEMENTAREA_MEDI                58.515956
BASEMENTAREA_AVG                  58.515956
BASEMENTAREA_MODE                 58.515956
EXT_SOURCE_1                       56.381073
NONLIVINGAREA_MODE                55.179164
NONLIVINGAREA_AVG                 55.179164
NONLIVINGAREA_MEDI                55.179164
ELEVATORS_MEDI                    53.295980
ELEVATORS_AVG                     53.295980
ELEVATORS_MODE                    53.295980
WALLSMATERIAL_MODE               50.840783
APARTMENTS_MEDI                  50.749729
APARTMENTS_AVG                    50.749729
APARTMENTS_MODE                   50.749729
ENTRANCES_MEDI                   50.348768
ENTRANCES_AVG                     50.348768
ENTRANCES_MODE                    50.348768
LIVINGAREA_AVG                    50.193326
LIVINGAREA_MODE                   50.193326
LIVINGAREA_MEDI                  50.193326
HOUSETYPE_MODE                     50.176091
FLOORSMAX_MODE                    49.760822
FLOORSMAX_MEDI                   49.760822
FLOORSMAX_AVG                     49.760822
YEARS_BEGINEXPLUATATION_MODE     48.781019
YEARS_BEGINEXPLUATATION_MEDI      48.781019
YEARS_BEGINEXPLUATATION_AVG       48.781019
TOTALAREA_MODE                     48.268517
EMERGENCYSTATE_MODE               47.398304
OCCUPATION_TYPE                   31.345545
EXT_SOURCE_3                        19.825307
AMT_REQ_CREDIT_BUREAU_HOUR        13.501631
AMT_REQ_CREDIT_BUREAU_DAY          13.501631
AMT_REQ_CREDIT_BUREAU_WEEK         13.501631
```

AMT_REQ_CREDIT_BUREAU_MON	13.501631
AMT_REQ_CREDIT_BUREAU_QRT	13.501631
AMT_REQ_CREDIT_BUREAU_YEAR	13.501631
NAME_TYPE_SUITE	0.420148
OBS_30_CNT_SOCIAL_CIRCLE	0.332021
DEF_30_CNT_SOCIAL_CIRCLE	0.332021
OBS_60_CNT_SOCIAL_CIRCLE	0.332021
DEF_60_CNT_SOCIAL_CIRCLE	0.332021
EXT_SOURCE_2	0.214626
AMT_GOODS_PRICE	0.090403
AMT_ANNUITY	0.003902
CNT_FAM_MEMBERS	0.000650
DAYSLAST_PHONE_CHANGE	0.000325
CNT_CHILDREN	0.000000
FLAG_DOCUMENT_8	0.000000
NAME_CONTRACT_TYPE	0.000000
CODE_GENDER	0.000000
FLAG_OWN_CAR	0.000000
FLAG_DOCUMENT_2	0.000000
FLAG_DOCUMENT_3	0.000000
FLAG_DOCUMENT_4	0.000000
FLAG_DOCUMENT_5	0.000000
FLAG_DOCUMENT_6	0.000000
FLAG_DOCUMENT_7	0.000000
FLAG_DOCUMENT_9	0.000000
FLAG_DOCUMENT_21	0.000000
FLAG_DOCUMENT_10	0.000000
FLAG_DOCUMENT_11	0.000000
FLAG_OWN_REALTY	0.000000
FLAG_DOCUMENT_13	0.000000
FLAG_DOCUMENT_14	0.000000
FLAG_DOCUMENT_15	0.000000
FLAG_DOCUMENT_16	0.000000
FLAG_DOCUMENT_17	0.000000
FLAG_DOCUMENT_18	0.000000
FLAG_DOCUMENT_19	0.000000
FLAG_DOCUMENT_20	0.000000
FLAG_DOCUMENT_12	0.000000
AMT_CREDIT	0.000000
AMT_INCOME_TOTAL	0.000000
FLAG_PHONE	0.000000
LIVE_CITY_NOT_WORK_CITY	0.000000
REG_CITY_NOT_WORK_CITY	0.000000
TARGET	0.000000
REG_CITY_NOT_LIVE_CITY	0.000000
LIVE_REGION_NOT_WORK_REGION	0.000000
REG_REGION_NOT_WORK_REGION	0.000000
REG_REGION_NOT_LIVE_REGION	0.000000
HOUR_APPR_PROCESS_START	0.000000
WEEKDAY_APPR_PROCESS_START	0.000000
REGION_RATING_CLIENT_W_CITY	0.000000
REGION_RATING_CLIENT	0.000000
FLAG_EMAIL	0.000000
FLAG_CONT_MOBILE	0.000000

```
ORGANIZATION_TYPE          0.000000
FLAG_WORK_PHONE             0.000000
FLAG_EMP_PHONE              0.000000
FLAG_MOBIL                  0.000000
DAYS_ID_PUBLISH             0.000000
DAYS_REGISTRATION            0.000000
DAYS_EMPLOYED                 0.000000
DAYS_BIRTH                   0.000000
REGION_POPULATION_RELATIVE   0.000000
NAME_HOUSING_TYPE             0.000000
NAME_FAMILY_STATUS              0.000000
NAME_EDUCATION_TYPE             0.000000
NAME_INCOME_TYPE                0.000000
SK_ID_CURR                     0.000000
dtype: float64
```

```
In [7]: appl_data=app_data.loc[:,missing_data<50]
appl_data.shape
```

```
Out[7]: (307511, 81)
```

```
In [8]: appl_data.head()
```

```
Out[8]: SK_ID_CURR  TARGET  NAME_CONTRACT_TYPE  CODE_GENDER  FLAG_OWN_CAR
0           100002      1        Cash loans       M           N
1           100003      0        Cash loans       F           N
2           100004      0    Revolving loans       M           Y
3           100006      0        Cash loans       F           N
4           100007      0        Cash loans       M           N
```

```
In [9]: #Checking for null values in each column
appl_data.isnull().sum()
```

```
Out[9]: SK_ID_CURR          0
TARGET          0
NAME_CONTRACT_TYPE      0
CODE_GENDER        0
FLAG_OWN_CAR        0
FLAG_OWN_REALTY      0
CNT_CHILDREN        0
AMT_INCOME_TOTAL      0
AMT_CREDIT          0
AMT_ANNUITY          12
AMT_GOODS_PRICE        278
NAME_TYPE_SUITE        1292
NAME_INCOME_TYPE        0
NAME_EDUCATION_TYPE      0
NAME_FAMILY_STATUS       0
NAME_HOUSING_TYPE        0
REGION_POPULATION_RELATIVE 0
DAYS_BIRTH          0
DAYS_EMPLOYED         0
DAYS_REGISTRATION       0
DAYS_ID_PUBLISH        0
FLAG_MOBIL          0
FLAG_EMP_PHONE         0
FLAG_WORK_PHONE        0
FLAG_CONT_MOBILE        0
FLAG_PHONE           0
FLAG_EMAIL           0
OCCUPATION_TYPE        96391
CNT_FAM_MEMBERS        2
REGION_RATING_CLIENT      0
REGION_RATING_CLIENT_W_CITY 0
WEEKDAY_APPR_PROCESS_START 0
HOUR_APPR_PROCESS_START    0
REG_REGION_NOT_LIVE_REGION 0
REG_REGION_NOT_WORK_REGION 0
LIVE_REGION_NOT_WORK_REGION 0
REG_CITY_NOT_LIVE_CITY      0
REG_CITY_NOT_WORK_CITY      0
LIVE_CITY_NOT_WORK_CITY      0
ORGANIZATION_TYPE        0
EXT_SOURCE_2            660
EXT_SOURCE_3            60965
YEARS_BEGINEXPLUATATION_AVG 150007
FLOORSMAX_AVG          153020
YEARS_BEGINEXPLUATATION_MODE 150007
FLOORSMAX_MODE          153020
YEARS_BEGINEXPLUATATION_MEDI 150007
FLOORSMAX_MEDI          153020
TOTALAREA_MODE          148431
EMERGENCYSTATE_MODE        145755
OBS_30_CNT_SOCIAL_CIRCLE 1021
DEF_30_CNT_SOCIAL_CIRCLE 1021
OBS_60_CNT_SOCIAL_CIRCLE 1021
DEF_60_CNT_SOCIAL_CIRCLE 1021
```

```
  DAYS_LAST_PHONE_CHANGE          1
  FLAG_DOCUMENT_2                 0
  FLAG_DOCUMENT_3                 0
  FLAG_DOCUMENT_4                 0
  FLAG_DOCUMENT_5                 0
  FLAG_DOCUMENT_6                 0
  FLAG_DOCUMENT_7                 0
  FLAG_DOCUMENT_8                 0
  FLAG_DOCUMENT_9                 0
  FLAG_DOCUMENT_10                0
  FLAG_DOCUMENT_11                0
  FLAG_DOCUMENT_12                0
  FLAG_DOCUMENT_13                0
  FLAG_DOCUMENT_14                0
  FLAG_DOCUMENT_15                0
  FLAG_DOCUMENT_16                0
  FLAG_DOCUMENT_17                0
  FLAG_DOCUMENT_18                0
  FLAG_DOCUMENT_19                0
  FLAG_DOCUMENT_20                0
  FLAG_DOCUMENT_21                0
  AMT_REQ_CREDIT_BUREAU_HOUR     41519
  AMT_REQ_CREDIT_BUREAU_DAY       41519
  AMT_REQ_CREDIT_BUREAU_WEEK      41519
  AMT_REQ_CREDIT_BUREAU_MON       41519
  AMT_REQ_CREDIT_BUREAU_QRT       41519
  AMT_REQ_CREDIT_BUREAU_YEAR      41519
dtype: int64
```

```
In [10]: #Checking for datatypes of variables in dataframe
appl_data.dtypes
```

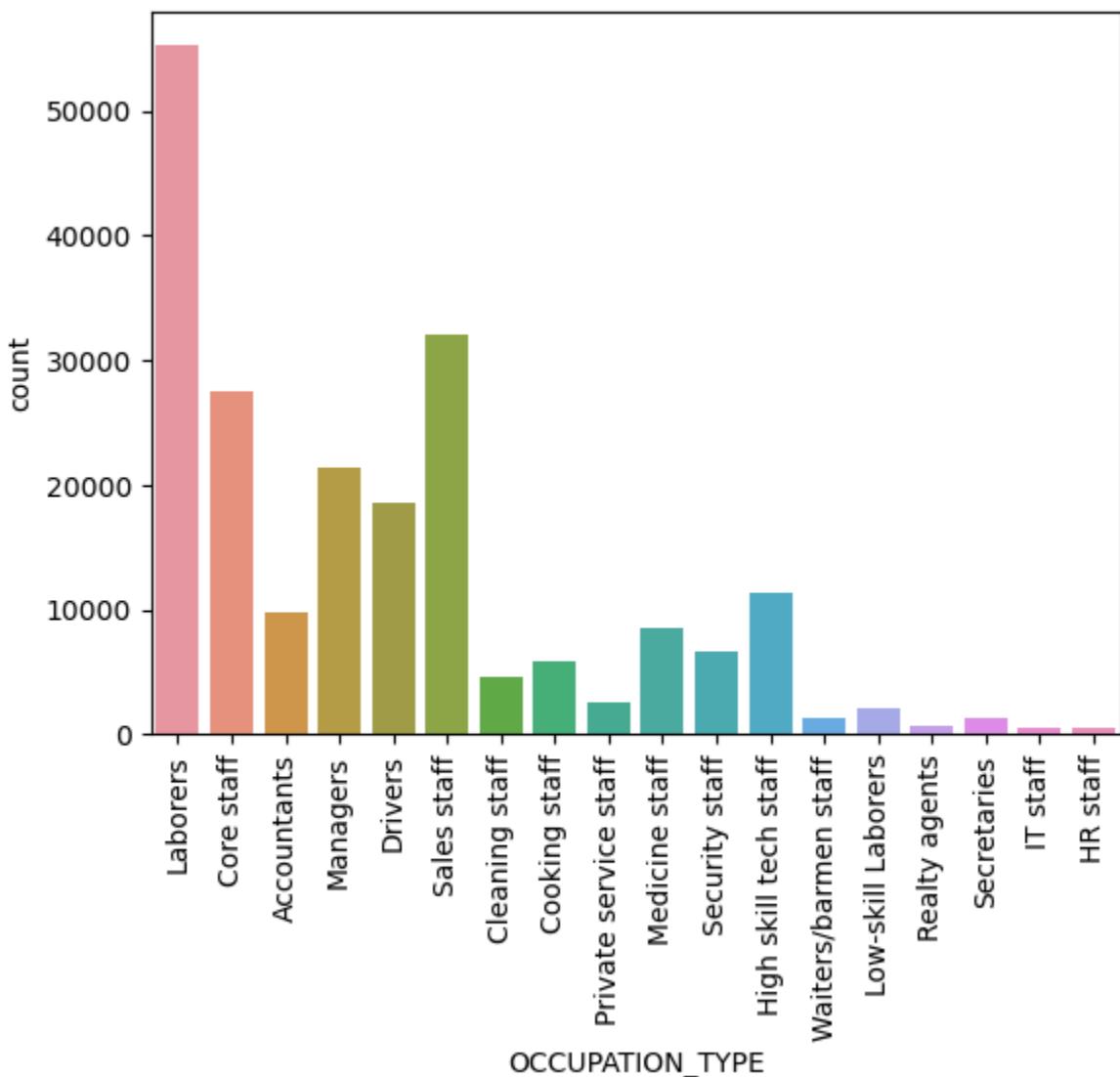
```
Out[10]: SK_ID_CURR                      int64
TARGET                         int64
NAME_CONTRACT_TYPE            object
CODE_GENDER                     object
FLAG_OWN_CAR                    object
FLAG_OWN_REALTY                  object
CNT_CHILDREN                      int64
AMT_INCOME_TOTAL                  float64
AMT_CREDIT                        float64
AMT_ANNUITY                        float64
AMT_GOODS_PRICE                    float64
NAME_TYPE_SUITE                   object
NAME_INCOME_TYPE                  object
NAME_EDUCATION_TYPE                object
NAME_FAMILY_STATUS                 object
NAME_HOUSING_TYPE                  object
REGION_POPULATION_RELATIVE        float64
DAYS_BIRTH                         int64
DAYS_EMPLOYED                      int64
DAYS_REGISTRATION                  float64
DAYS_ID_PUBLISH                    int64
FLAG_MOBIL                          int64
FLAG_EMP_PHONE                      int64
FLAG_WORK_PHONE                     int64
FLAG_CONT_MOBILE                     int64
FLAG_PHONE                          int64
FLAG_EMAIL                           int64
OCCUPATION_TYPE                   object
CNT_FAM_MEMBERS                     float64
REGION_RATING_CLIENT                int64
REGION_RATING_CLIENT_W_CITY          int64
WEEKDAY_APPR_PROCESS_START           object
HOUR_APPR_PROCESS_START               int64
REG_REGION_NOT_LIVE_REGION             int64
REG_REGION_NOT_WORK_REGION              int64
LIVE_REGION_NOT_WORK_REGION             int64
REG_CITY_NOT_LIVE_CITY                int64
REG_CITY_NOT_WORK_CITY                 int64
LIVE_CITY_NOT_WORK_CITY                 int64
ORGANIZATION_TYPE                   object
EXT_SOURCE_2                         float64
EXT_SOURCE_3                         float64
YEARS_BEGINEXPLUATATION_AVG           float64
FLOORSMAX_AVG                       float64
YEARS_BEGINEXPLUATATION_MODE           float64
FLOORSMAX_MODE                      float64
YEARS_BEGINEXPLUATATION_MEDI             float64
FLOORSMAX_MEDI                      float64
TOTALAREA_MODE                      float64
EMERGENCYSTATE_MODE                  object
OBS_30_CNT_SOCIAL_CIRCLE              float64
DEF_30_CNT_SOCIAL_CIRCLE                float64
OBS_60_CNT_SOCIAL_CIRCLE              float64
DEF_60_CNT_SOCIAL_CIRCLE                float64
```

```
    DAYS_LAST_PHONE_CHANGE           float64
    FLAG_DOCUMENT_2                  int64
    FLAG_DOCUMENT_3                  int64
    FLAG_DOCUMENT_4                  int64
    FLAG_DOCUMENT_5                  int64
    FLAG_DOCUMENT_6                  int64
    FLAG_DOCUMENT_7                  int64
    FLAG_DOCUMENT_8                  int64
    FLAG_DOCUMENT_9                  int64
    FLAG_DOCUMENT_10                 int64
    FLAG_DOCUMENT_11                 int64
    FLAG_DOCUMENT_12                 int64
    FLAG_DOCUMENT_13                 int64
    FLAG_DOCUMENT_14                 int64
    FLAG_DOCUMENT_15                 int64
    FLAG_DOCUMENT_16                 int64
    FLAG_DOCUMENT_17                 int64
    FLAG_DOCUMENT_18                 int64
    FLAG_DOCUMENT_19                 int64
    FLAG_DOCUMENT_20                 int64
    FLAG_DOCUMENT_21                 int64
    AMT_REQ_CREDIT_BUREAU_HOUR      float64
    AMT_REQ_CREDIT_BUREAU_DAY        float64
    AMT_REQ_CREDIT_BUREAU_WEEK       float64
    AMT_REQ_CREDIT_BUREAU_MON        float64
    AMT_REQ_CREDIT_BUREAU_QRT        float64
    AMT_REQ_CREDIT_BUREAU_YEAR       float64
dtype: object
```

```
In [11]: sns.countplot(x='OCCUPATION_TYPE', data=appl_data)
plt.xticks(rotation=90)
plt.show()
```

```
Out[11]: <Axes: xlabel='OCCUPATION_TYPE', ylabel='count'>
```

```
Out[11]: (array([ 0,  1,  2,  3,  4,  5,  6,  7,  8,  9, 10, 11, 12, 13, 14, 15, 16,
   17]),
 [Text(0, 0, 'Laborers'),
  Text(1, 0, 'Core staff'),
  Text(2, 0, 'Accountants'),
  Text(3, 0, 'Managers'),
  Text(4, 0, 'Drivers'),
  Text(5, 0, 'Sales staff'),
  Text(6, 0, 'Cleaning staff'),
  Text(7, 0, 'Cooking staff'),
  Text(8, 0, 'Private service staff'),
  Text(9, 0, 'Medicine staff'),
  Text(10, 0, 'Security staff'),
  Text(11, 0, 'High skill tech staff'),
  Text(12, 0, 'Waiters/barmen staff'),
  Text(13, 0, 'Low-skill Laborers'),
  Text(14, 0, 'Realty agents'),
  Text(15, 0, 'Secretaries'),
  Text(16, 0, 'IT staff'),
  Text(17, 0, 'HR staff')])
```



```
In [12]: mean=appl_data['FLOORSMAX_AVG'].mean()
```

```
In [13]: appl_data['FLOORSMAX_AVG'].median()
```

```
Out[13]: 0.1667
```

```
In [14]: appl_data['FLOORSMAX_AVG'].mode()
```

```
Out[14]: 0    0.1667  
Name: FLOORSMAX_AVG, dtype: float64
```

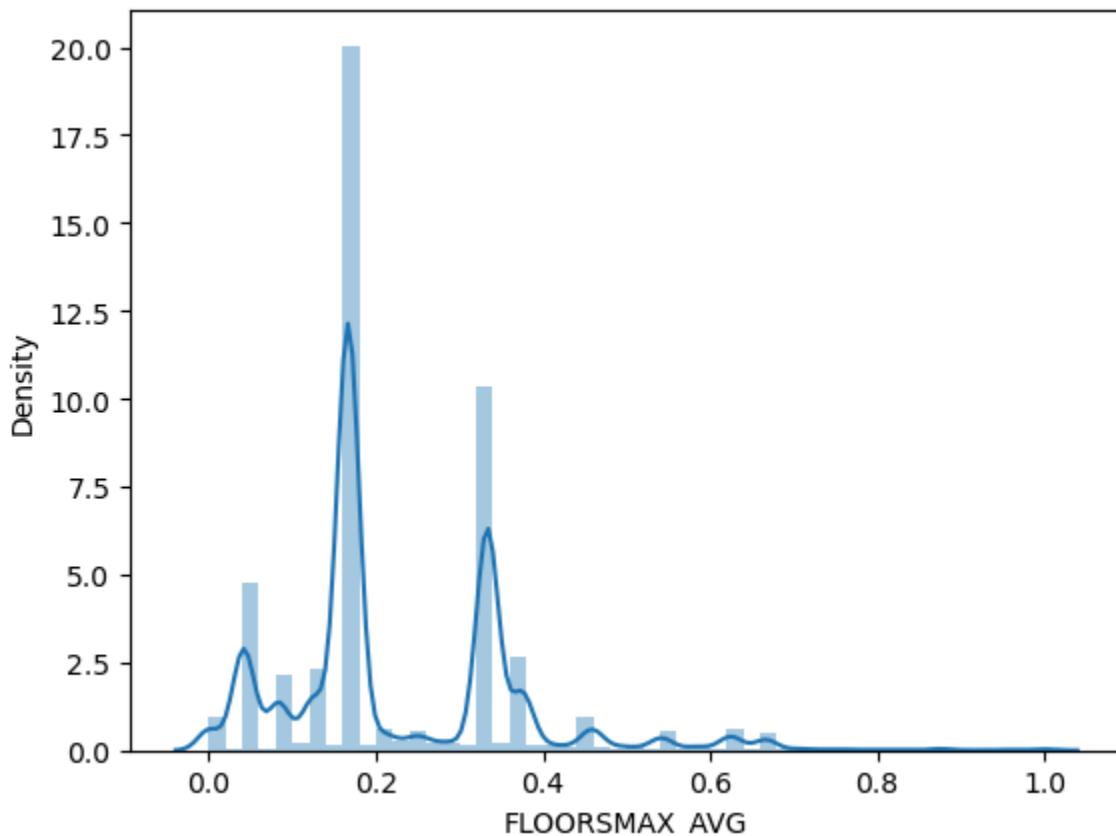
```
In [15]: appl_data['FLOORSMAX_AVG'].isnull().sum()
```

```
Out[15]: 153020
```

```
In [16]: import warnings  
warnings.filterwarnings('ignore')
```

```
In [17]: sns.distplot(appl_data['FLOORSMAX_AVG'])  
plt.show()
```

```
Out[17]: <Axes: xlabel='FLOORSMAX_AVG', ylabel='Density'>
```



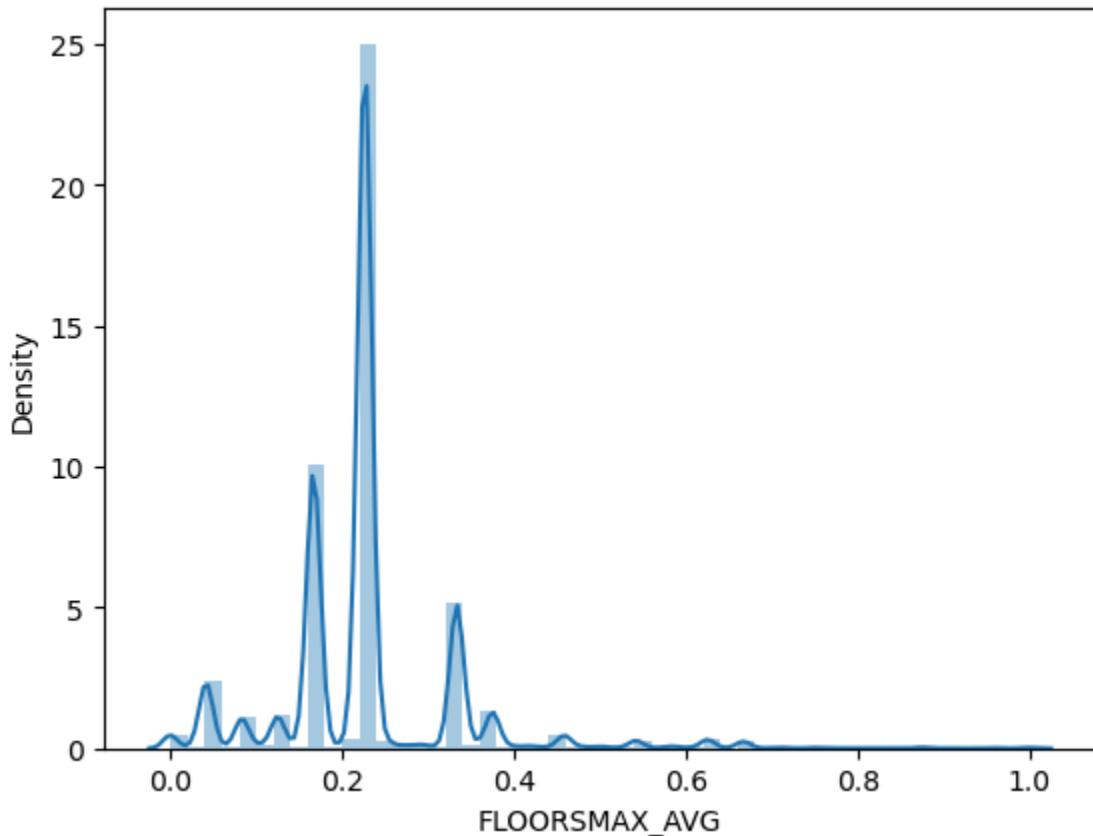
```
In [18]: appl_data.FLOORSMAX_AVG.fillna(mean,inplace=True)
```

```
In [19]: appl_data.FLOORSMAX_AVG.isnull().sum()
```

```
Out[19]: 0
```

```
In [20]: sns.distplot(appl_data['FLOORSMAX_AVG'])
```

```
Out[20]: <Axes: xlabel='FLOORSMAX_AVG', ylabel='Density'>
```



```
In [21]: appl_data.AMT_REQ_CREDIT_BUREAU_QRT.value_counts()
```

```
Out[21]: AMT_REQ_CREDIT_BUREAU_QRT
0.0      215417
1.0      33862
2.0      14412
3.0      1717
4.0       476
5.0        64
6.0        28
8.0         7
7.0         7
261.0      1
19.0        1
Name: count, dtype: int64
```

```
In [22]: appl_data['AMT_REQ_CREDIT_BUREAU_QRT'].mean()
```

```
Out[22]: 0.26547414959848414
```

```
In [23]: appl_data['AMT_REQ_CREDIT_BUREAU_QRT'].mode()
```

```
Out[23]: 0      0.0  
         Name: AMT_REQ_CREDIT_BUREAU_QRT, dtype: float64
```

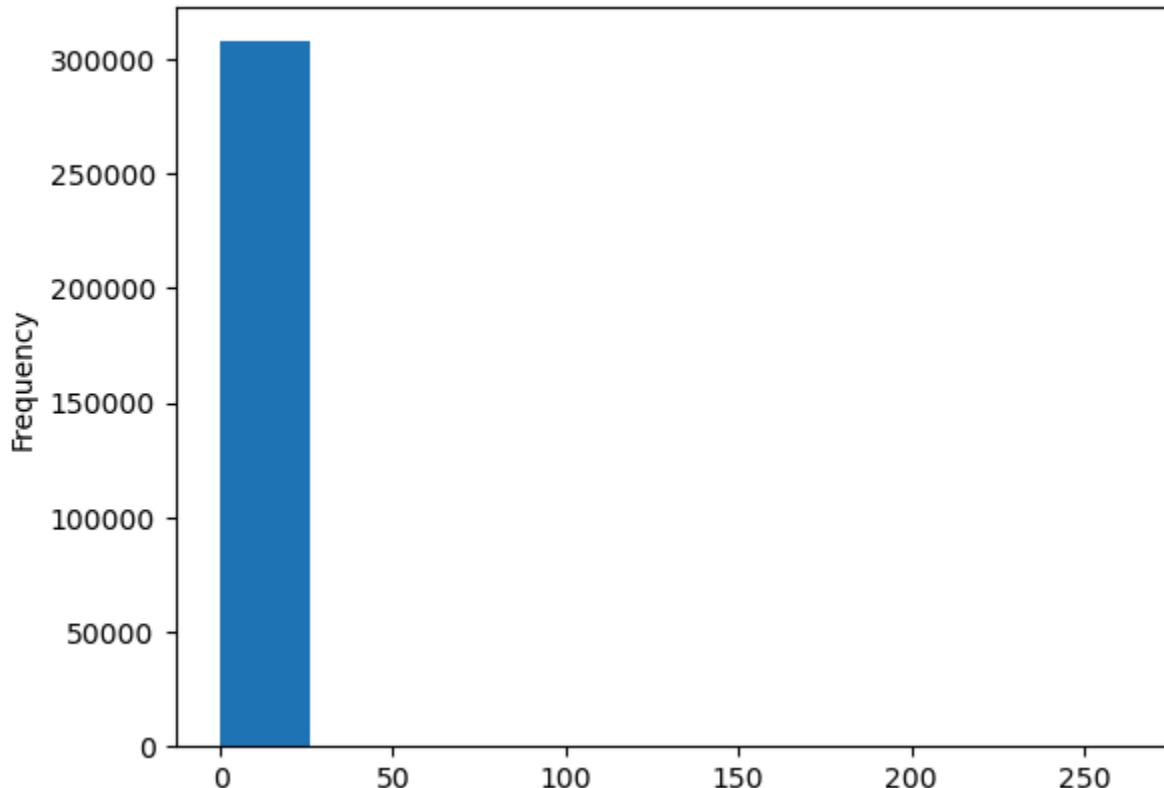
```
In [24]: appl_data['AMT_REQ_CREDIT_BUREAU_QRT']=appl_data['AMT_REQ_CREDIT_BUREAU_QRT'].
```

```
In [25]: appl_data['AMT_REQ_CREDIT_BUREAU_QRT'].isnull().sum()
```

```
Out[25]: 0
```

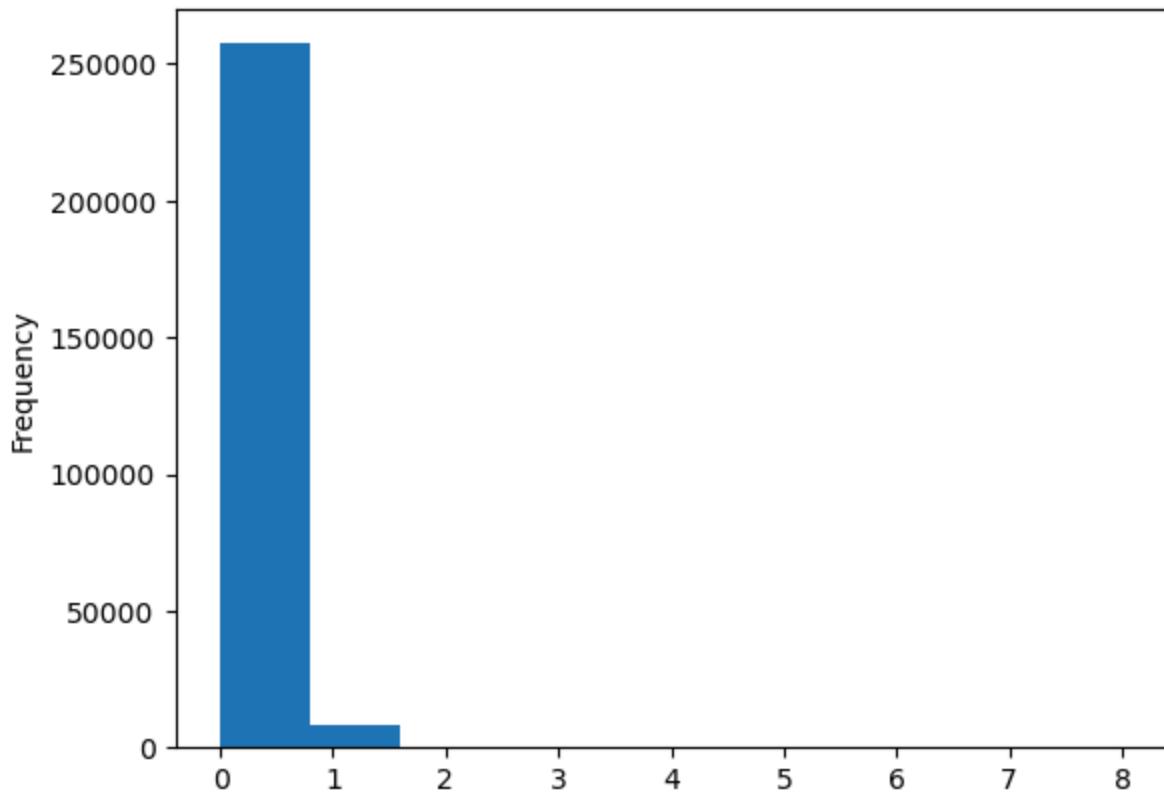
```
In [26]: appl_data['AMT_REQ_CREDIT_BUREAU_QRT'].plot.hist()
```

```
Out[26]: <Axes: ylabel='Frequency'>
```



```
In [27]: num_unique_values=len(appl_data.AMT_REQ_CREDIT_BUREAU_WEEK.unique())  
appl_data['AMT_REQ_CREDIT_BUREAU_WEEK'].plot.hist(bins=num_unique_values)
```

```
Out[27]: <Axes: ylabel='Frequency'>
```



Since in this graph most of the values are zeros, this column will not cause much variation in analysis and so we will not use it for analysis.

```
In [28]: appl_data.isnull().sum()
```

```
Out[28]: SK_ID_CURR          0
TARGET           0
NAME_CONTRACT_TYPE 0
CODE_GENDER       0
FLAG_OWN_CAR      0
FLAG_OWN_REALTY   0
CNT_CHILDREN      0
AMT_INCOME_TOTAL  0
AMT_CREDIT         0
AMT_ANNUITY        12
AMT_GOODS_PRICE    278
NAME_TYPE_SUITE    1292
NAME_INCOME_TYPE   0
NAME_EDUCATION_TYPE 0
NAME_FAMILY_STATUS  0
NAME_HOUSING_TYPE  0
REGION_POPULATION_RELATIVE 0
DAYS_BIRTH         0
DAYS_EMPLOYED      0
DAYS_REGISTRATION   0
DAYS_ID_PUBLISH    0
FLAG_MOBIL          0
FLAG_EMP_PHONE      0
FLAG_WORK_PHONE     0
FLAG_CONT_MOBILE    0
FLAG_PHONE          0
FLAG_EMAIL          0
OCCUPATION_TYPE     96391
CNT_FAM_MEMBERS     2
REGION_RATING_CLIENT 0
REGION_RATING_CLIENT_W_CITY 0
WEEKDAY_APPR_PROCESS_START 0
HOUR_APPR_PROCESS_START 0
REG_REGION_NOT_LIVE_REGION 0
REG_REGION_NOT_WORK_REGION 0
LIVE_REGION_NOT_WORK_REGION 0
REG_CITY_NOT_LIVE_CITY 0
REG_CITY_NOT_WORK_CITY 0
LIVE_CITY_NOT_WORK_CITY 0
ORGANIZATION_TYPE    0
EXT_SOURCE_2          660
EXT_SOURCE_3          60965
YEARS_BEGINEXPLUATATION_AVG 150007
FLOORSMAX_AVG        0
YEARS_BEGINEXPLUATATION_MODE 150007
FLOORSMAX_MODE        153020
YEARS_BEGINEXPLUATATION_MEDI 150007
FLOORSMAX_MEDI        153020
TOTALAREA_MODE        148431
EMERGENCYSTATE_MODE   145755
OBS_30_CNT_SOCIAL_CIRCLE 1021
DEF_30_CNT_SOCIAL_CIRCLE 1021
OBS_60_CNT_SOCIAL_CIRCLE 1021
DEF_60_CNT_SOCIAL_CIRCLE 1021
```

```
DAYSLASTPHONECHANGE 1
FLAG_DOCUMENT_2 0
FLAG_DOCUMENT_3 0
FLAG_DOCUMENT_4 0
FLAG_DOCUMENT_5 0
FLAG_DOCUMENT_6 0
FLAG_DOCUMENT_7 0
FLAG_DOCUMENT_8 0
FLAG_DOCUMENT_9 0
FLAG_DOCUMENT_10 0
FLAG_DOCUMENT_11 0
FLAG_DOCUMENT_12 0
FLAG_DOCUMENT_13 0
FLAG_DOCUMENT_14 0
FLAG_DOCUMENT_15 0
FLAG_DOCUMENT_16 0
FLAG_DOCUMENT_17 0
FLAG_DOCUMENT_18 0
FLAG_DOCUMENT_19 0
FLAG_DOCUMENT_20 0
FLAG_DOCUMENT_21 0
AMT_REQ_CREDIT_BUREAU_HOUR 41519
AMT_REQ_CREDIT_BUREAU_DAY 41519
AMT_REQ_CREDIT_BUREAU_WEEK 41519
AMT_REQ_CREDIT_BUREAU_MON 41519
AMT_REQ_CREDIT_BUREAU_QRT 0
AMT_REQ_CREDIT_BUREAU_YEAR 41519
dtype: int64
```

```
In [29]: appl_data['OBS_30_CNT_SOCIAL_CIRCLE'].describe()
```

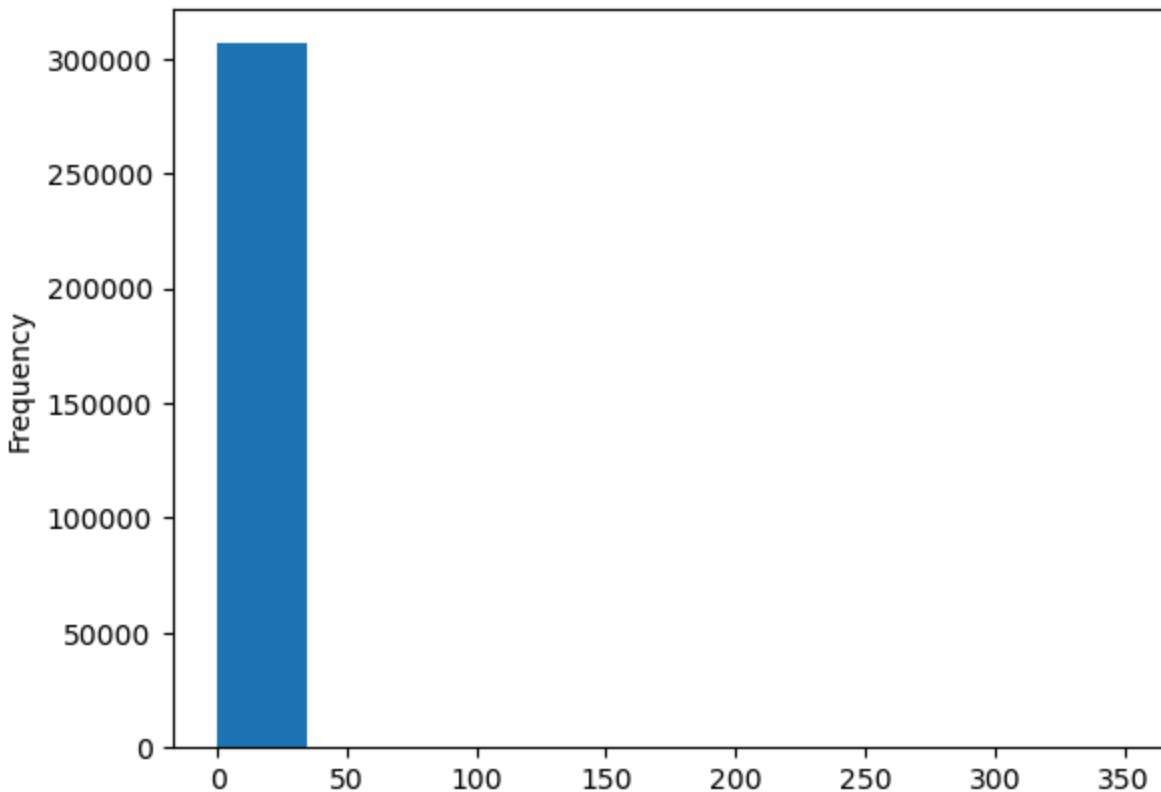
```
Out[29]: count    306490.000000
mean      1.422245
std       2.400989
min       0.000000
25%      0.000000
50%      0.000000
75%      2.000000
max      348.000000
Name: OBS_30_CNT_SOCIAL_CIRCLE, dtype: float64
```

```
In [30]: appl_data['OBS_30_CNT_SOCIAL_CIRCLE'].mean()
```

```
Out[30]: 1.4222454239942575
```

```
In [31]: appl_data['OBS_30_CNT_SOCIAL_CIRCLE'].plot.hist()
```

```
Out[31]: <Axes: ylabel='Frequency'>
```



```
In [32]: appl_data['DEF_60_CNT_SOCIAL_CIRCLE'].isnull().sum()
```

```
Out[32]: 1021
```

```
In [33]: mode=appl_data['DEF_60_CNT_SOCIAL_CIRCLE'].mode()  
mode
```

```
Out[33]: 0    0.0  
Name: DEF_60_CNT_SOCIAL_CIRCLE, dtype: float64
```

```
In [34]: appl_data['DEF_60_CNT_SOCIAL_CIRCLE'].mean()
```

```
Out[34]: 0.10004894123788705
```

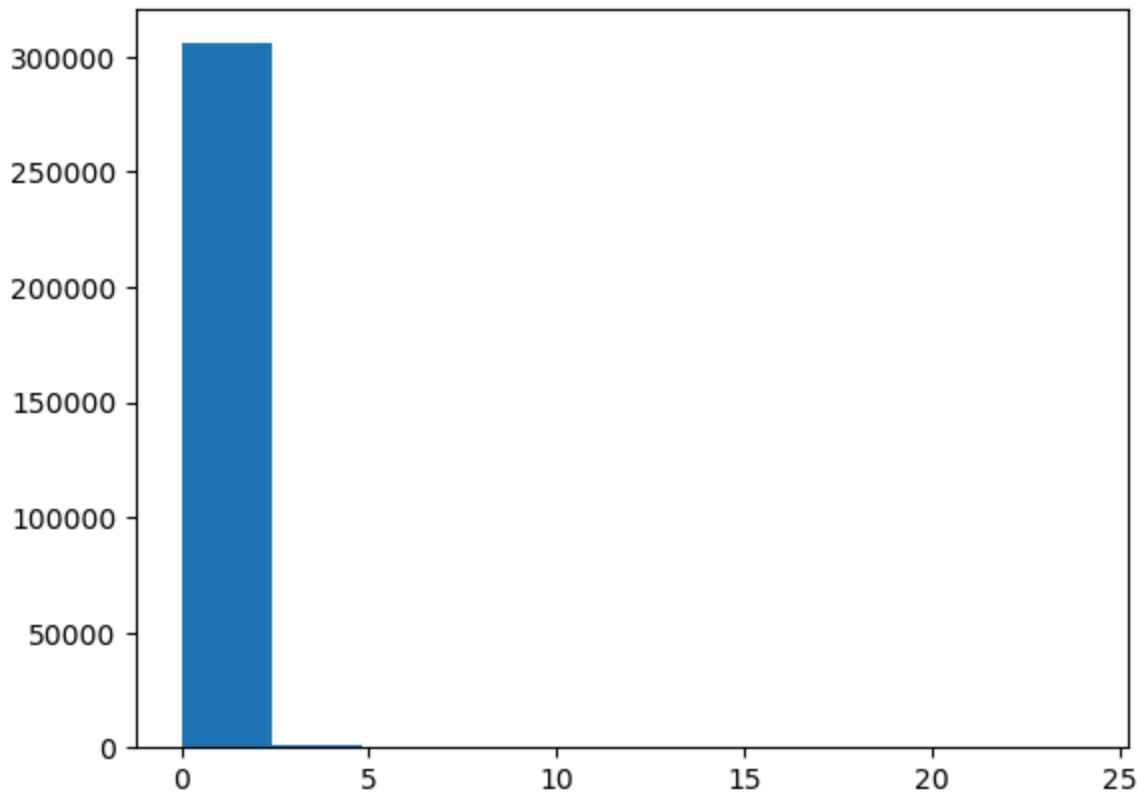
```
In [35]: appl_data['DEF_60_CNT_SOCIAL_CIRCLE'].fillna(mode,inplace=True)
```

```
In [36]: appl_data['DEF_60_CNT_SOCIAL_CIRCLE'].isnull().sum()
```

```
Out[36]: 1021
```

```
In [37]: plt.hist(appl_data['DEF_60_CNT_SOCIAL_CIRCLE'])
```

```
Out[37]: (array([3.05732e+05, 7.33000e+02, 2.40000e+01, 0.00000e+00, 0.00000e+00,  
0.00000e+00, 0.00000e+00, 0.00000e+00, 0.00000e+00, 1.00000e+00]),  
array([ 0. ,  2.4,  4.8,  7.2,  9.6, 12. , 14.4, 16.8, 19.2, 21.6, 24. ]),  
<BarContainer object of 10 artists>)
```



```
In [38]: nulls = appl_data.isnull().sum()  
nulls=nulls > 0
```

```
Out[38]: AMT_ANNUITY           12  
AMT_GOODS_PRICE            278  
NAME_TYPE_SUITE          1292  
OCCUPATION_TYPE         96391  
CNT_FAM_MEMBERS             2  
EXT_SOURCE_2              660  
EXT_SOURCE_3              60965  
YEARS_BEGINEXPLUATATION_AVG 150007  
YEARS_BEGINEXPLUATATION_MODE 150007  
FLOORSMAX_MODE           153020  
YEARS_BEGINEXPLUATATION_MEDI 150007  
FLOORSMAX_MEDI            153020  
TOTALAREA_MODE            148431  
EMERGENCYSTATE_MODE        145755  
OBS_30_CNT_SOCIAL_CIRCLE    1021  
DEF_30_CNT_SOCIAL_CIRCLE    1021  
OBS_60_CNT_SOCIAL_CIRCLE    1021  
DEF_60_CNT_SOCIAL_CIRCLE    1021  
DAYS_LAST_PHONE_CHANGE      1  
AMT_REQ_CREDIT_BUREAU_HOUR 41519  
AMT_REQ_CREDIT_BUREAU_DAY   41519  
AMT_REQ_CREDIT_BUREAU_WEEK  41519  
AMT_REQ_CREDIT_BUREAU_MON   41519  
AMT_REQ_CREDIT_BUREAU_YEAR  41519  
dtype: int64
```

```
In [39]: prev_app=pd.read_csv(r"C:\Users\nikhi\Data Science\EDA Assignment\previous_app
```

```
In [40]: prev_app.head()
```

```
Out[40]: SK_ID_PREV SK_ID_CURR NAME_CONTRACT_TYPE AMT_ANNUITY AMT_APPLI
```

0	2030495	271877	Consumer loans	1730.430	:
1	2802425	108129	Cash loans	25188.615	60
2	2523466	122040	Cash loans	15060.735	12
3	2819243	176158	Cash loans	47041.335	4!
4	1784265	202054	Cash loans	31924.395	3:

```
In [41]: prev_app.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1670214 entries, 0 to 1670213
Data columns (total 37 columns):
 #   Column           Non-Null Count   Dtype  
--- 
 0   SK_ID_PREV       1670214 non-null   int64  
 1   SK_ID_CURR       1670214 non-null   int64  
 2   NAME_CONTRACT_TYPE 1670214 non-null   object  
 3   AMT_ANNUITY      1297979 non-null   float64 
 4   AMT_APPLICATION  1670214 non-null   float64 
 5   AMT_CREDIT        1670213 non-null   float64 
 6   AMT_DOWN_PAYMENT  774370 non-null   float64 
 7   AMT_GOODS_PRICE   1284699 non-null   float64 
 8   WEEKDAY_APPR_PROCESS_START 1670214 non-null   object  
 9   HOUR_APPR_PROCESS_START 1670214 non-null   int64  
 10  FLAG_LAST_APPL_PER_CONTRACT 1670214 non-null   object  
 11  NFLAG_LAST_APPL_IN_DAY    1670214 non-null   int64  
 12  RATE_DOWN_PAYMENT     774370 non-null   float64 
 13  RATE_INTEREST_PRIMARY 5951 non-null    float64 
 14  RATE_INTEREST_PRIVILEGED 5951 non-null   float64 
 15  NAME_CASH_LOAN_PURPOSE 1670214 non-null   object  
 16  NAME_CONTRACT_STATUS  1670214 non-null   object  
 17  DAYS_DECISION       1670214 non-null   int64  
 18  NAME_PAYMENT_TYPE   1670214 non-null   object  
 19  CODE_REJECT_REASON  1670214 non-null   object  
 20  NAME_TYPE_SUITE     849809 non-null   object  
 21  NAME_CLIENT_TYPE    1670214 non-null   object  
 22  NAME_GOODS_CATEGORY 1670214 non-null   object  
 23  NAME_PORTFOLIO      1670214 non-null   object  
 24  NAME_PRODUCT_TYPE   1670214 non-null   object  
 25  CHANNEL_TYPE        1670214 non-null   object  
 26  SELLERPLACE_AREA    1670214 non-null   int64  
 27  NAME_SELLER_INDUSTRY 1670214 non-null   object  
 28  CNT_PAYMENT         1297984 non-null   float64 
 29  NAME_YIELD_GROUP   1670214 non-null   object  
 30  PRODUCT_COMBINATION 1669868 non-null   object  
 31  DAYS_FIRST_DRAWING 997149 non-null   float64 
 32  DAYS_FIRST_DUE     997149 non-null   float64 
 33  DAYS_LAST_DUE_1ST_VERSION 997149 non-null   float64 
 34  DAYS_LAST_DUE      997149 non-null   float64 
 35  DAYS_TERMINATION   997149 non-null   float64 
 36  NFLAG_INSURED_ON_APPROVAL 997149 non-null   float64 
dtypes: float64(15), int64(6), object(16)
memory usage: 471.5+ MB
```

In [42]: `prev_app.shape`

Out[42]: (1670214, 37)

In [43]: `prev_app.isnull().sum()`

```
Out[43]: SK_ID_PREV          0  
SK_ID_CURR           0  
NAME_CONTRACT_TYPE    0  
AMT_ANNUITY          372235  
AMT_APPLICATION      0  
AMT_CREDIT            1  
AMT_DOWN_PAYMENT     895844  
AMT_GOODS_PRICE       385515  
WEEKDAY_APPR_PROCESS_START 0  
HOUR_APPR_PROCESS_START 0  
FLAG_LAST_APPL_PER_CONTRACT 0  
NFLAG_LAST_APPL_IN_DAY 0  
RATE_DOWN_PAYMENT    895844  
RATE_INTEREST_PRIMARY 1664263  
RATE_INTEREST_PRIVILEGED 1664263  
NAME_CASH_LOAN_PURPOSE 0  
NAME_CONTRACT_STATUS   0  
DAYS_DECISION         0  
NAME_PAYMENT_TYPE     0  
CODE_REJECT_REASON    0  
NAME_TYPE_SUITE        820405  
NAME_CLIENT_TYPE       0  
NAME_GOODS_CATEGORY    0  
NAME_PORTFOLIO         0  
NAME_PRODUCT_TYPE      0  
CHANNEL_TYPE           0  
SELLERPLACE_AREA       0  
NAME_SELLER_INDUSTRY   0  
CNT_PAYMENT           372230  
NAME_YIELD_GROUP       0  
PRODUCT_COMBINATION    346  
DAYS_FIRST_DRAWING    673065  
DAYS_FIRST_DUE         673065  
DAYS_LAST_DUE_1ST_VERSION 673065  
DAYS_LAST_DUE          673065  
DAYS_TERMINATION       673065  
NFLAG_INSURED_ON_APPROVAL 673065  
dtype: int64
```

```
In [44]: prev_app.PRODUCT_COMBINATION.isnull().sum()
```

```
Out[44]: 346
```

```
In [45]: float(100*346/1670214)
```

```
Out[45]: 0.02071590826085759
```

```
In [46]: prev_appl=prev_app[~prev_app.PRODUCT_COMBINATION.isnull()].copy()
```

```
In [47]: prev_appl.shape
```

```
Out[47]: (1669868, 37)
```

```
In [48]: prev_appl['AMT_GOODS_PRICE']=prev_appl.groupby('PRODUCT_COMBINATION')['AMT_GOC  
Out[48]: 1638.0
```

```
In [49]: dp=prev_appl.AMT_DOWN_PAYMENT.median()  
dp
```

```
Out[49]: 1638.0
```

```
In [50]: prev_appl.AMT_DOWN_PAYMENT.fillna(dp,inplace=True)
```

```
In [51]: rate_dp=prev_appl.RATE_DOWN_PAYMENT.median()  
rate_dp
```

```
Out[51]: 0.051605084707529295
```

```
In [52]: prev_appl.RATE_DOWN_PAYMENT.fillna(rate_dp,inplace=True)
```

```
In [53]: prev_appl=prev_appl[-prev_appl.AMT_CREDIT.isnull()]
```

```
In [54]: prev_appl.isnull().sum()
```

```
Out[54]: SK_ID_PREV          0  
SK_ID_CURR           0  
NAME_CONTRACT_TYPE    0  
AMT_ANNUITY          371889  
AMT_APPLICATION      0  
AMT_CREDIT            0  
AMT_DOWN_PAYMENT      0  
AMT_GOODS_PRICE        0  
WEEKDAY_APPR_PROCESS_START 0  
HOUR_APPR_PROCESS_START 0  
FLAG_LAST_APPL_PER_CONTRACT 0  
NFLAG_LAST_APPL_IN_DAY 0  
RATE_DOWN_PAYMENT      0  
RATE_INTEREST_PRIMARY   1663916  
RATE_INTEREST_PRIVILEGED 1663916  
NAME_CASH_LOAN_PURPOSE 0  
NAME_CONTRACT_STATUS    0  
DAYS_DECISION          0  
NAME_PAYMENT_TYPE       0  
CODE_REJECT_REASON      0  
NAME_TYPE_SUITE         820058  
NAME_CLIENT_TYPE        0  
NAME_GOODS_CATEGORY      0  
NAME_PORTFOLIO          0  
NAME_PRODUCT_TYPE        0  
CHANNEL_TYPE            0  
SELLERPLACE_AREA        0  
NAME_SELLER_INDUSTRY    0  
CNT_PAYMENT             371884  
NAME_YIELD_GROUP         0  
PRODUCT_COMBINATION      0  
DAYS_FIRST_DRAWING     672718  
DAYS_FIRST_DUE          672718  
DAYS_LAST_DUE_1ST_VERSION 672718  
DAYS_LAST_DUE            672718  
DAYS_TERMINATION         672718  
NFLAG_INSURED_ON_APPROVAL 672718  
dtype: int64
```

Analysing Target column

```
In [55]: appl_data['TARGET'].head()
```

```
Out[55]: 0    1  
1    0  
2    0  
3    0  
4    0  
Name: TARGET, dtype: int64
```

```
In [56]: appl_data['TARGET'].isnull().sum()
```

```
Out[56]: 0
```

```
In [57]: appl_data['TARGET'].unique()
```

```
Out[57]: array([1, 0], dtype=int64)
```

```
In [58]: appl_data['TARGET'].value_counts()[0]
```

```
Out[58]: 282686
```

```
In [59]: appl_data['TARGET'].value_counts()[1]
```

```
Out[59]: 24825
```

Imbalance Ratio

```
In [60]: r=appl_data['TARGET'].value_counts()[0]/appl_data['TARGET'].value_counts()[1]  
r
```

```
Out[60]: 11.387150050352467
```

```
In [61]: train_0 = appl_data.loc[appl_data['TARGET'] == 0]  
train_1 = appl_data.loc[appl_data['TARGET'] == 1]  
train_0  
train_1
```

```
Out[61]:
```

	SK_ID_CURR	TARGET	NAME_CONTRACT_TYPE	CODE_GENDER	FLAG_OW
1	100003	0	Cash loans	F	
2	100004	0	Revolving loans	M	
3	100006	0	Cash loans	F	
4	100007	0	Cash loans	M	
5	100008	0	Cash loans	M	
...
307505	456249	0	Cash loans	F	
307506	456251	0	Cash loans	M	
307507	456252	0	Cash loans	F	
307508	456253	0	Cash loans	F	
307510	456255	0	Cash loans	F	

282686 rows × 81 columns

Out[61]:

	SK_ID_CURR	TARGET	NAME_CONTRACT_TYPE	CODE_GENDER	FLAG_OWN_CAR
0	100002	1	Cash loans	M	
26	100031	1	Cash loans	F	
40	100047	1	Cash loans	M	
42	100049	1	Cash loans	F	
81	100096	1	Cash loans	F	
...
307448	456186	1	Cash loans	M	
307475	456215	1	Cash loans	F	
307481	456225	1	Cash loans	M	
307489	456233	1	Cash loans	F	
307509	456254	1	Cash loans	F	

24825 rows × 81 columns

In [62]:

```
cat_col = list(set(appl_data.columns) - set(appl_data.describe().columns))
cat_col
```

Out[62]:

```
['FLAG_OWN_CAR',
 'NAME_EDUCATION_TYPE',
 'NAME_CONTRACT_TYPE',
 'WEEKDAY_APPR_PROCESS_START',
 'FLAG_OWN_REALTY',
 'EMERGENCYSTATE_MODE',
 'CODE_GENDER',
 'NAME_TYPE_SUITE',
 'ORGANIZATION_TYPE',
 'NAME_HOUSING_TYPE',
 'NAME_FAMILY_STATUS',
 'NAME_INCOME_TYPE',
 'OCCUPATION_TYPE']
```

In [63]:

```
num_col = appl_data.describe().columns
num_col
```

```
Out[63]: Index(['SK_ID_CURR', 'TARGET', 'CNT_CHILDREN', 'AMT_INCOME_TOTAL',
       'AMT_CREDIT', 'AMT_ANNUITY', 'AMT_GOODS_PRICE',
       'REGION_POPULATION_RELATIVE', 'DAYS_BIRTH', 'DAYS_EMPLOYED',
       'DAYS_REGISTRATION', 'DAYS_ID_PUBLISH', 'FLAG_MOBIL', 'FLAG_EMP_PHONE',
       'FLAG_WORK_PHONE', 'FLAG_CONT_MOBILE', 'FLAG_PHONE', 'FLAG_EMAIL',
       'CNT_FAM_MEMBERS', 'REGION_RATING_CLIENT',
       'REGION_RATING_CLIENT_W_CITY', 'HOUR_APPR_PROCESS_START',
       'REG_REGION_NOT_LIVE_REGION', 'REG_REGION_NOT_WORK_REGION',
       'LIVE_REGION_NOT_WORK_REGION', 'REG_CITY_NOT_LIVE_CITY',
       'REG_CITY_NOT_WORK_CITY', 'LIVE_CITY_NOT_WORK_CITY', 'EXT_SOURCE_2',
       'EXT_SOURCE_3', 'YEARS_BEGINEXPLUATATION_AVG', 'FLOORSMAX_AVG',
       'YEARS_BEGINEXPLUATATION_MODE', 'FLOORSMAX_MODE',
       'YEARS_BEGINEXPLUATATION_MEDI', 'FLOORSMAX_MEDI', 'TOTALAREA_MODE',
       'OBS_30_CNT_SOCIAL_CIRCLE', 'DEF_30_CNT_SOCIAL_CIRCLE',
       'OBS_60_CNT_SOCIAL_CIRCLE', 'DEF_60_CNT_SOCIAL_CIRCLE',
       'DAYS_LAST_PHONE_CHANGE', 'FLAG_DOCUMENT_2', 'FLAG_DOCUMENT_3',
       'FLAG_DOCUMENT_4', 'FLAG_DOCUMENT_5', 'FLAG_DOCUMENT_6',
       'FLAG_DOCUMENT_7', 'FLAG_DOCUMENT_8', 'FLAG_DOCUMENT_9',
       'FLAG_DOCUMENT_10', 'FLAG_DOCUMENT_11', 'FLAG_DOCUMENT_12',
       'FLAG_DOCUMENT_13', 'FLAG_DOCUMENT_14', 'FLAG_DOCUMENT_15',
       'FLAG_DOCUMENT_16', 'FLAG_DOCUMENT_17', 'FLAG_DOCUMENT_18',
       'FLAG_DOCUMENT_19', 'FLAG_DOCUMENT_20', 'FLAG_DOCUMENT_21',
       'AMT_REQ_CREDIT_BUREAU_HOUR', 'AMT_REQ_CREDIT_BUREAU_DAY',
       'AMT_REQ_CREDIT_BUREAU_WEEK', 'AMT_REQ_CREDIT_BUREAU_MON',
       'AMT_REQ_CREDIT_BUREAU_QRT', 'AMT_REQ_CREDIT_BUREAU_YEAR'],
      dtype='object')
```

```
In [64]: def plotting(train, train0, train1, column):
    train = train
    train_0 = train0
    train_1 = train1
    col = column

    print(f"Processing column: {col}, dtype: {train[col].dtype}") # Check the
    fig = plt.figure(figsize=(13,10))

    ax1 = plt.subplot(221)
    # Check data type before plotting
    if train[col].dtype == object: # Assuming categorical columns are objects
        train[col].value_counts().plot.pie(autopct="%1.0f%%", ax=ax1)
    else:
        print(f"Column {col} is not of type object.")
        # Handle other data types if necessary

    plt.title('Plotting data for the column: '+ column)

    ax2 = plt.subplot(222)

    if pd.api.types.is_categorical_dtype(column): # Check if column is categorical
        sns.countplot(x=column, hue='TARGET', data=train, ax=ax2)
        plt.xticks(rotation=90)
```

```

        plt.title('Plotting data for target in terms of total count')
    else:
        sns.histplot(data=train, x=column, bins=10, kde=False, ax=ax2)
        plt.title('Histogram of Numeric Column')

    ax3 = plt.subplot(223)
    df = pd.DataFrame()
    df['0'] = ((train_0[column].value_counts() / len(train_0)))
    df['1'] = ((train_1[column].value_counts() / len(train_1)))
    df.plot.bar(ax=ax3)
    plt.title('Plotting data for target in terms of percentage')

    fig.tight_layout()
    plt.show()

```

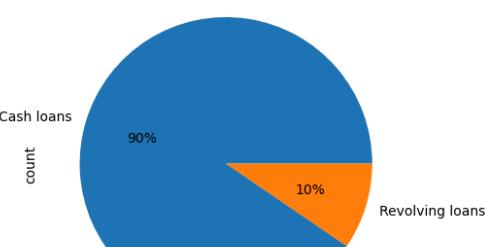
In [65]: train_categorical = appl_data.select_dtypes(include=['object']).columns
train_categorical

Out[65]: Index(['NAME_CONTRACT_TYPE', 'CODE_GENDER', 'FLAG_OWN_CAR', 'FLAG_OWN_REALTY',
 'NAME_TYPE_SUITE', 'NAME_INCOME_TYPE', 'NAME_EDUCATION_TYPE',
 'NAME_FAMILY_STATUS', 'NAME_HOUSING_TYPE', 'OCCUPATION_TYPE',
 'WEEKDAY_APPR_PROCESS_START', 'ORGANIZATION_TYPE',
 'EMERGENCYSTATE_MODE'],
 dtype='object')

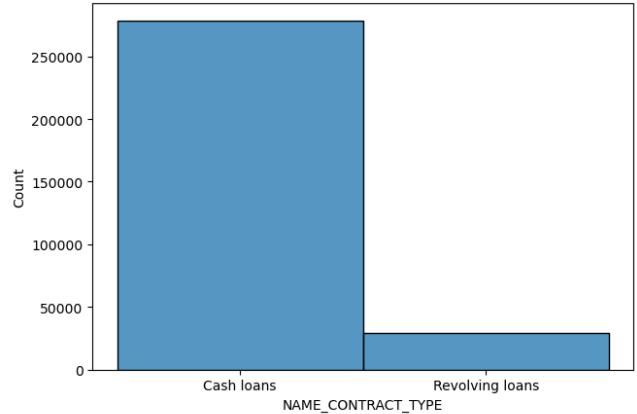
In [66]: **for** column **in** train_categorical:
 print("Plotting ", column)
 plotting(appl_data, train_0, train_1, column)
 print('-----')

Plotting NAME_CONTRACT_TYPE
 Processing column: NAME_CONTRACT_TYPE, dtype: object

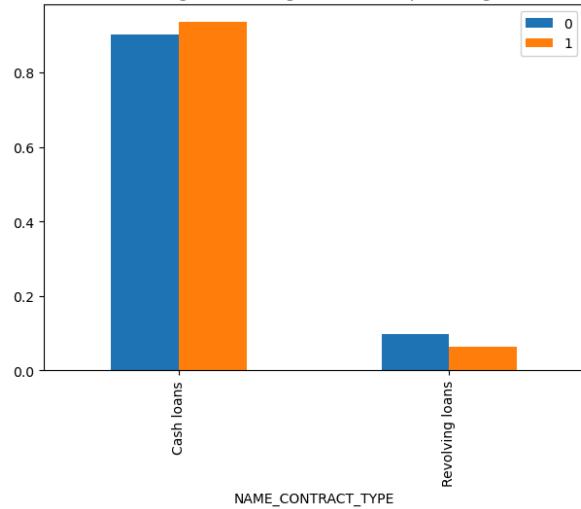
Plotting data for the column: NAME_CONTRACT_TYPE



Histogram of Numeric Column



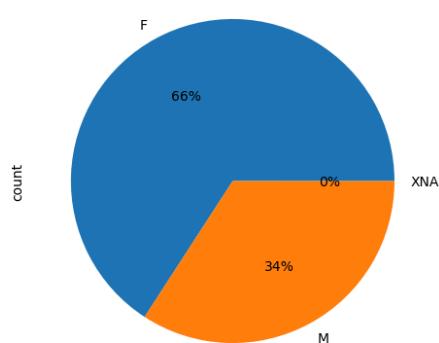
Plotting data for target in terms of percentage



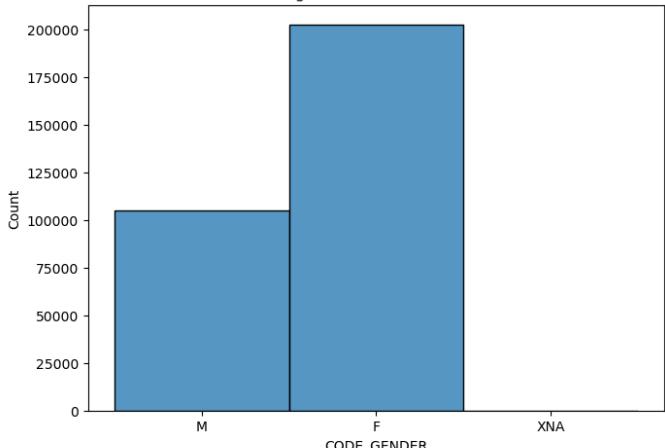
Plotting CODE_GENDER

Processing column: CODE_GENDER, dtype: object

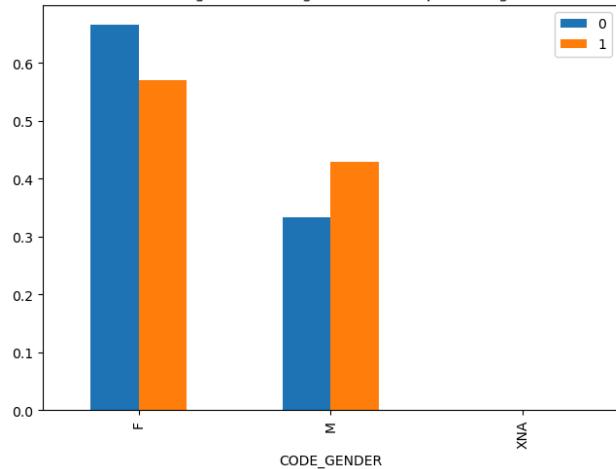
Plotting data for the column: CODE_GENDER



Histogram of Numeric Column



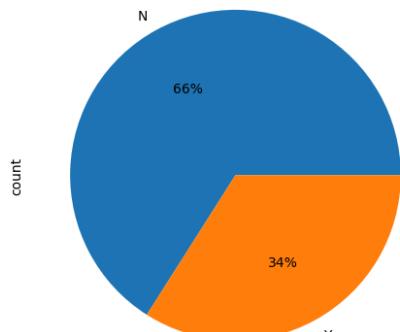
Plotting data for target in terms of percentage



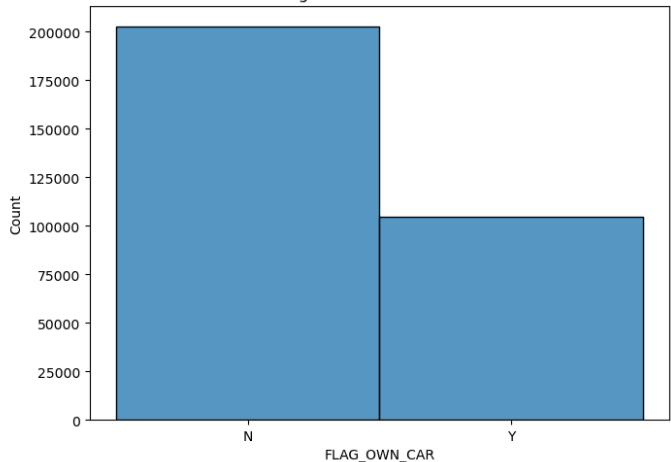
Plotting FLAG_own_car

Processing column: FLAG_own_car, dtype: object

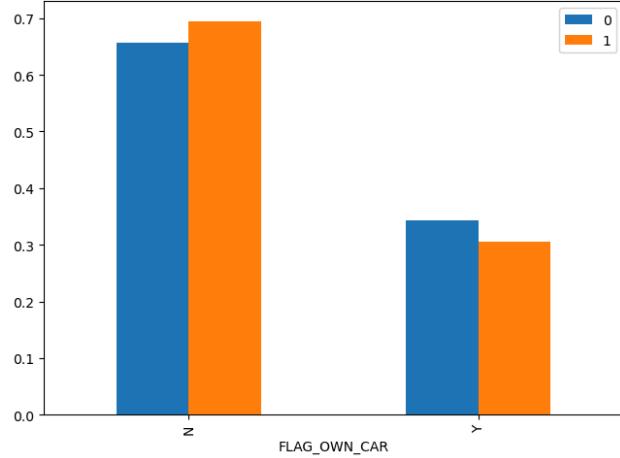
Plotting data for the column: FLAG_own_CAR



Histogram of Numeric Column



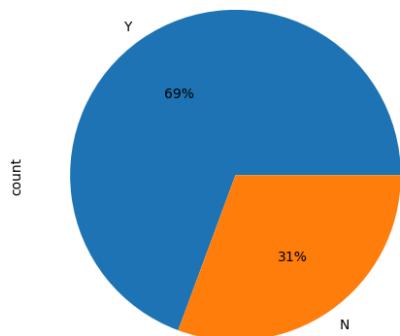
Plotting data for target in terms of percentage



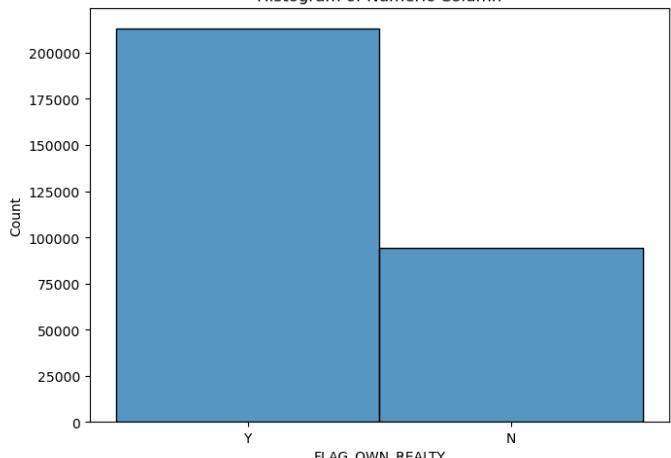
Plotting FLAG_own_REALTY

Processing column: FLAG_own_REALTY, dtype: object

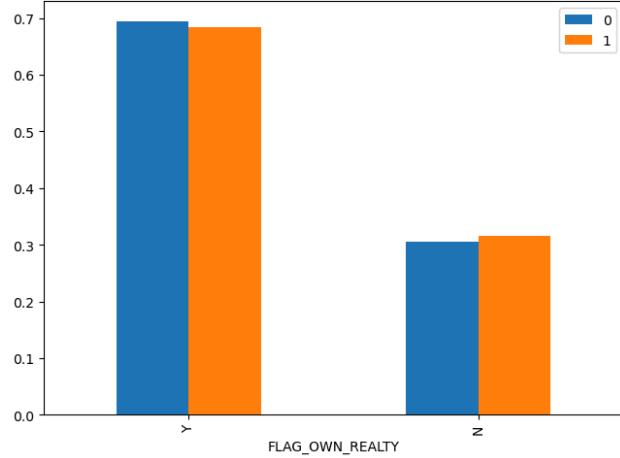
Plotting data for the column: FLAG_OWN_REALTY



Histogram of Numeric Column



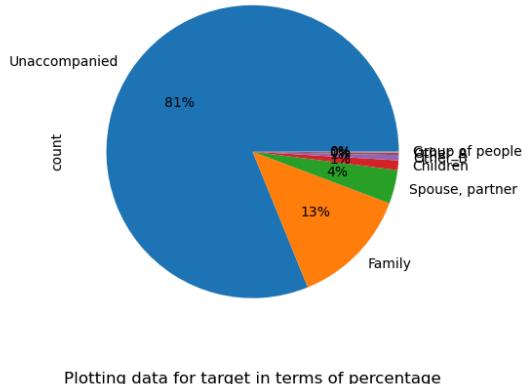
Plotting data for target in terms of percentage



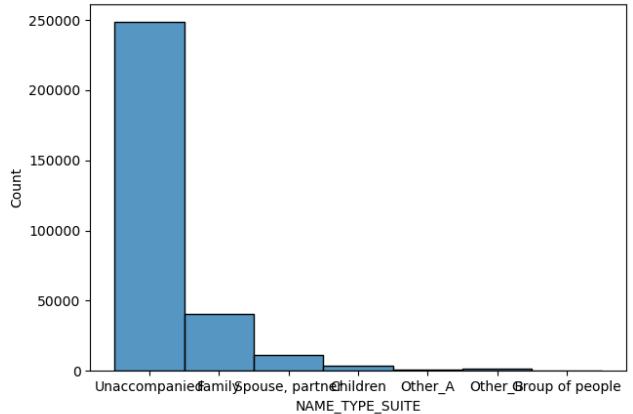
Plotting NAME_TYPE_SUITE

Processing column: NAME_TYPE_SUITE, dtype: object

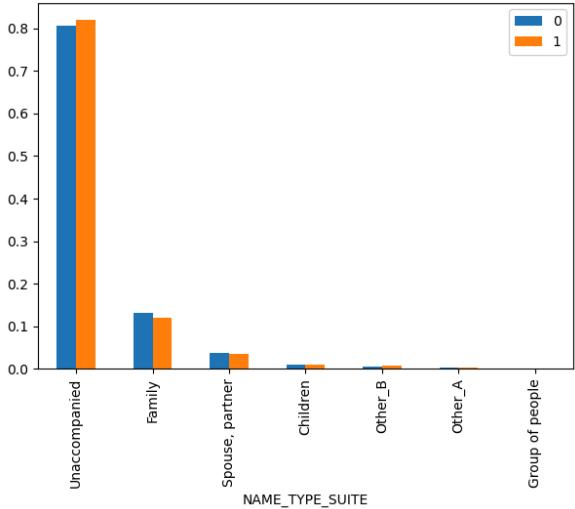
Plotting data for the column: NAME_TYPE_SUITE



Histogram of Numeric Column



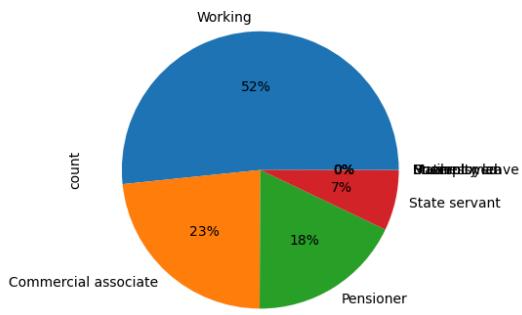
Plotting data for target in terms of percentage



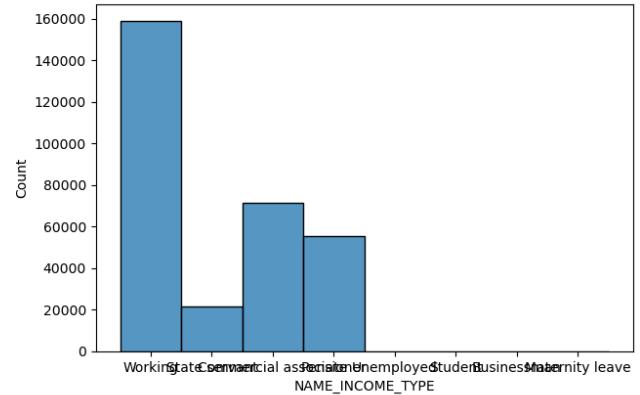
Plotting NAME_INCOME_TYPE

Processing column: NAME_INCOME_TYPE, dtype: object

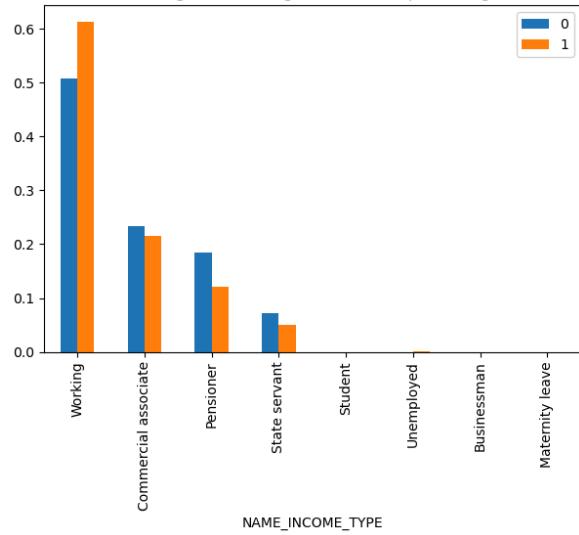
Plotting data for the column: NAME_INCOME_TYPE



Histogram of Numeric Column



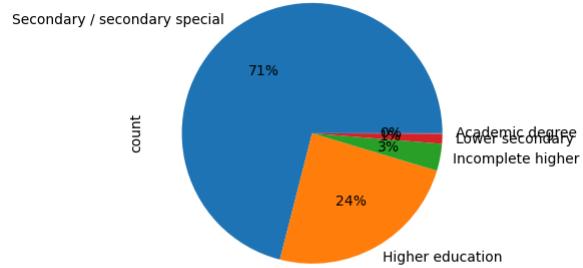
Plotting data for target in terms of percentage



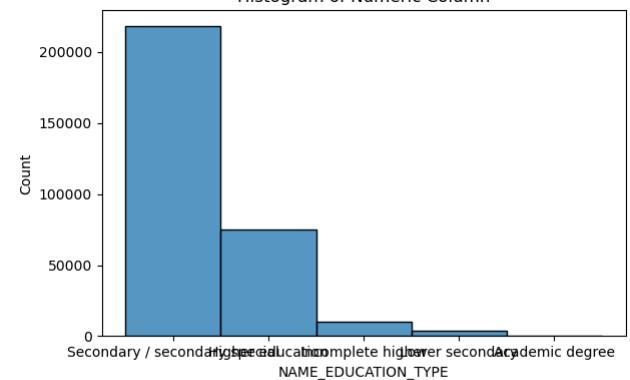
Plotting NAME_EDUCATION_TYPE

Processing column: NAME_EDUCATION_TYPE, dtype: object

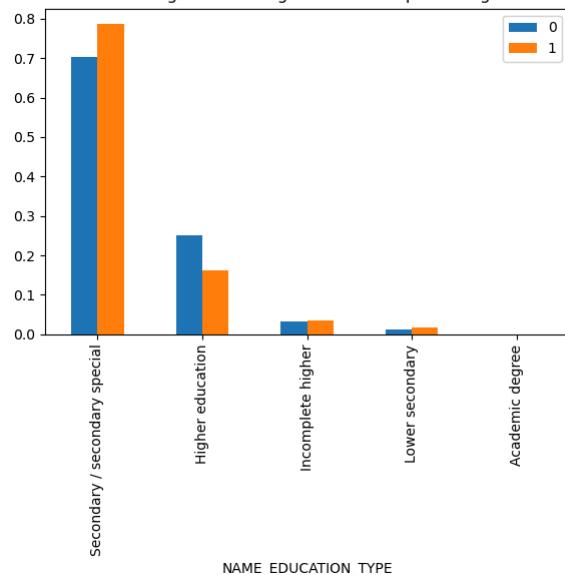
Plotting data for the column: NAME_EDUCATION_TYPE



Histogram of Numeric Column

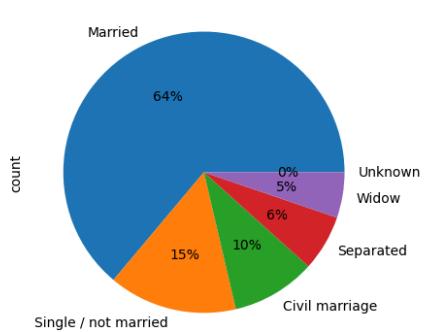


Plotting data for target in terms of percentage

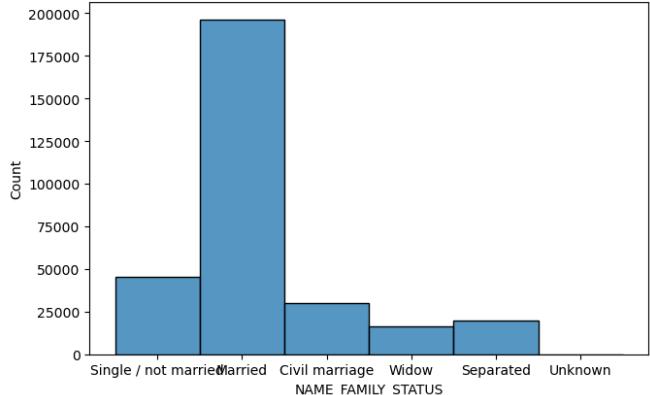


Plotting NAME_FAMILY_STATUS
Processing column: NAME_FAMILY_STATUS, dtype: object

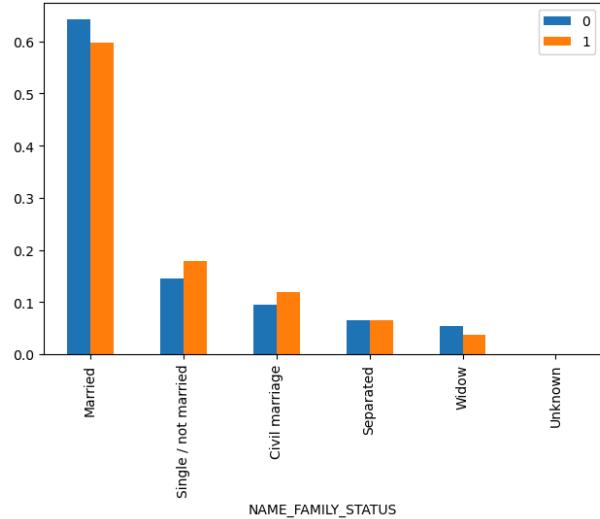
Plotting data for the column: NAME_FAMILY_STATUS



Histogram of Numeric Column



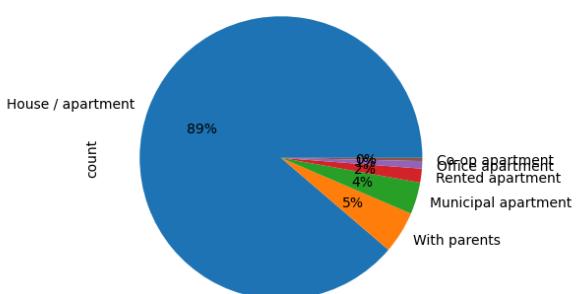
Plotting data for target in terms of percentage



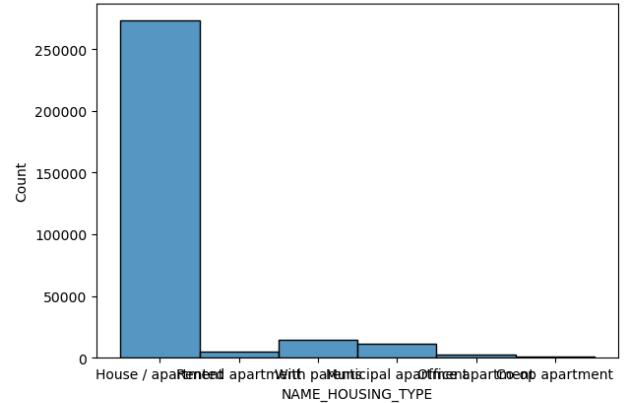
Plotting NAME_HOUSING_TYPE

Processing column: NAME_HOUSING_TYPE, dtype: object

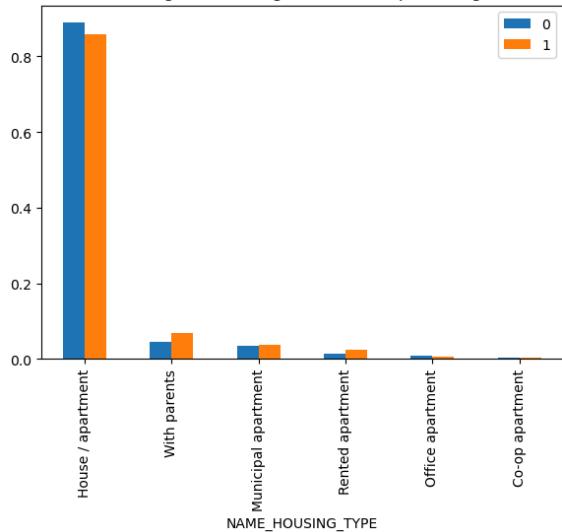
Plotting data for the column: NAME_HOUSING_TYPE



Histogram of Numeric Column



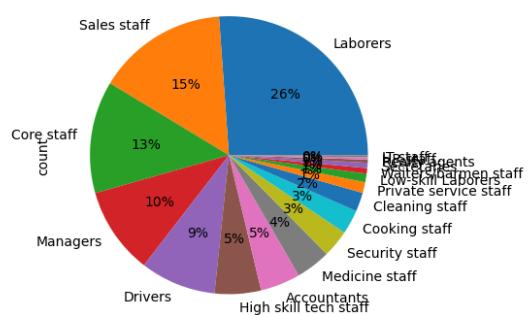
Plotting data for target in terms of percentage



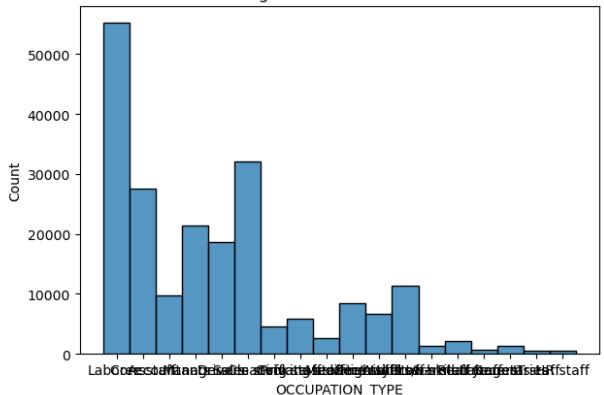
Plotting OCCUPATION_TYPE

Processing column: OCCUPATION_TYPE, dtype: object

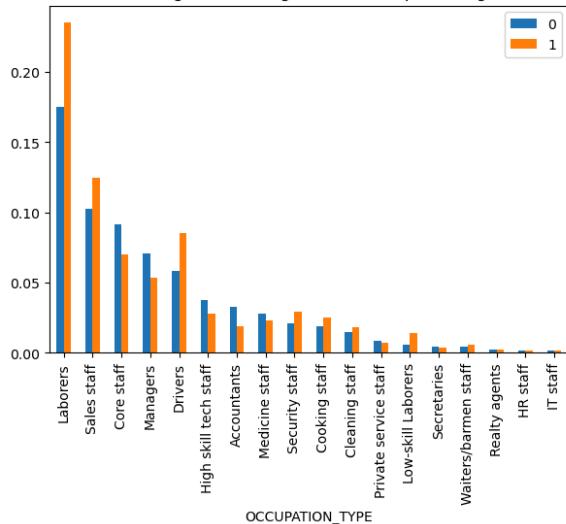
Plotting data for the column: OCCUPATION_TYPE



Histogram of Numeric Column



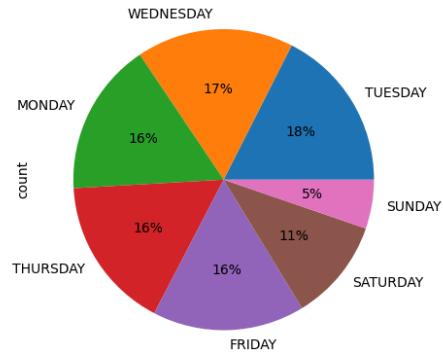
Plotting data for target in terms of percentage



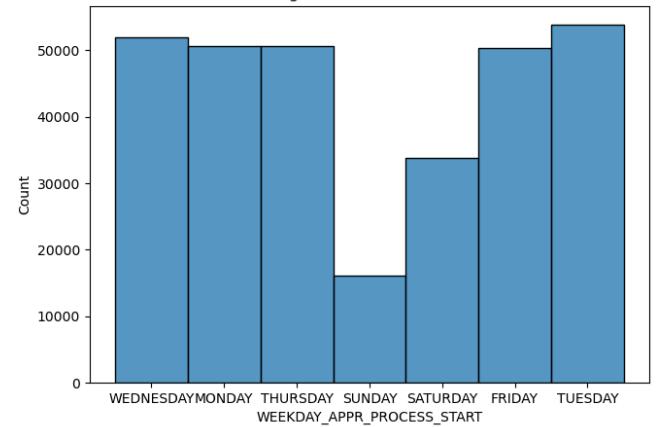
Plotting WEEKDAY_APPR_PROCESS_START

Processing column: WEEKDAY_APPR_PROCESS_START, dtype: object

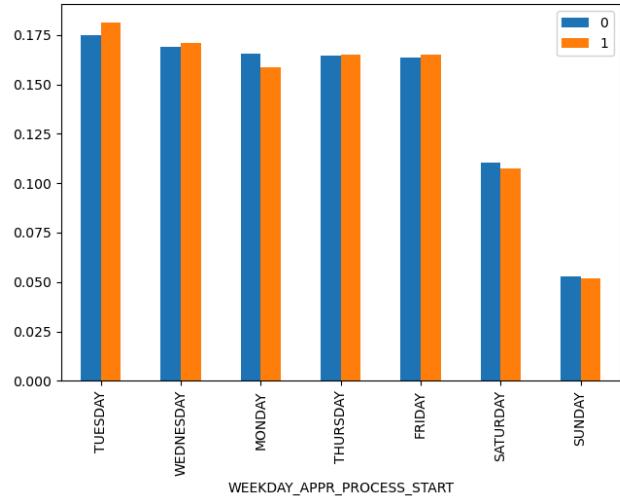
Plotting data for the column: WEEKDAY_APPR_PROCESS_START



Histogram of Numeric Column



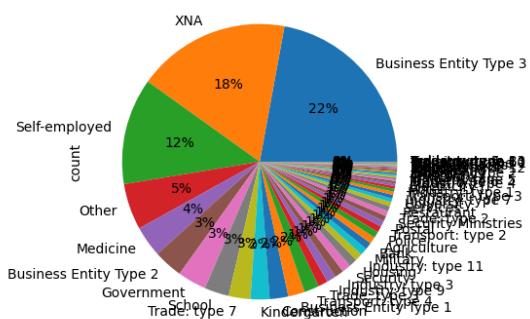
Plotting data for target in terms of percentage



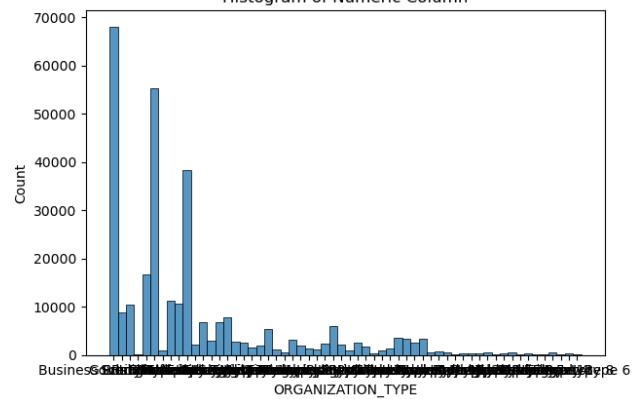
Plotting ORGANIZATION_TYPE

Processing column: ORGANIZATION_TYPE, dtype: object

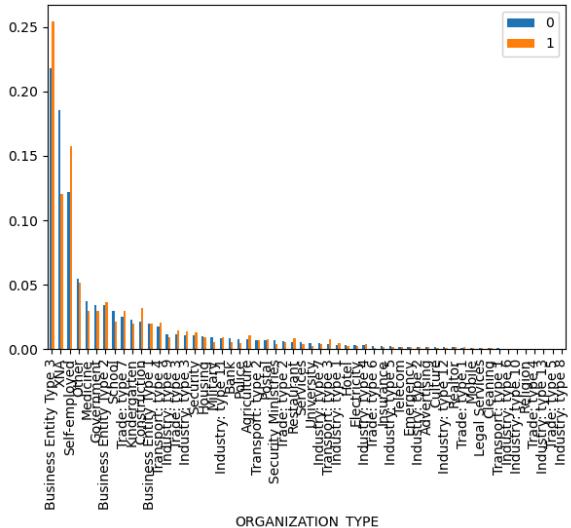
Plotting data for the column: ORGANIZATION_TYPE



Histogram of Numeric Column



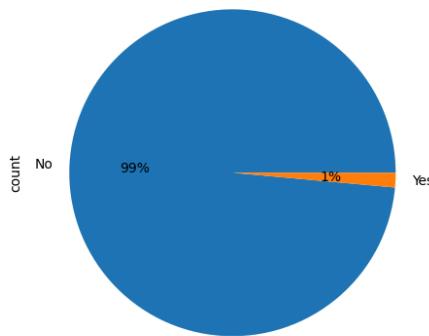
Plotting data for target in terms of percentage



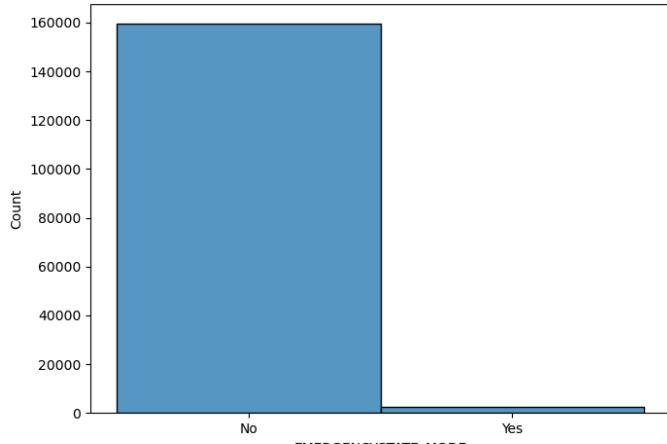
Plotting EMERGENCYSTATE_MODE

Processing column: EMERGENCYSTATE_MODE, dtype: object

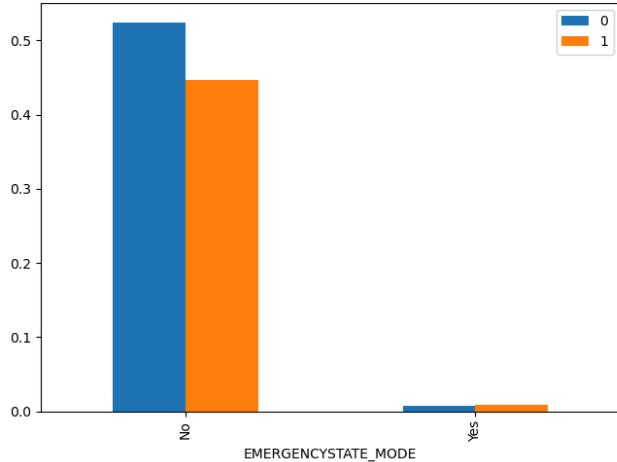
Plotting data for the column: EMERGENCYSTATE_MODE



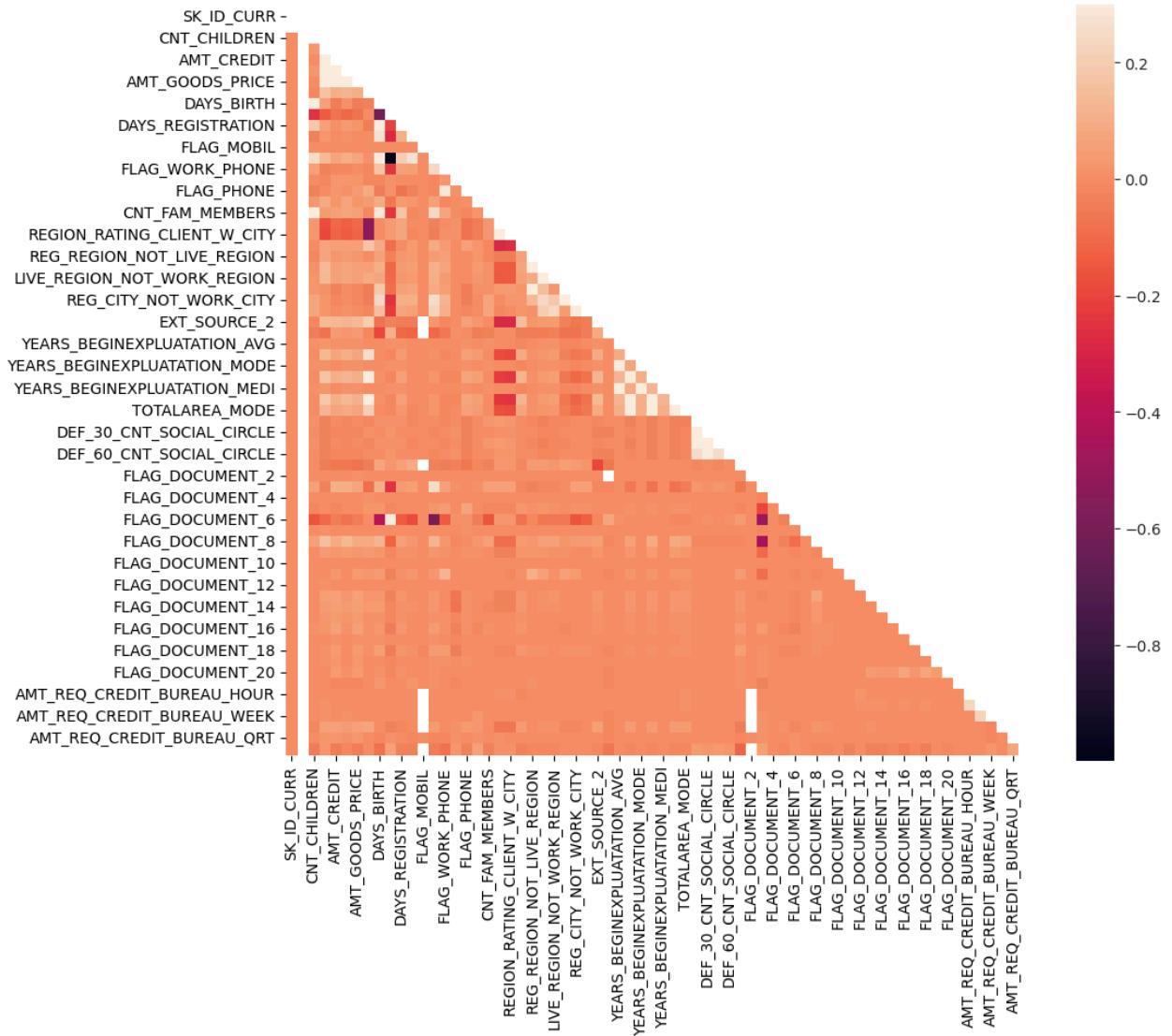
Histogram of Numeric Column



Plotting data for target in terms of percentage

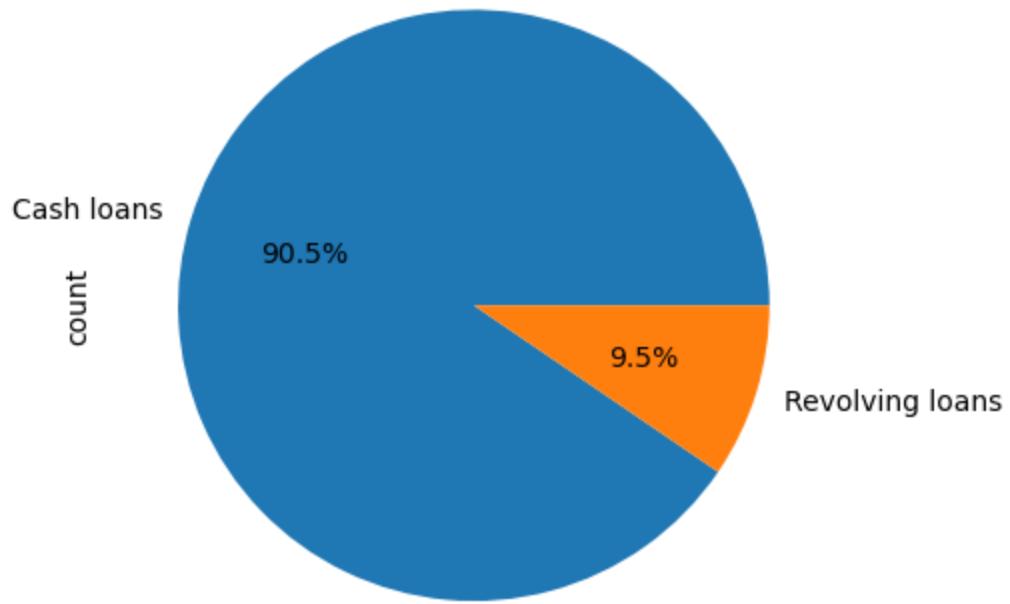


```
In [67]: import numpy as np
corr = train_0.corr(numeric_only=True)
mask = np.zeros_like(corr)
mask[np.triu_indices_from(mask)] = True
f, ax = plt.subplots(figsize=(11, 9))
with sns.axes_style("white"):
    ax = sns.heatmap(corr, mask=mask, vmax=.3, square=True)
```



```
In [68]: appl_data['NAME_CONTRACT_TYPE'].value_counts().plot(kind='pie', autopct='%1.1f
```

```
Out[68]: <Axes: ylabel='count'>
```



```
In [69]: train_0.corr(numeric_only=True)
```

Out[69]:

	SK_ID_CURR	TARGET	CNT_CHILDREN	AMT_I
SK_ID_CURR	1.000000	NaN	-0.000716	
TARGET	NaN	NaN	NaN	
CNT_CHILDREN	-0.000716	NaN	1.000000	
AMT_INCOME_TOTAL	0.001739	NaN	0.027397	
AMT_CREDIT	-0.000342	NaN	0.003081	
AMT_ANNUITY	0.000068	NaN	0.020905	
AMT_GOODS_PRICE	-0.000205	NaN	-0.000525	
REGION_POPULATION_RELATIVE	0.000360	NaN	-0.024363	
DAYS_BIRTH	-0.001346	NaN	0.336966	
DAYS_EMPLOYED	0.001744	NaN	-0.243356	
DAYS_REGISTRATION	-0.001475	NaN	0.185792	
DAYS_ID_PUBLISH	-0.000077	NaN	-0.028751	
FLAG_MOBIL	0.002925	NaN	0.001080	
FLAG_EMP_PHONE	-0.001717	NaN	0.244274	
FLAG_WORK_PHONE	-0.000854	NaN	0.055453	
FLAG_CONT_MOBILE	0.003319	NaN	-0.001201	
FLAG_PHONE	0.001408	NaN	-0.029929	
FLAG_EMAIL	0.000229	NaN	0.023460	
CNT_FAM_MEMBERS	-0.002790	NaN	0.878571	
REGION_RATING_CLIENT	-0.000519	NaN	0.022842	
REGION_RATING_CLIENT_W_CITY	-0.000739	NaN	0.021866	
HOUR_APPR_PROCESS_START	-0.000115	NaN	-0.005244	
REG_REGION_NOT_LIVE_REGION	0.000104	NaN	-0.012342	
REG_REGION_NOT_WORK_REGION	0.000836	NaN	0.010857	
LIVE_REGION_NOT_WORK_REGION	0.002784	NaN	0.017326	
REG_CITY_NOT_LIVE_CITY	-0.002885	NaN	0.021587	
REG_CITY_NOT_WORK_CITY	-0.001696	NaN	0.072193	
LIVE_CITY_NOT_WORK_CITY	0.000428	NaN	0.070988	
EXT_SOURCE_2	0.001589	NaN	-0.015455	
EXT_SOURCE_3	0.000309	NaN	-0.041729	
YEARS_BEGINEXPLUATATION_AVG	0.001202	NaN	0.007298	

	SK_ID_CURR	TARGET	CNT_CHILDREN	AMT_I
FLOORSMAX_AVG	0.002925	NaN	-0.005905	
YEARS_BEGINEXPLUATATION_MODE	0.001442	NaN	0.006545	
FLOORSMAX_MODE	0.003747	NaN	-0.008199	
YEARS_BEGINEXPLUATATION_MEDI	0.000952	NaN	0.006783	
FLOORSMAX_MEDI	0.003810	NaN	-0.008063	
TOTALAREA_MODE	0.002040	NaN	-0.006972	
OBS_30_CNT_SOCIAL_CIRCLE	-0.000695	NaN	0.014471	
DEF_30_CNT_SOCIAL_CIRCLE	0.000556	NaN	-0.002246	
OBS_60_CNT_SOCIAL_CIRCLE	-0.000741	NaN	0.014137	
DEF_60_CNT_SOCIAL_CIRCLE	0.002382	NaN	-0.002172	
DAYS_LAST_PHONE_CHANGE	-0.000998	NaN	-0.008501	
FLAG_DOCUMENT_2	0.001342	NaN	0.003736	
FLAG_DOCUMENT_3	-0.003217	NaN	0.055718	
FLAG_DOCUMENT_4	-0.004321	NaN	-0.003833	
FLAG_DOCUMENT_5	-0.001426	NaN	-0.016239	
FLAG_DOCUMENT_6	0.001972	NaN	-0.158963	
FLAG_DOCUMENT_7	-0.003047	NaN	-0.001442	
FLAG_DOCUMENT_8	0.001618	NaN	0.053919	
FLAG_DOCUMENT_9	0.002007	NaN	-0.002358	
FLAG_DOCUMENT_10	-0.000852	NaN	-0.002859	
FLAG_DOCUMENT_11	-0.001900	NaN	-0.005059	
FLAG_DOCUMENT_12	-0.001092	NaN	0.000322	
FLAG_DOCUMENT_13	0.001006	NaN	0.004340	
FLAG_DOCUMENT_14	-0.001236	NaN	-0.005455	
FLAG_DOCUMENT_15	0.002330	NaN	0.003710	
FLAG_DOCUMENT_16	-0.000753	NaN	0.011775	
FLAG_DOCUMENT_17	0.001776	NaN	0.001160	
FLAG_DOCUMENT_18	0.001194	NaN	0.004181	
FLAG_DOCUMENT_19	0.000676	NaN	0.001078	
FLAG_DOCUMENT_20	0.000832	NaN	0.001302	
FLAG_DOCUMENT_21	-0.000485	NaN	-0.002983	

	SK_ID_CURR	TARGET	CNT_CHILDREN	AMT_I
AMT_REQ_CREDIT_BUREAU_HOUR	-0.001879	NaN		-0.000432
AMT_REQ_CREDIT_BUREAU_DAY	-0.001725	NaN		0.000648
AMT_REQ_CREDIT_BUREAU_WEEK	0.002524	NaN		-0.001632
AMT_REQ_CREDIT_BUREAU_MON	0.000054	NaN		-0.010455
AMT_REQ_CREDIT_BUREAU_QRT	0.000963	NaN		-0.004218
AMT_REQ_CREDIT_BUREAU_YEAR	0.004349	NaN		-0.042547

```
In [70]: train_0.corr(numeric_only=True).abs()
```

Out[70]:

	SK_ID_CURR	TARGET	CNT_CHILDREN	AMT_I
SK_ID_CURR	1.000000	NaN	0.000716	
TARGET	NaN	NaN	NaN	
CNT_CHILDREN	0.000716	NaN	1.000000	
AMT_INCOME_TOTAL	0.001739	NaN	0.027397	
AMT_CREDIT	0.000342	NaN	0.003081	
AMT_ANNUITY	0.000068	NaN	0.020905	
AMT_GOODS_PRICE	0.000205	NaN	0.000525	
REGION_POPULATION_RELATIVE	0.000360	NaN	0.024363	
DAYS_BIRTH	0.001346	NaN	0.336966	
DAYS_EMPLOYED	0.001744	NaN	0.243356	
DAYS_REGISTRATION	0.001475	NaN	0.185792	
DAYS_ID_PUBLISH	0.000077	NaN	0.028751	
FLAG_MOBIL	0.002925	NaN	0.001080	
FLAG_EMP_PHONE	0.001717	NaN	0.244274	
FLAG_WORK_PHONE	0.000854	NaN	0.055453	
FLAG_CONT_MOBILE	0.003319	NaN	0.001201	
FLAG_PHONE	0.001408	NaN	0.029929	
FLAG_EMAIL	0.000229	NaN	0.023460	
CNT_FAM_MEMBERS	0.002790	NaN	0.878571	
REGION_RATING_CLIENT	0.000519	NaN	0.022842	
REGION_RATING_CLIENT_W_CITY	0.000739	NaN	0.021866	
HOUR_APPR_PROCESS_START	0.000115	NaN	0.005244	
REG_REGION_NOT_LIVE_REGION	0.000104	NaN	0.012342	
REG_REGION_NOT_WORK_REGION	0.000836	NaN	0.010857	
LIVE_REGION_NOT_WORK_REGION	0.002784	NaN	0.017326	
REG_CITY_NOT_LIVE_CITY	0.002885	NaN	0.021587	
REG_CITY_NOT_WORK_CITY	0.001696	NaN	0.072193	
LIVE_CITY_NOT_WORK_CITY	0.000428	NaN	0.070988	
EXT_SOURCE_2	0.001589	NaN	0.015455	
EXT_SOURCE_3	0.000309	NaN	0.041729	
YEARS_BEGINEXPLUATATION_AVG	0.001202	NaN	0.007298	

	SK_ID_CURR	TARGET	CNT_CHILDREN	AMT_I
FLOORSMAX_AVG	0.002925	NaN	0.005905	
YEARS_BEGINEXPLUATATION_MODE	0.001442	NaN	0.006545	
FLOORSMAX_MODE	0.003747	NaN	0.008199	
YEARS_BEGINEXPLUATATION_MEDI	0.000952	NaN	0.006783	
FLOORSMAX_MEDI	0.003810	NaN	0.008063	
TOTALAREA_MODE	0.002040	NaN	0.006972	
OBS_30_CNT_SOCIAL_CIRCLE	0.000695	NaN	0.014471	
DEF_30_CNT_SOCIAL_CIRCLE	0.000556	NaN	0.002246	
OBS_60_CNT_SOCIAL_CIRCLE	0.000741	NaN	0.014137	
DEF_60_CNT_SOCIAL_CIRCLE	0.002382	NaN	0.002172	
DAYS_LAST_PHONE_CHANGE	0.000998	NaN	0.008501	
FLAG_DOCUMENT_2	0.001342	NaN	0.003736	
FLAG_DOCUMENT_3	0.003217	NaN	0.055718	
FLAG_DOCUMENT_4	0.004321	NaN	0.003833	
FLAG_DOCUMENT_5	0.001426	NaN	0.016239	
FLAG_DOCUMENT_6	0.001972	NaN	0.158963	
FLAG_DOCUMENT_7	0.003047	NaN	0.001442	
FLAG_DOCUMENT_8	0.001618	NaN	0.053919	
FLAG_DOCUMENT_9	0.002007	NaN	0.002358	
FLAG_DOCUMENT_10	0.000852	NaN	0.002859	
FLAG_DOCUMENT_11	0.001900	NaN	0.005059	
FLAG_DOCUMENT_12	0.001092	NaN	0.000322	
FLAG_DOCUMENT_13	0.001006	NaN	0.004340	
FLAG_DOCUMENT_14	0.001236	NaN	0.005455	
FLAG_DOCUMENT_15	0.002330	NaN	0.003710	
FLAG_DOCUMENT_16	0.000753	NaN	0.011775	
FLAG_DOCUMENT_17	0.001776	NaN	0.001160	
FLAG_DOCUMENT_18	0.001194	NaN	0.004181	
FLAG_DOCUMENT_19	0.000676	NaN	0.001078	
FLAG_DOCUMENT_20	0.000832	NaN	0.001302	
FLAG_DOCUMENT_21	0.000485	NaN	0.002983	

	SK_ID_CURR	TARGET	CNT_CHILDREN	AMT_I
AMT_REQ_CREDIT_BUREAU_HOUR	0.001879	NaN	0.000432	
AMT_REQ_CREDIT_BUREAU_DAY	0.001725	NaN	0.000648	
AMT_REQ_CREDIT_BUREAU_WEEK	0.002524	NaN	0.001632	
AMT_REQ_CREDIT_BUREAU_MON	0.000054	NaN	0.010455	
AMT_REQ_CREDIT_BUREAU_QRT	0.000963	NaN	0.004218	
AMT_REQ_CREDIT_BUREAU_YEAR	0.004349	NaN	0.042547	

```
In [71]: train_0.corr(numeric_only=True).abs().unstack()
```

```
Out[71]: SK_ID_CURR           SK_ID_CURR      1.000000
          TARGET             NaN
          CNT_CHILDREN       0.000716
          AMT_INCOME_TOTAL   0.001739
          AMT_CREDIT          0.000342
          ...
          AMT_REQ_CREDIT_BUREAU_YEAR AMT_REQ_CREDIT_BUREAU_DAY 0.003798
                                      AMT_REQ_CREDIT_BUREAU_WEEK 0.019085
                                      AMT_REQ_CREDIT_BUREAU_MON 0.004889
                                      AMT_REQ_CREDIT_BUREAU_QRT 0.074664
                                      AMT_REQ_CREDIT_BUREAU_YEAR 1.000000
Length: 4624, dtype: float64
```

```
In [73]: train_0.corr(numeric_only=True).abs().unstack().sort_values()
```

```
Out[73]: FLAG_DOCUMENT_2           EXT_SOURCE_2      0.000004
          EXT_SOURCE_2             FLAG_DOCUMENT_2  0.000004
          FLAG_MOBIL                FLAG_DOCUMENT_12  0.000005
          FLAG_DOCUMENT_12            FLAG_MOBIL      0.000005
          YEARS_BEGINEXPLUATATION_MEDI FLAG_DOCUMENT_4  0.000008
          ...
          AMT_REQ_CREDIT_BUREAU_MON  FLAG_DOCUMENT_2      NaN
          AMT_REQ_CREDIT_BUREAU_QRT  TARGET            NaN
          AMT_REQ_CREDIT_BUREAU_YEAR TARGET            NaN
                                      FLAG_MOBIL      NaN
                                      FLAG_DOCUMENT_2  NaN
Length: 4624, dtype: float64
```

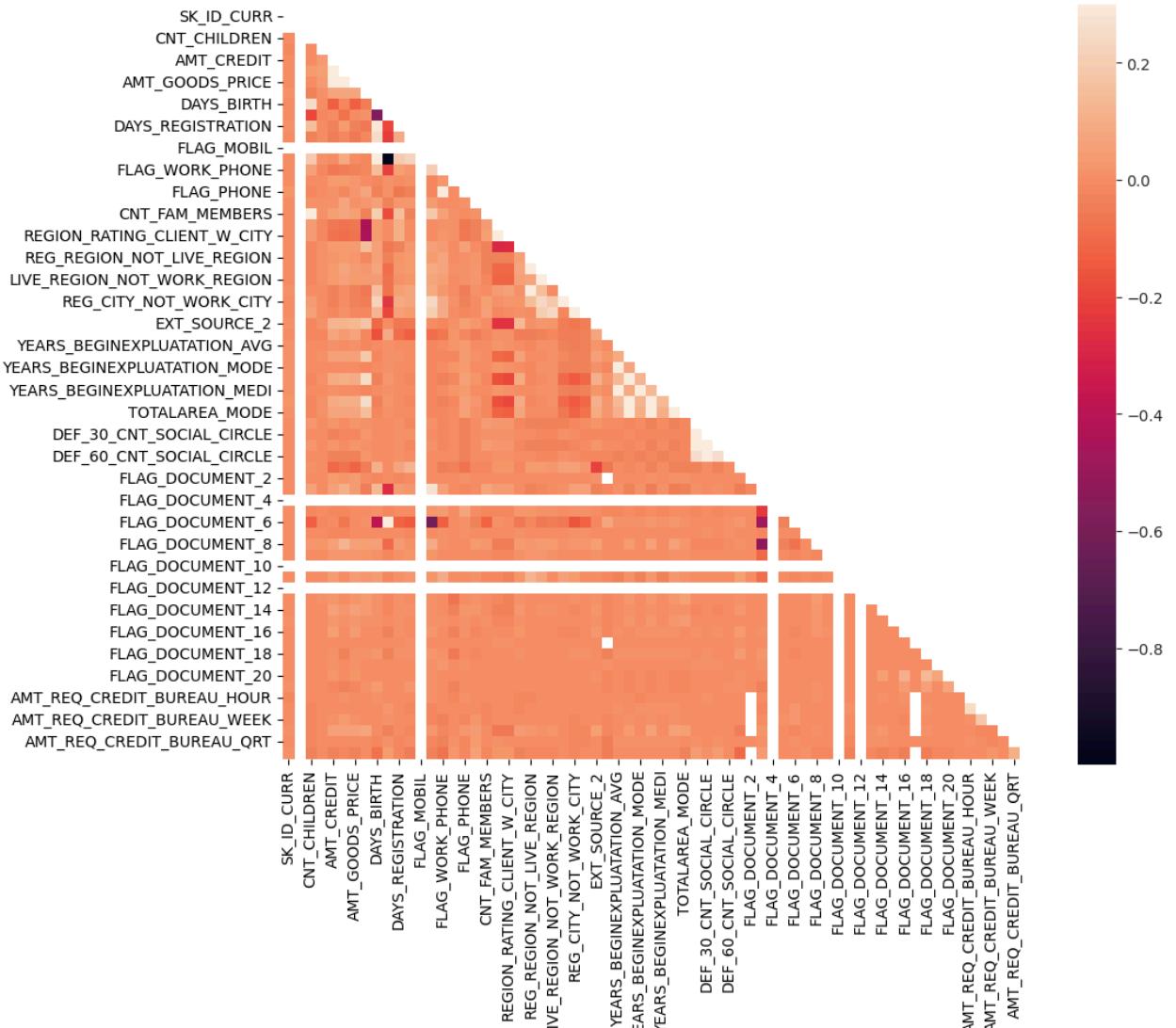
```
In [76]: correlation_0 = train_0.corr(numeric_only=True).abs().unstack().sort_values().
```

```
In [77]: correlation_0 = correlation_0[ correlation_0 != 1.0 ]
print(correlation_0)
```

FLAG_DOCUMENT_2	EXT_SOURCE_2	0.000004
EXT_SOURCE_2	FLAG_DOCUMENT_2	0.000004
FLAG_MOBIL	FLAG_DOCUMENT_12	0.000005
FLAG_DOCUMENT_12	FLAG_MOBIL	0.000005
YEARS_BEGINEXPLUATATION_MEDI	FLAG_DOCUMENT_4	0.000008
		...
FLOORSMAX_AVG	FLOORSMAX_MEDI	0.997018
OBS_30_CNT_SOCIAL_CIRCLE	OBS_60_CNT_SOCIAL_CIRCLE	0.998508
OBS_60_CNT_SOCIAL_CIRCLE	OBS_30_CNT_SOCIAL_CIRCLE	0.998508
DAYS_EMPLOYED	FLAG_EMP_PHONE	0.999758
FLAG_EMP_PHONE	DAYS_EMPLOYED	0.999758

Length: 4394, dtype: float64

```
In [79]: corr = train_1.corr(numeric_only=True)
mask = np.zeros_like(corr)
mask[np.triu_indices_from(mask)] = True
f, ax = plt.subplots(figsize=(11, 9))
with sns.axes_style("white"):
    ax = sns.heatmap(corr, mask=mask, vmax=.3, square=True)
```

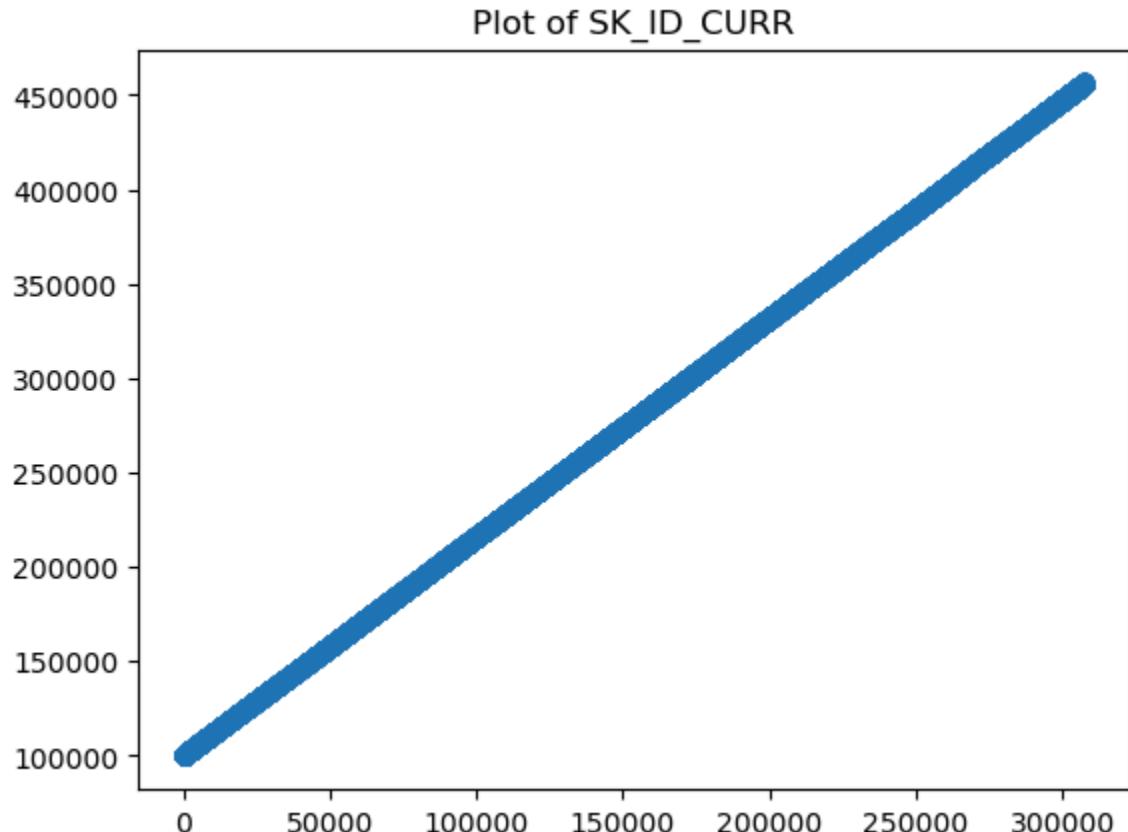


```
In [80]: train_categorical = appl_data.select_dtypes(include=['int64', 'float64']).columns
```

```
In [82]: for column in train_categorical:  
    title = "Plot of "+column  
    plt.scatter(appl_data.index, appl_data[column])  
    plt.title(title)  
    plt.show()
```

```
Out[82]: <matplotlib.collections.PathCollection at 0x22fdbba91a50>
```

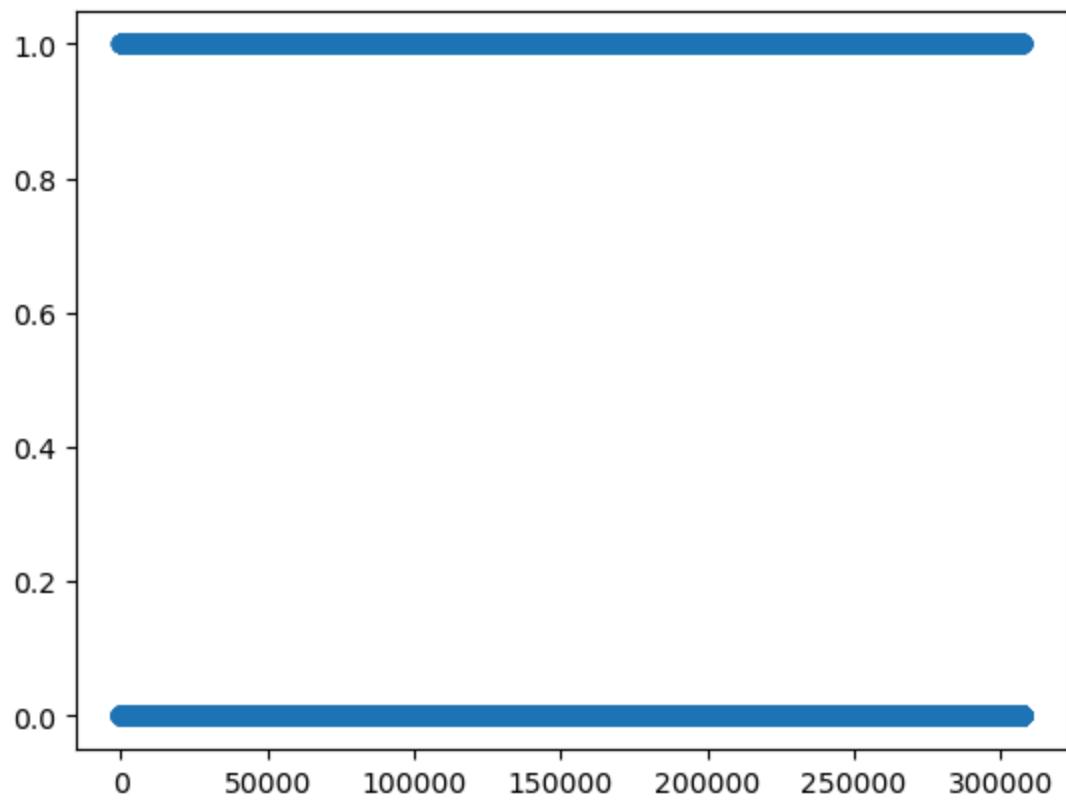
```
Out[82]: Text(0.5, 1.0, 'Plot of SK_ID_CURR')
```



```
Out[82]: <matplotlib.collections.PathCollection at 0x22fdbbb63ad0>
```

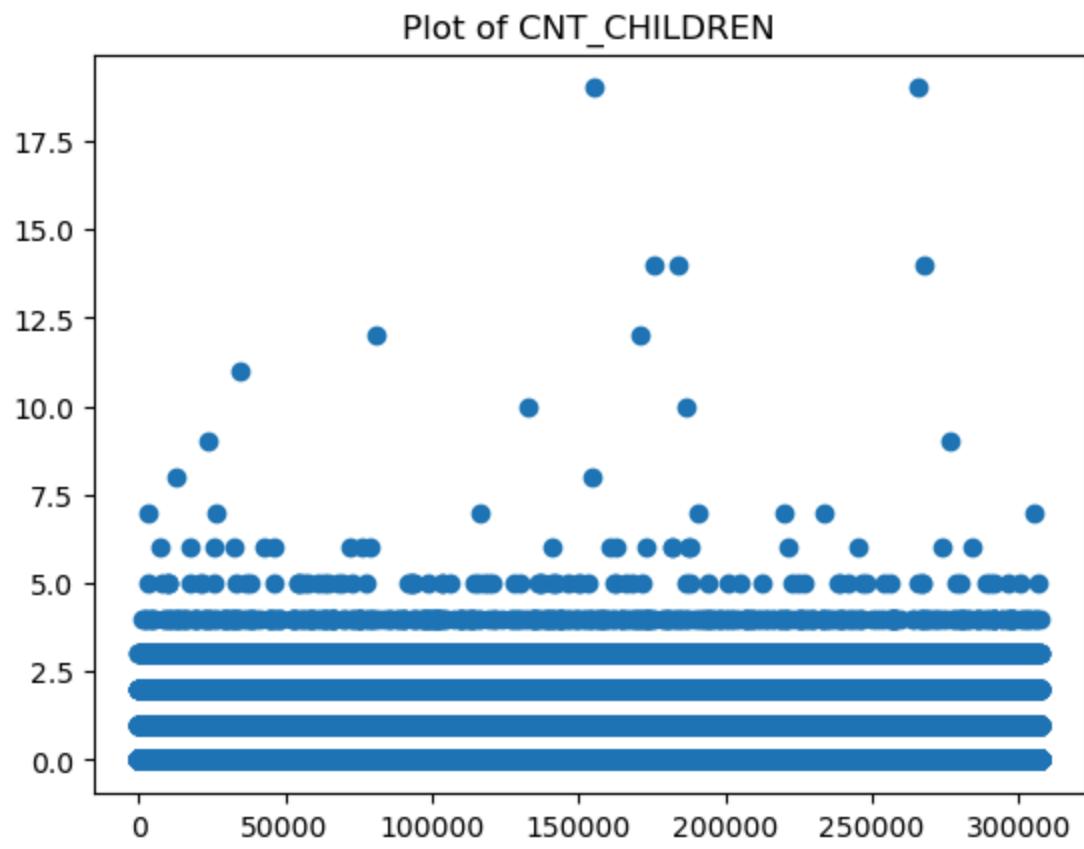
```
Out[82]: Text(0.5, 1.0, 'Plot of TARGET')
```

Plot of TARGET



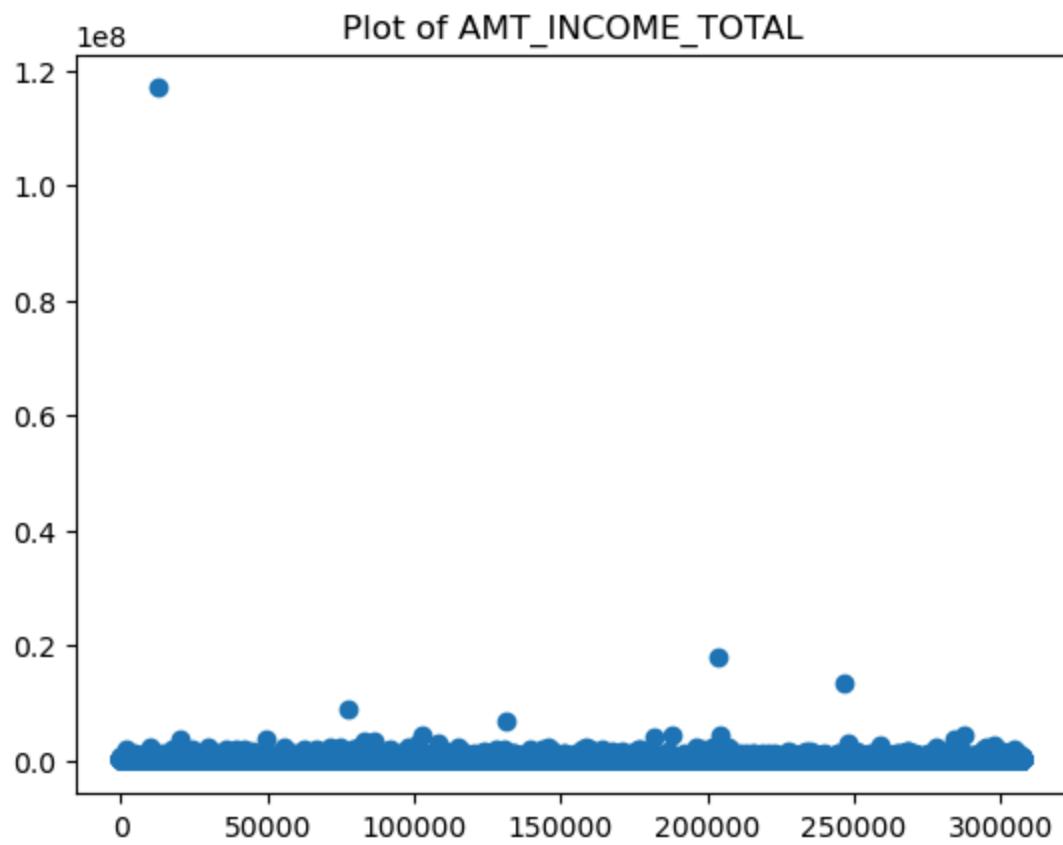
Out[82]: <matplotlib.collections.PathCollection at 0x22fdbbaf450>

Out[82]: Text(0.5, 1.0, 'Plot of CNT_CHILDREN')



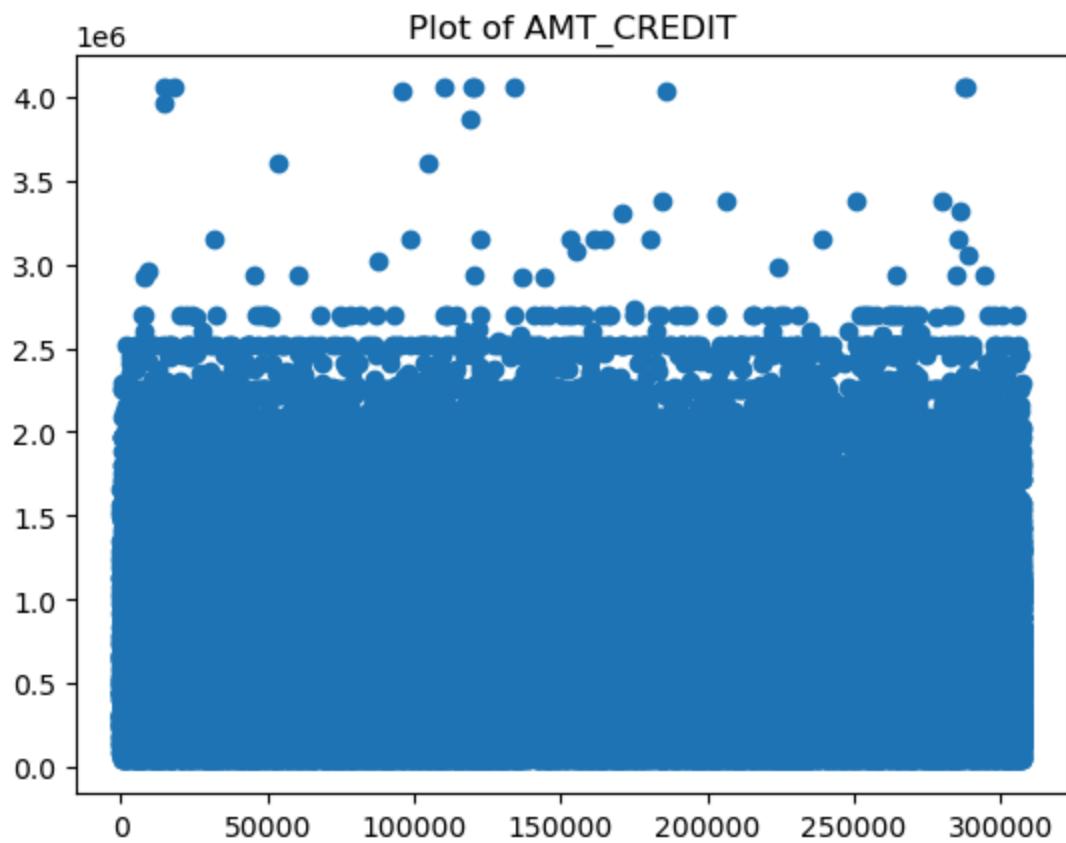
```
Out[82]: <matplotlib.collections.PathCollection at 0x22fdbc39c90>
```

```
Out[82]: Text(0.5, 1.0, 'Plot of AMT_INCOME_TOTAL')
```



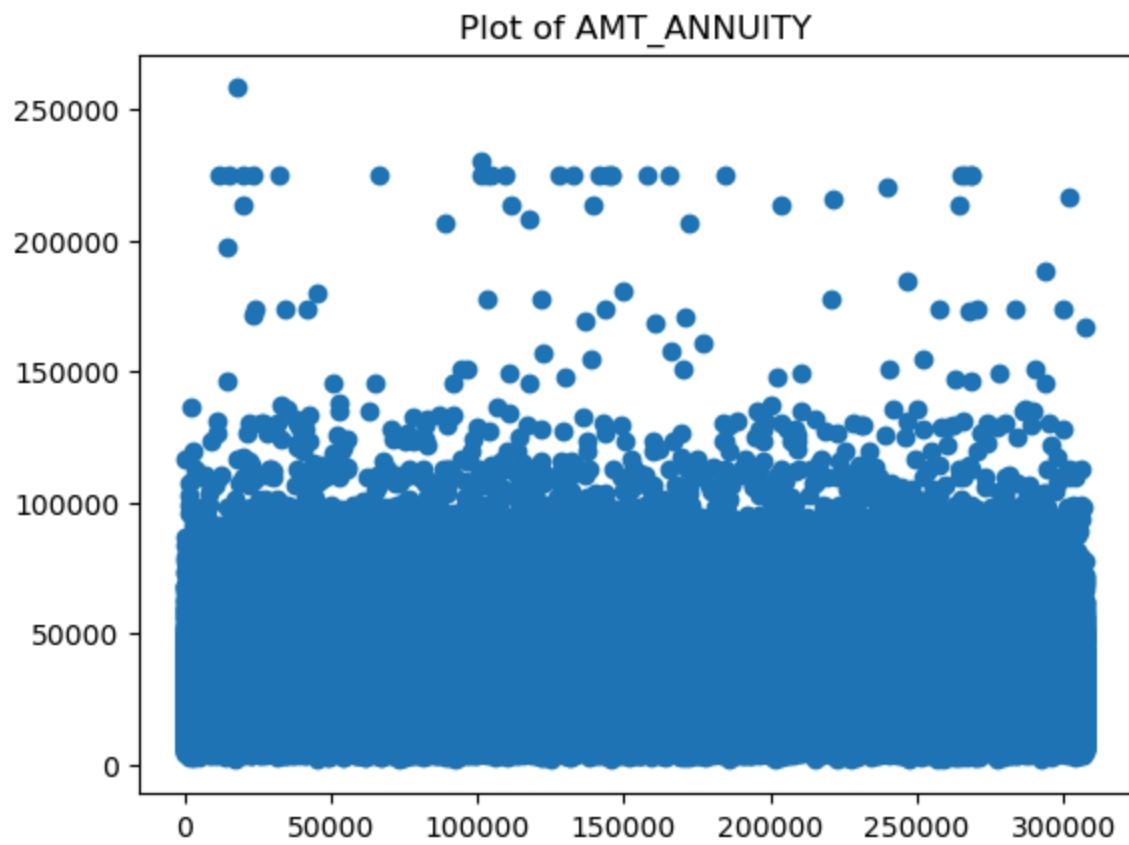
Out[82]: <matplotlib.collections.PathCollection at 0x22fdbbe9f10>

Out[82]: Text(0.5, 1.0, 'Plot of AMT_CREDIT')



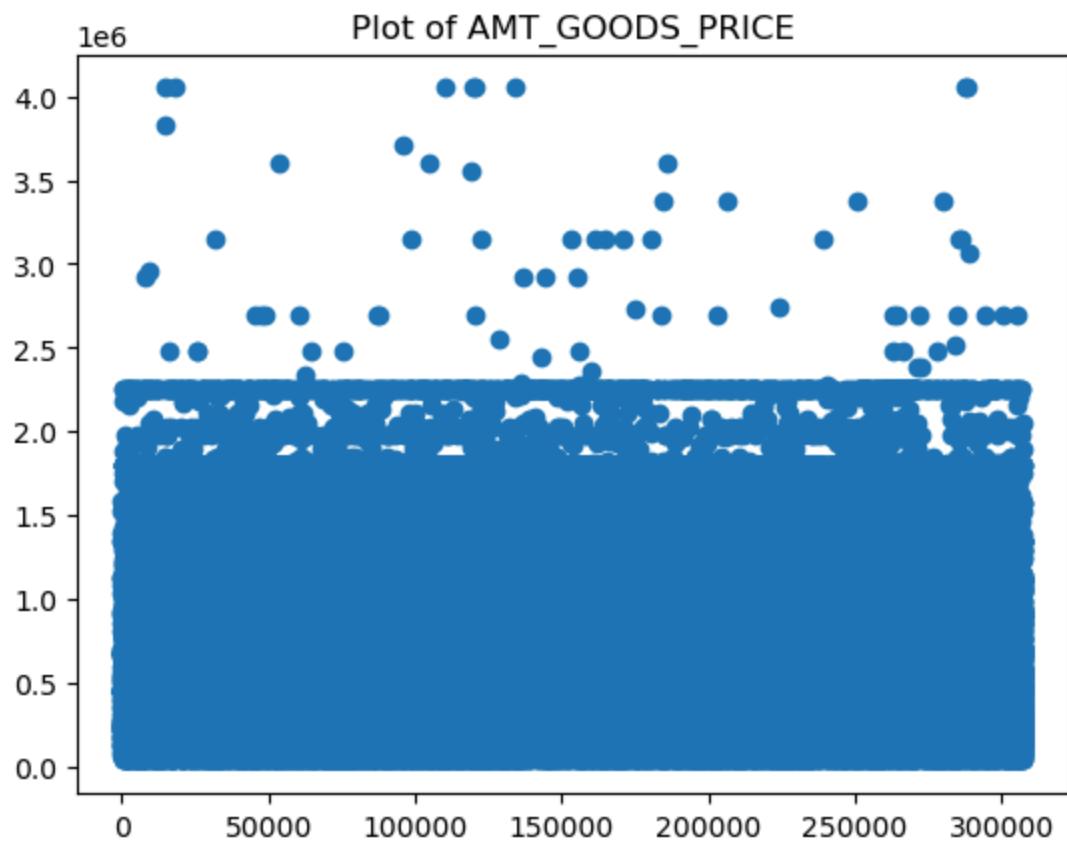
```
Out[82]: <matplotlib.collections.PathCollection at 0x22fdbc0bc10>
```

```
Out[82]: Text(0.5, 1.0, 'Plot of AMT_ANNUITY')
```



```
Out[82]: <matplotlib.collections.PathCollection at 0x22fdbcb2050>
```

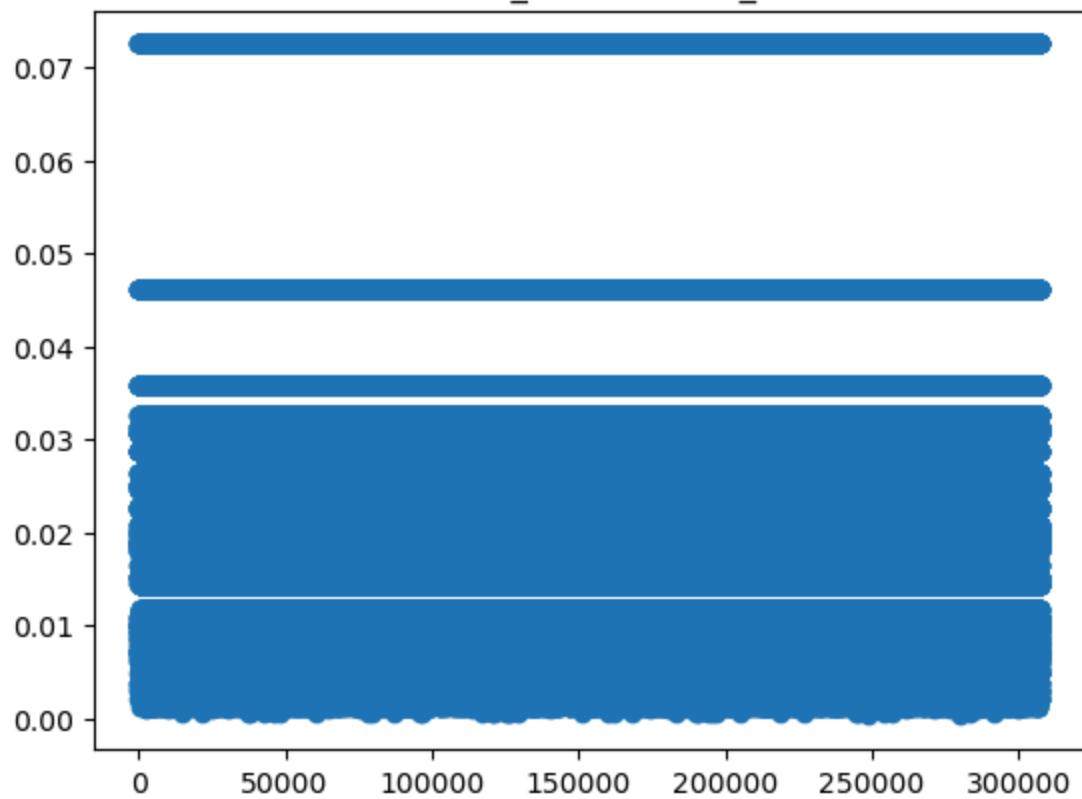
```
Out[82]: Text(0.5, 1.0, 'Plot of AMT_GOODS_PRICE')
```



```
Out[82]: <matplotlib.collections.PathCollection at 0x22fdbdcc210>
```

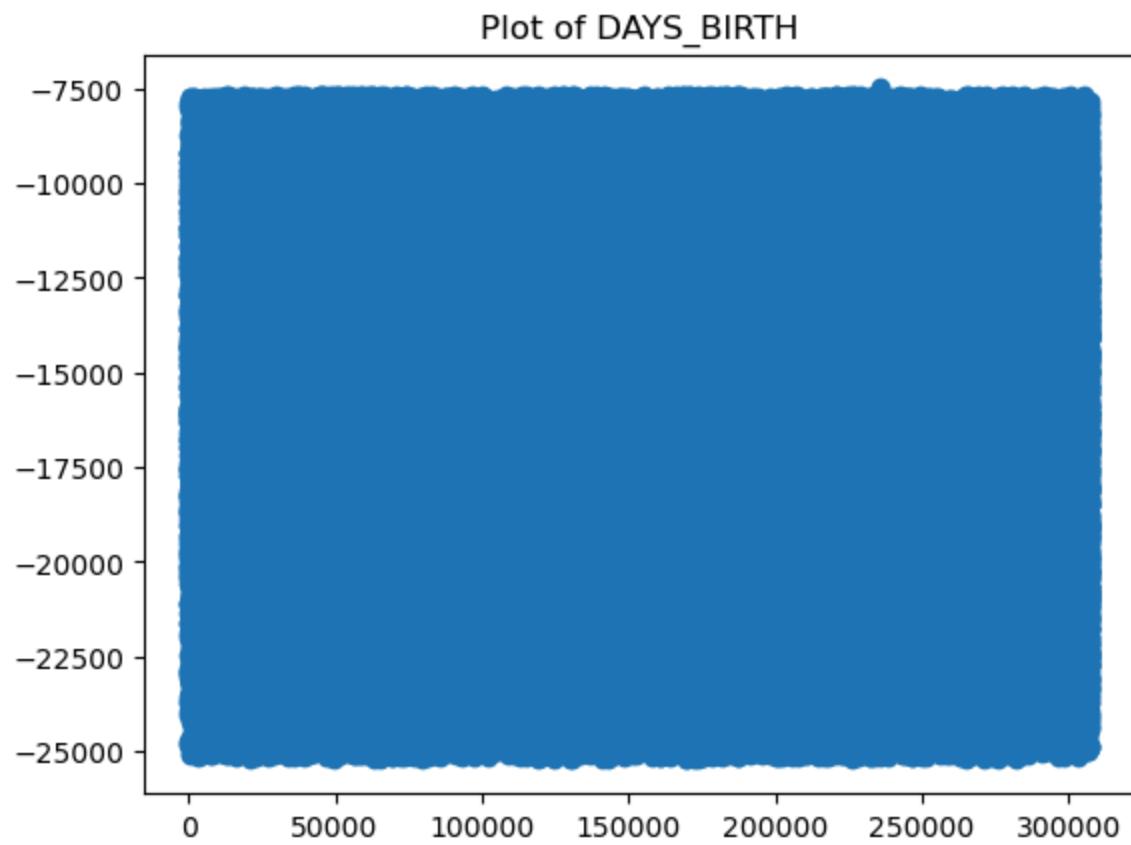
```
Out[82]: Text(0.5, 1.0, 'Plot of REGION_POPULATION_RELATIVE')
```

Plot of REGION_POPULATION_RELATIVE



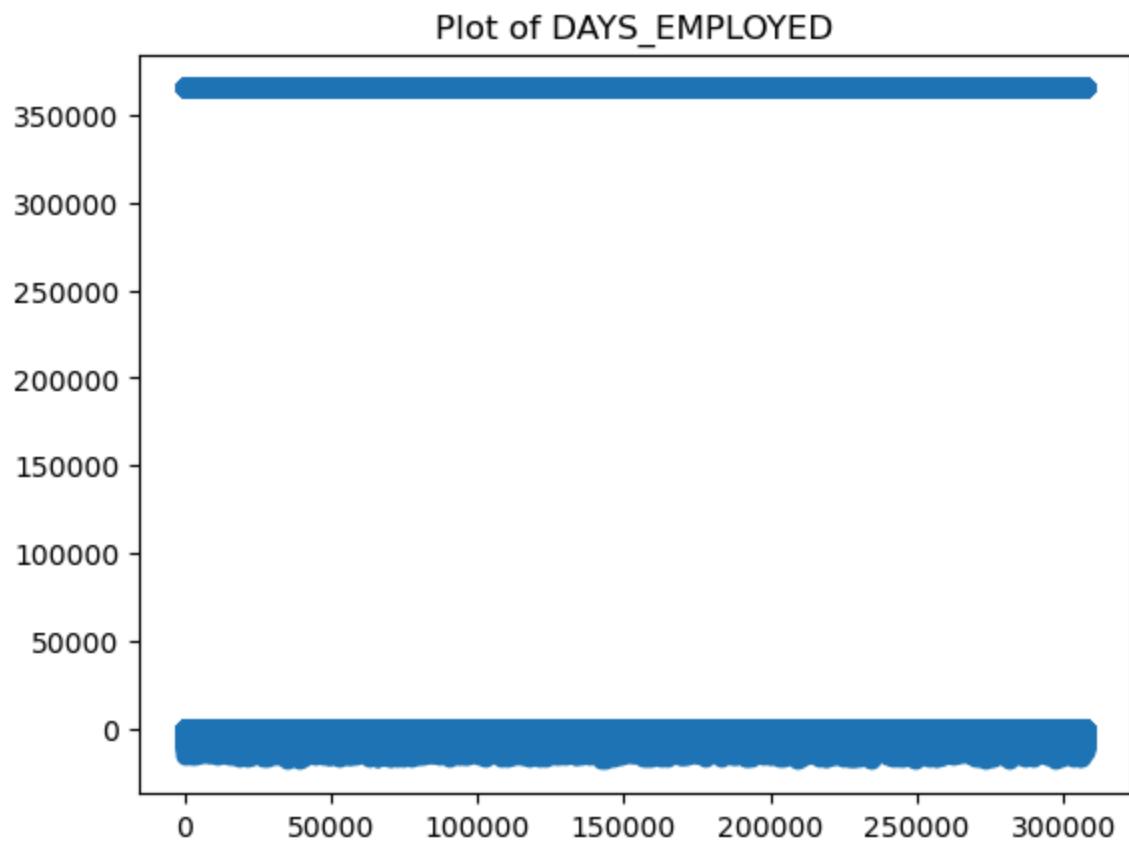
Out[82]: <matplotlib.collections.PathCollection at 0x22fdbbe36390>

Out[82]: Text(0.5, 1.0, 'Plot of DAYS_BIRTH')



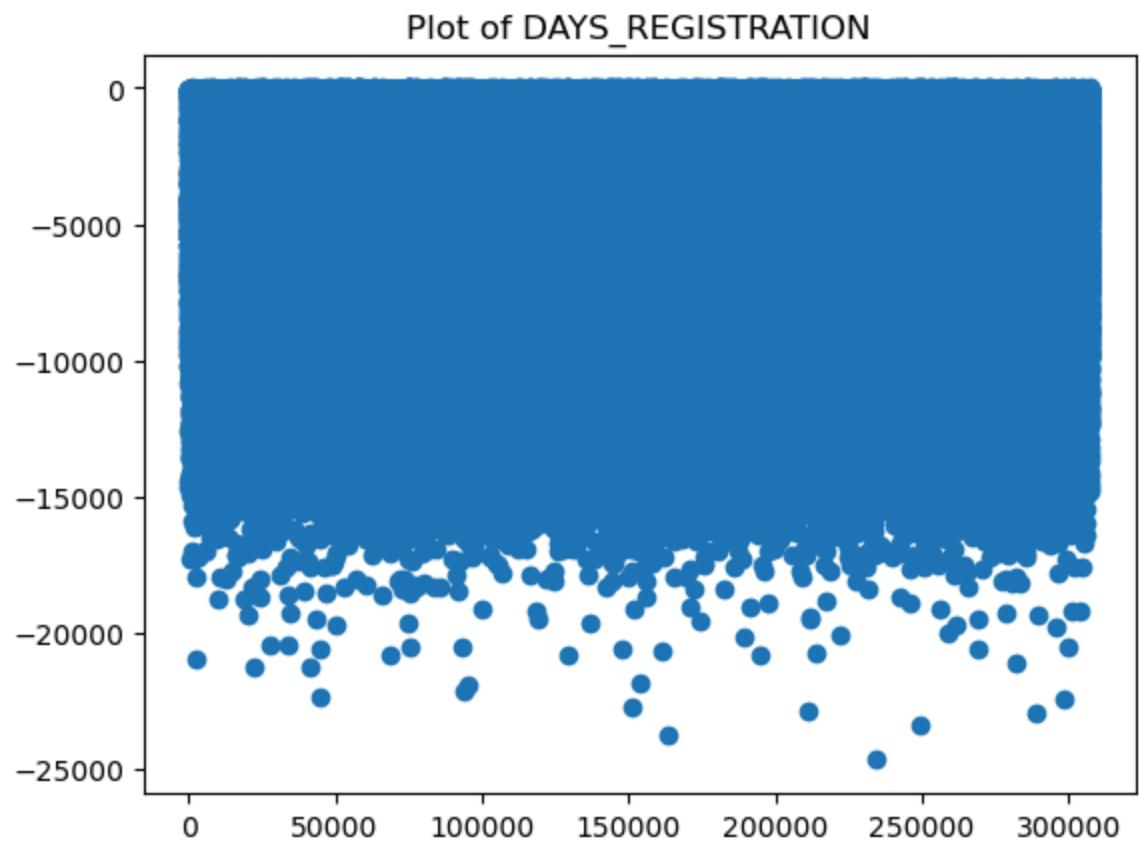
```
Out[82]: <matplotlib.collections.PathCollection at 0x22ffa01bbd0>
```

```
Out[82]: Text(0.5, 1.0, 'Plot of DAYS_EMPLOYED')
```



```
Out[82]: <matplotlib.collections.PathCollection at 0x22fdbdfa390>
```

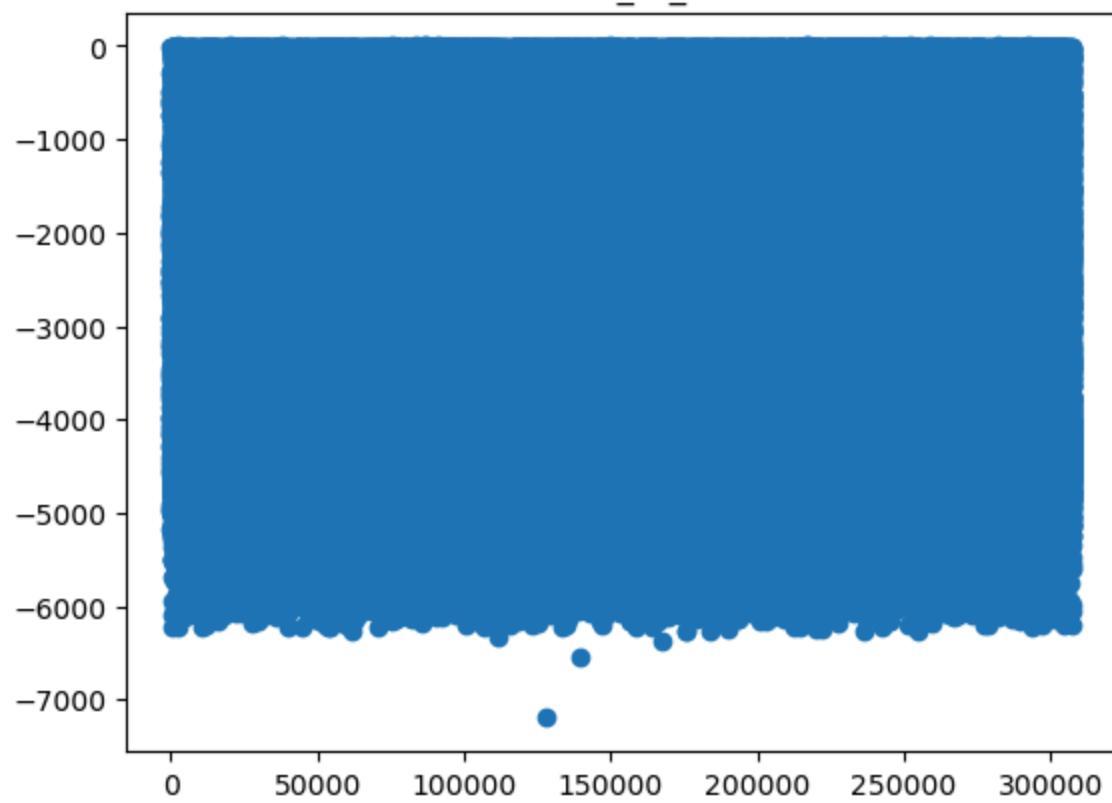
```
Out[82]: Text(0.5, 1.0, 'Plot of DAYS_REGISTRATION')
```



```
Out[82]: <matplotlib.collections.PathCollection at 0x22fdbdb6390>
```

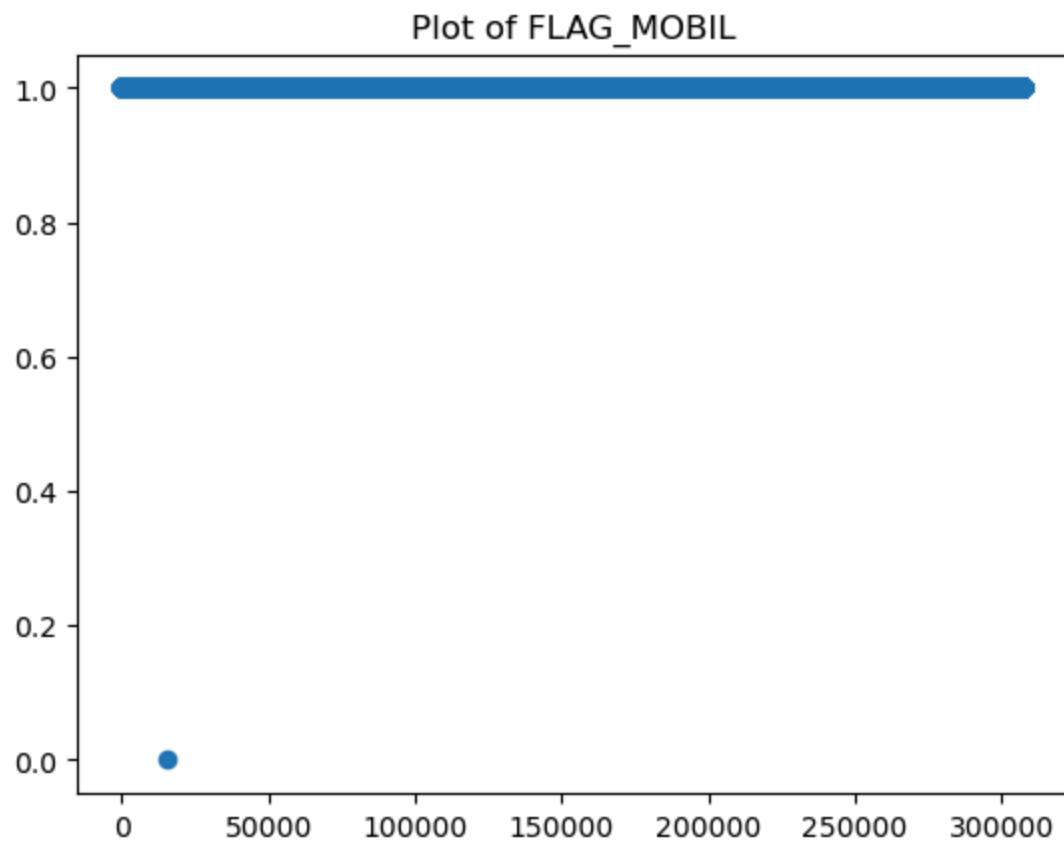
```
Out[82]: Text(0.5, 1.0, 'Plot of DAYS_ID_PUBLISH')
```

Plot of DAYS_ID_PUBLISH



Out[82]: <matplotlib.collections.PathCollection at 0x22fdbcf390>

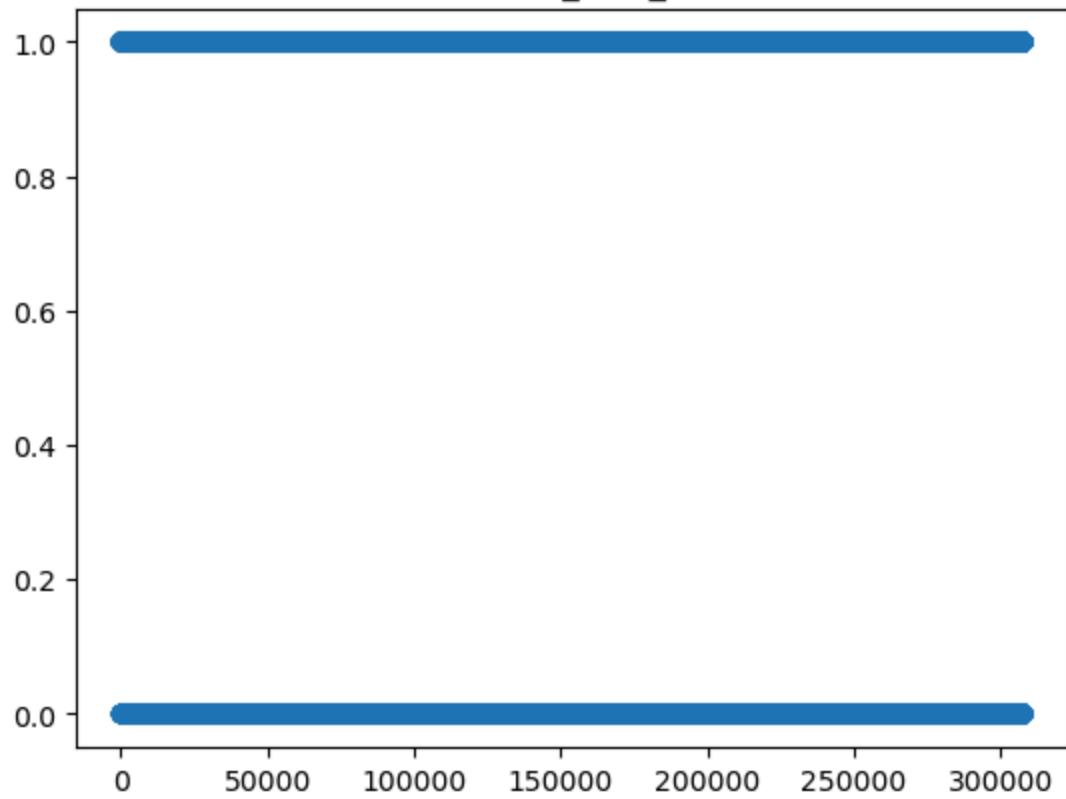
Out[82]: Text(0.5, 1.0, 'Plot of FLAG_MOBIL')



```
Out[82]: <matplotlib.collections.PathCollection at 0x22fdbce8690>
```

```
Out[82]: Text(0.5, 1.0, 'Plot of FLAG_EMP_PHONE')
```

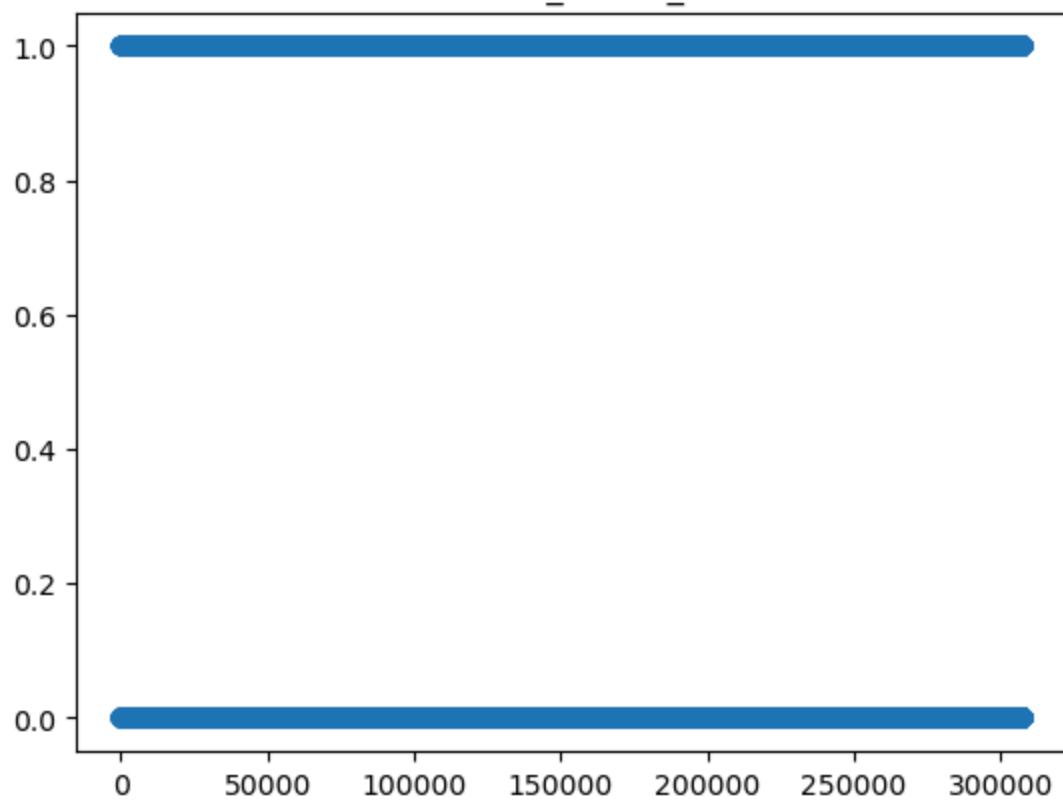
Plot of FLAG_EMP_PHONE



Out[82]: <matplotlib.collections.PathCollection at 0x22fdbcd4210>

Out[82]: Text(0.5, 1.0, 'Plot of FLAG_WORK_PHONE')

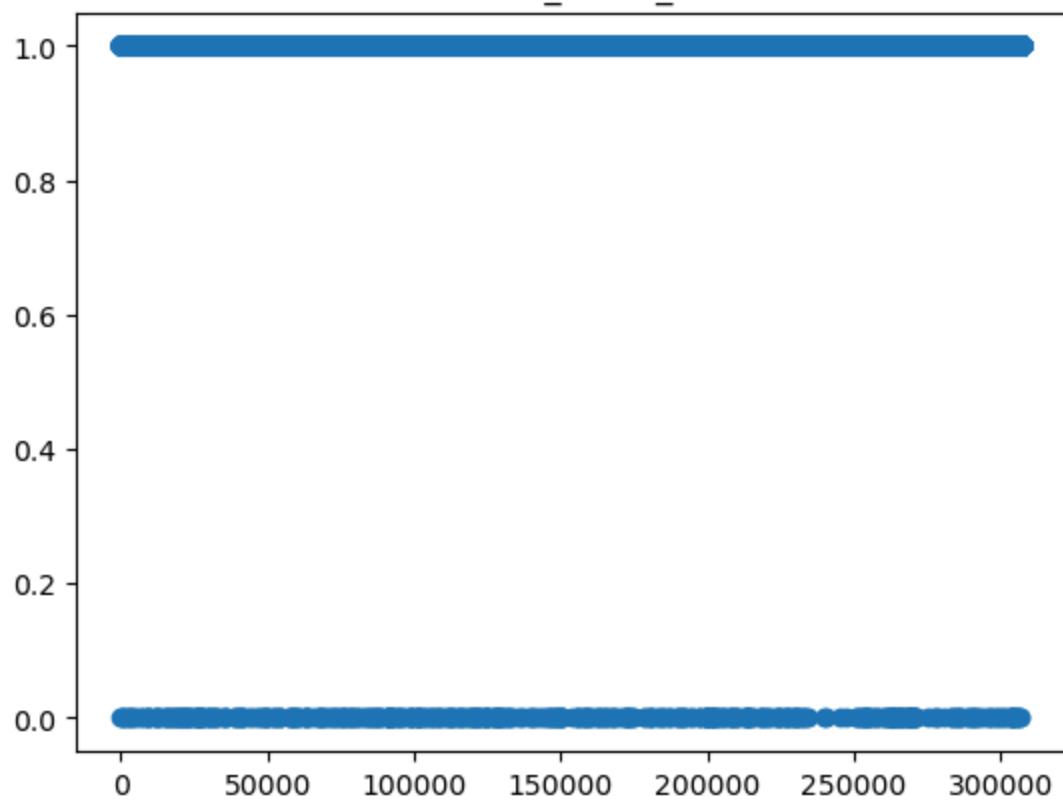
Plot of FLAG_WORK_PHONE



Out[82]: <matplotlib.collections.PathCollection at 0x22fdb39b50>

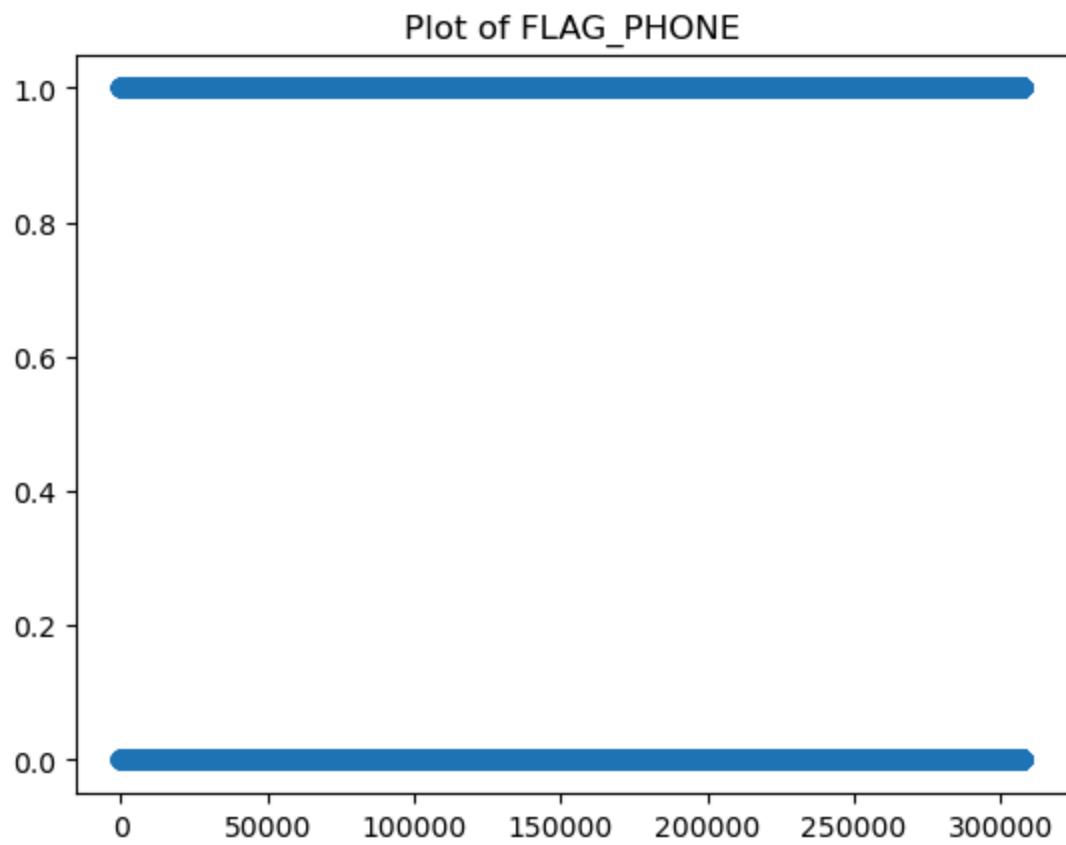
Out[82]: Text(0.5, 1.0, 'Plot of FLAG_CONT_MOBILE')

Plot of FLAG_CONT_MOBILE



Out[82]: <matplotlib.collections.PathCollection at 0x22fdbb74210>

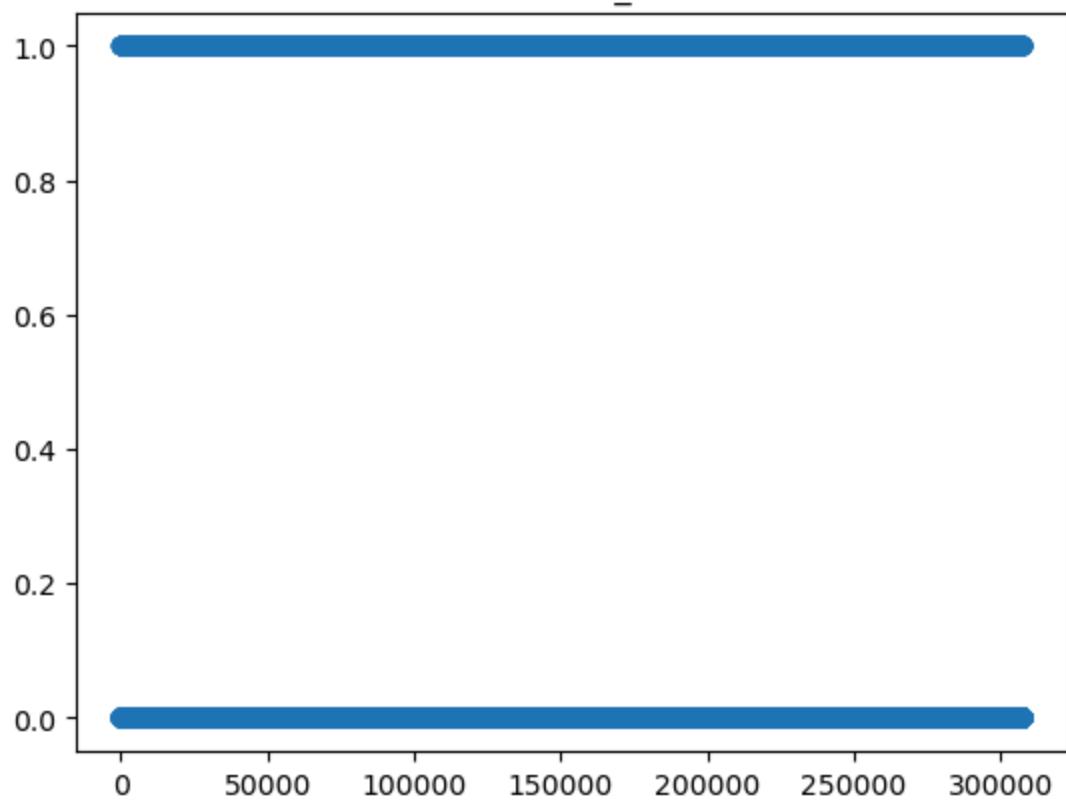
Out[82]: Text(0.5, 1.0, 'Plot of FLAG_PHONE')



```
Out[82]: <matplotlib.collections.PathCollection at 0x22fdbba21210>
```

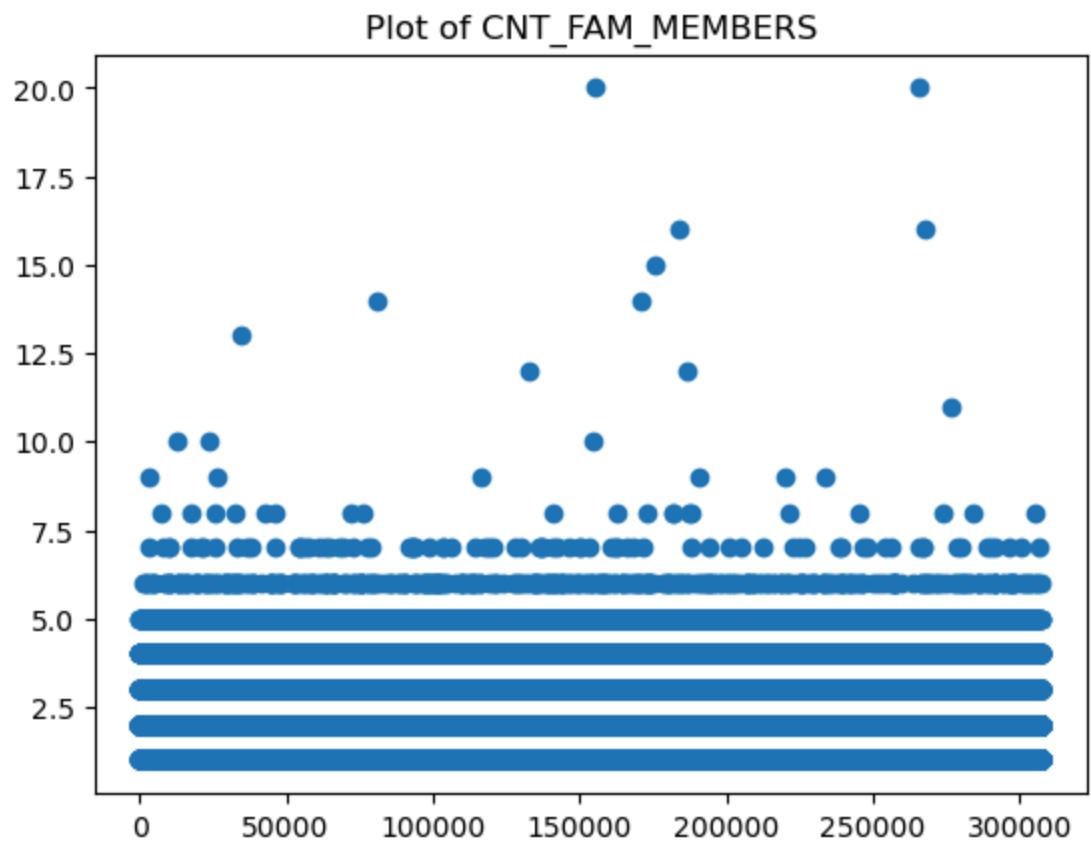
```
Out[82]: Text(0.5, 1.0, 'Plot of FLAG_EMAIL')
```

Plot of FLAG_EMAIL



Out[82]: <matplotlib.collections.PathCollection at 0x22fdc9befd0>

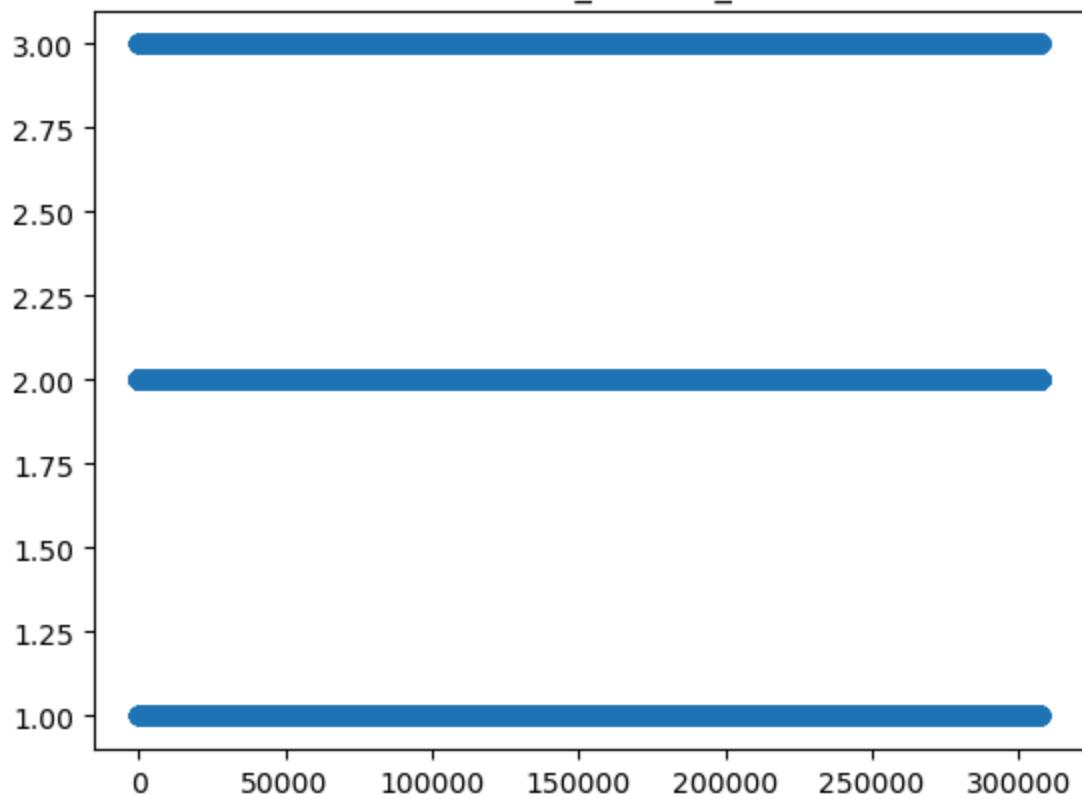
Out[82]: Text(0.5, 1.0, 'Plot of CNT_FAM_MEMBERS')



```
Out[82]: <matplotlib.collections.PathCollection at 0x22fdbb75090>
```

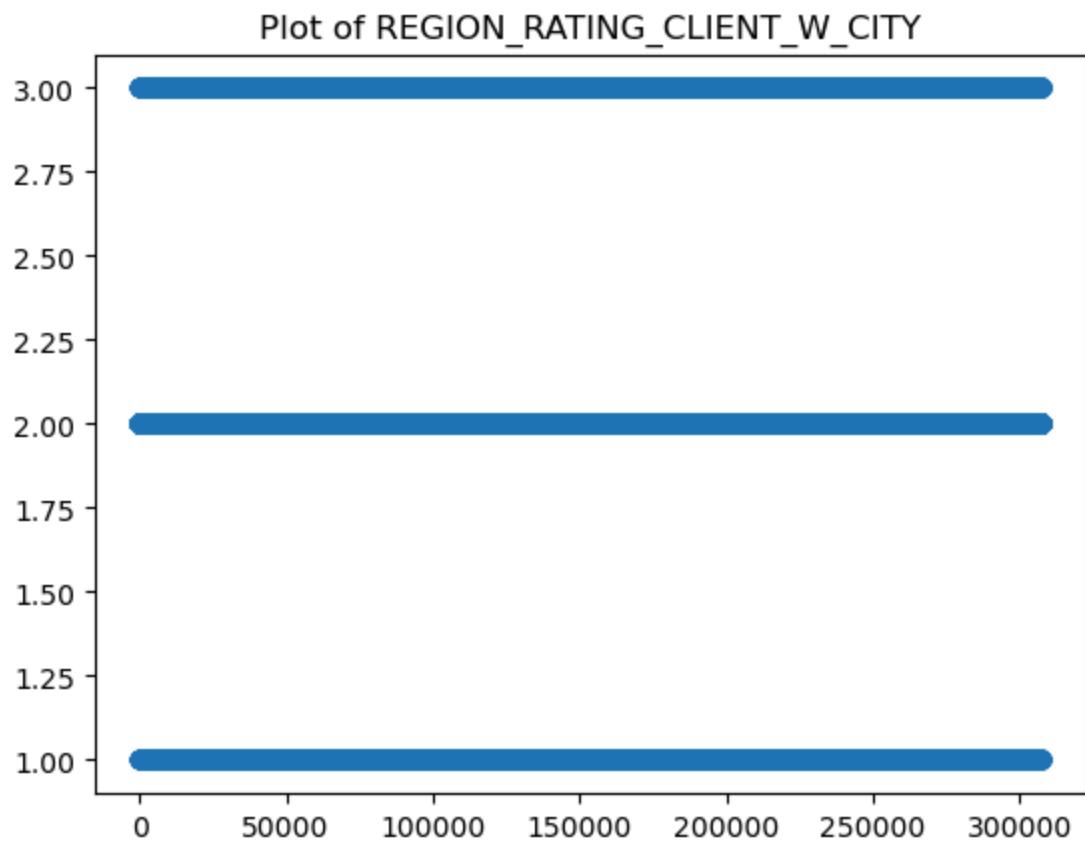
```
Out[82]: Text(0.5, 1.0, 'Plot of REGION_RATING_CLIENT')
```

Plot of REGION_RATING_CLIENT



Out[82]: <matplotlib.collections.PathCollection at 0x22f9ca654d0>

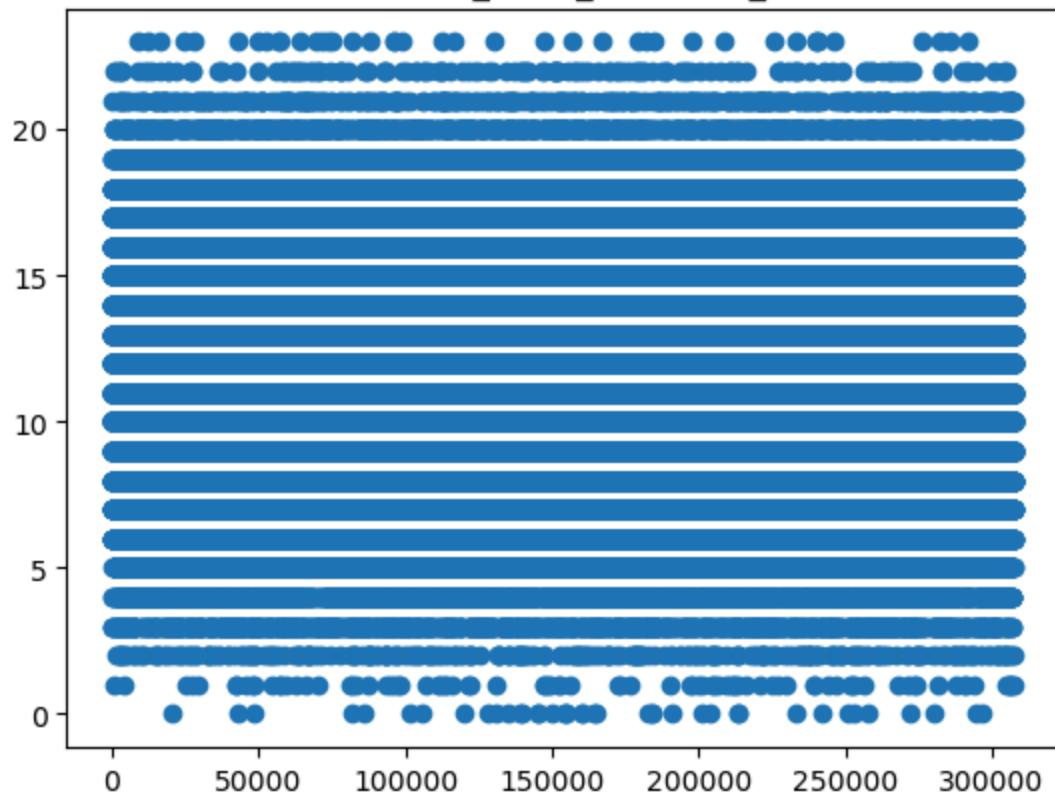
Out[82]: Text(0.5, 1.0, 'Plot of REGION_RATING_CLIENT_W_CITY')



```
Out[82]: <matplotlib.collections.PathCollection at 0x22fcec31350>
```

```
Out[82]: Text(0.5, 1.0, 'Plot of HOUR_APPR_PROCESS_START')
```

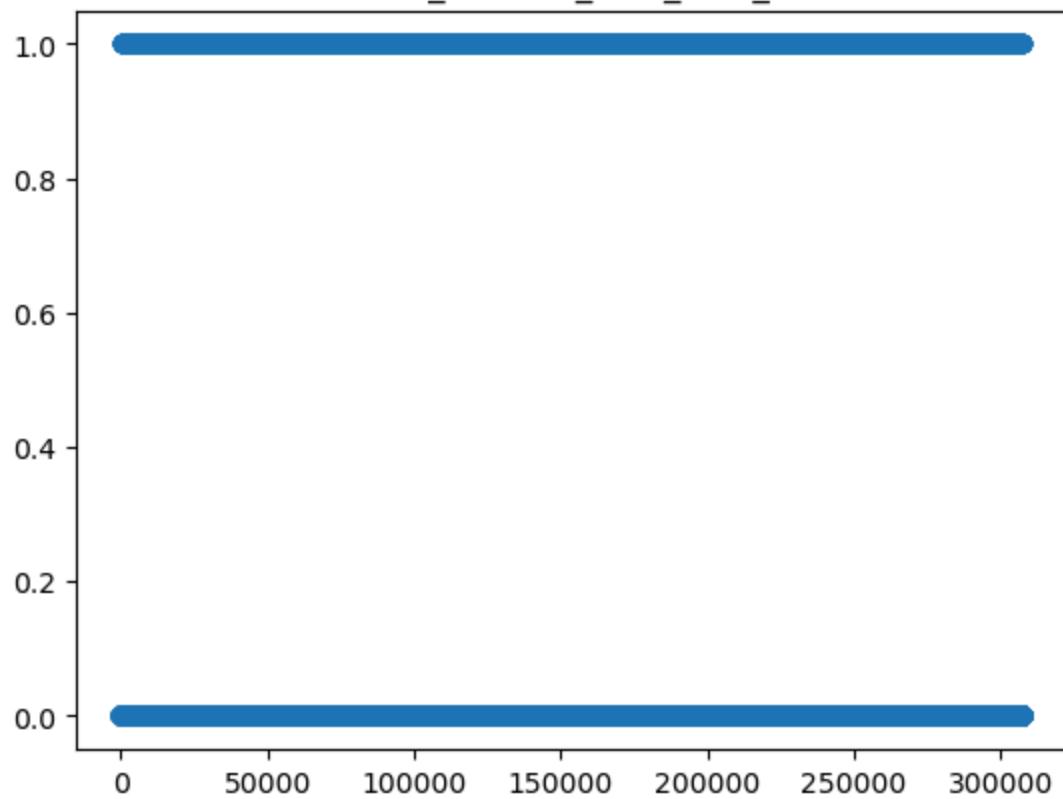
Plot of HOUR_APPR_PROCESS_START



Out[82]: <matplotlib.collections.PathCollection at 0x22fcec7d2d0>

Out[82]: Text(0.5, 1.0, 'Plot of REG_REGION_NOT_LIVE_REGION')

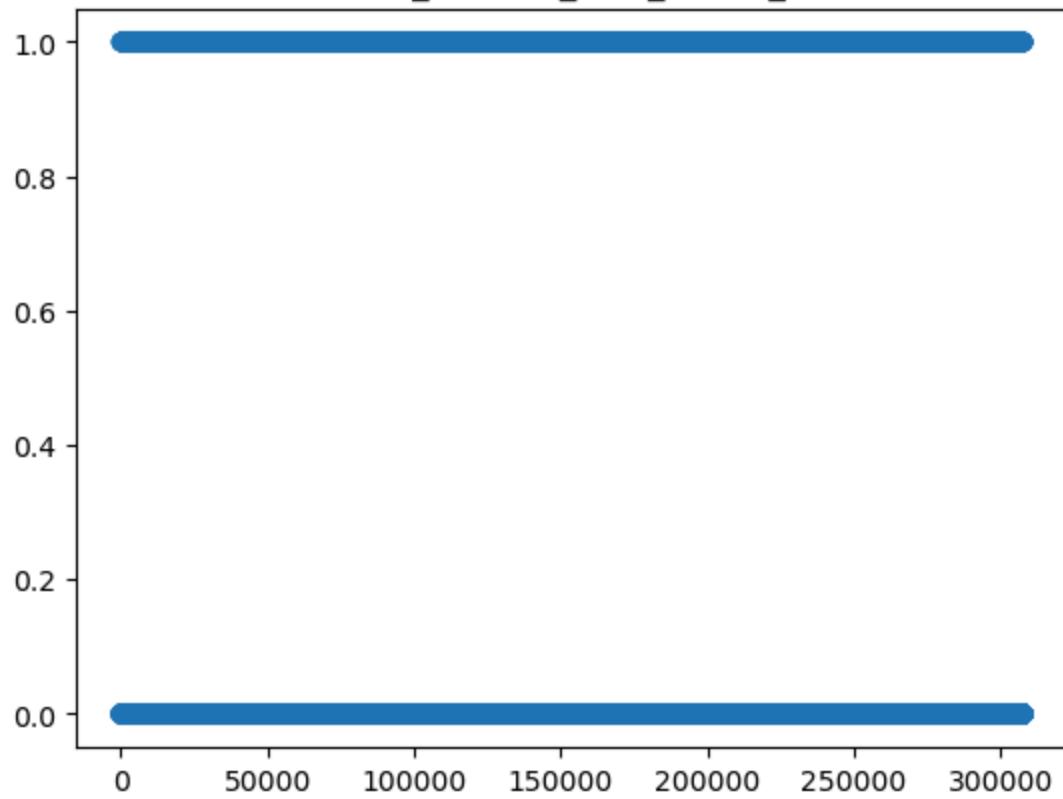
Plot of REG_REGION_NOT_LIVE_REGION



Out[82]: <matplotlib.collections.PathCollection at 0x22f8ce74210>

Out[82]: Text(0.5, 1.0, 'Plot of REG_REGION_NOT_WORK_REGION')

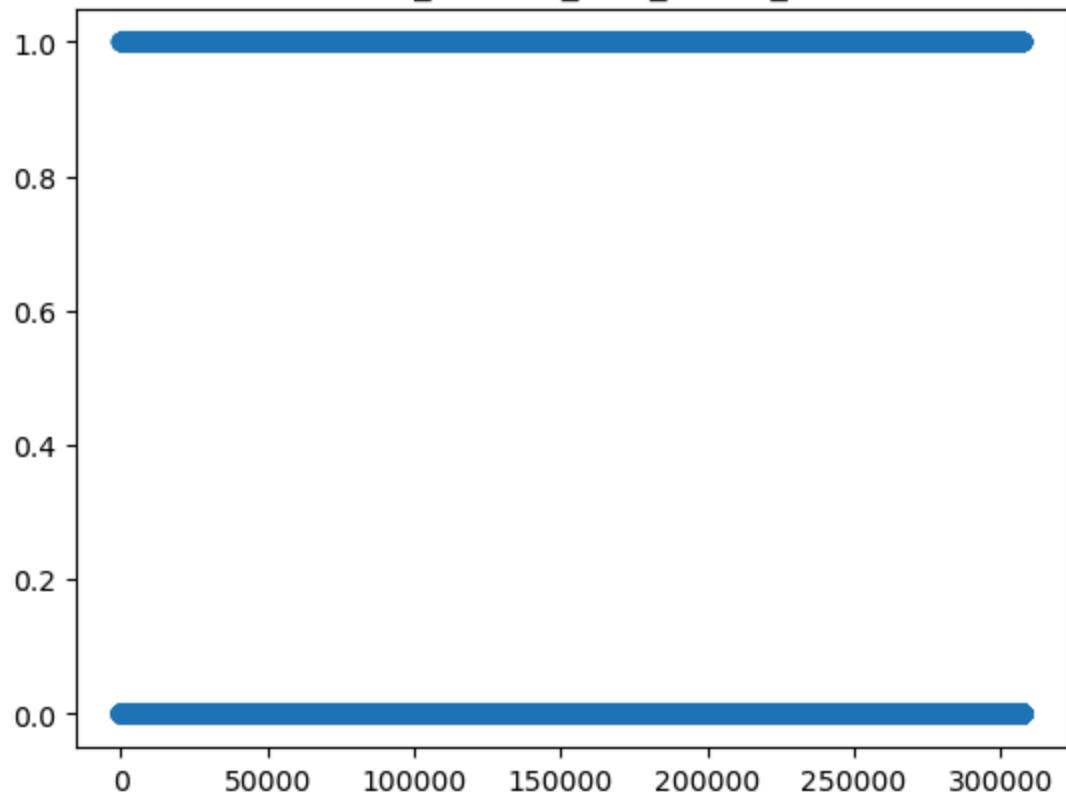
Plot of REG_REGION_NOT_WORK_REGION



Out[82]: <matplotlib.collections.PathCollection at 0x22fdcb4e390>

Out[82]: Text(0.5, 1.0, 'Plot of LIVE_REGION_NOT_WORK_REGION')

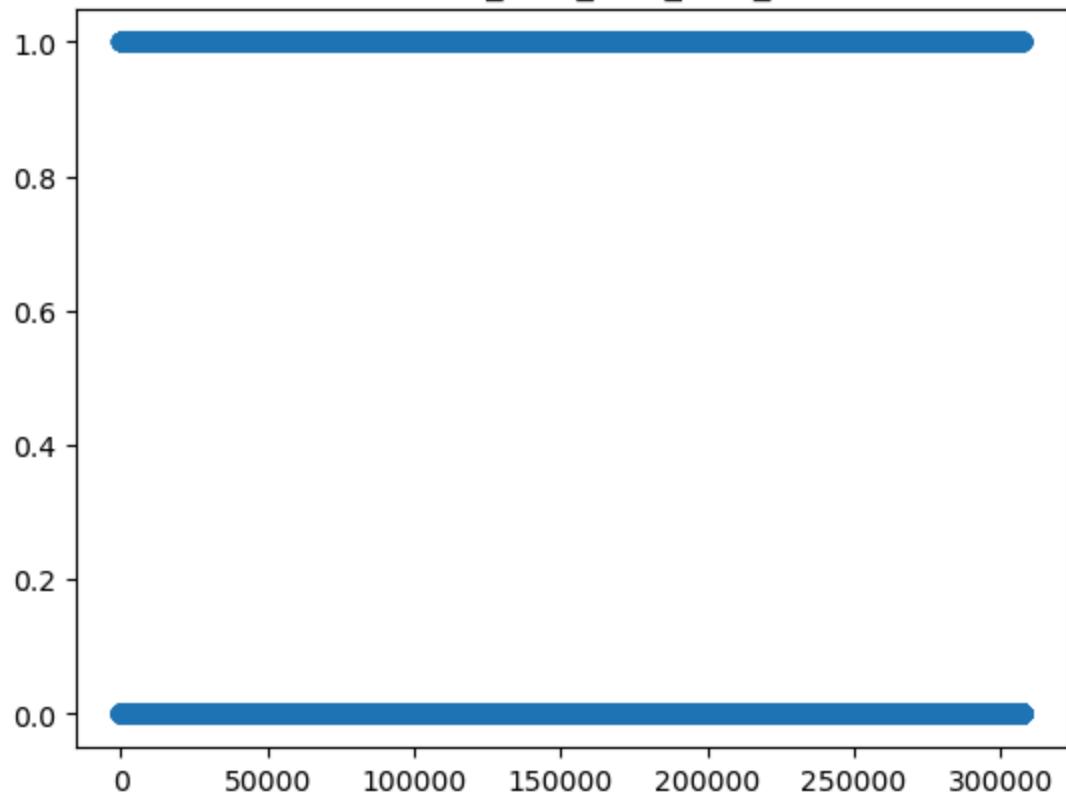
Plot of LIVE_REGION_NOT_WORK_REGION



Out[82]: <matplotlib.collections.PathCollection at 0x22fdc8c0a90>

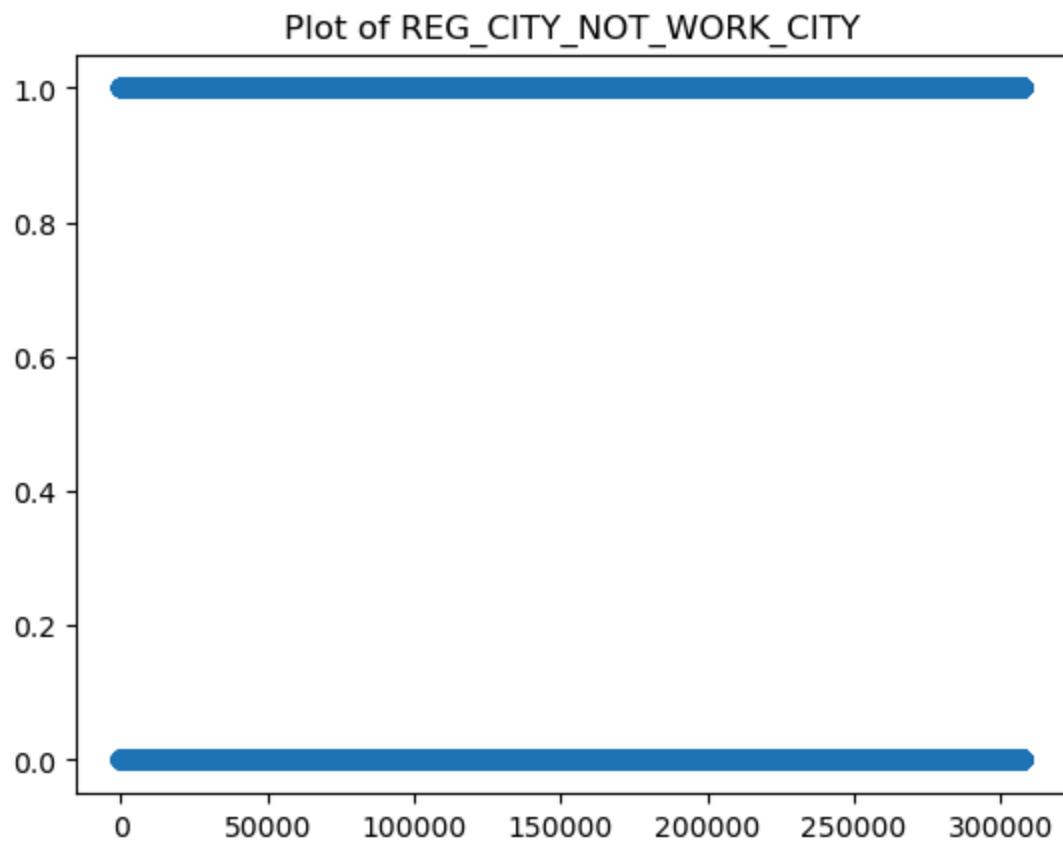
Out[82]: Text(0.5, 1.0, 'Plot of REG_CITY_NOT_LIVE_CITY')

Plot of REG_CITY_NOT_LIVE_CITY



Out[82]: <matplotlib.collections.PathCollection at 0x22fdcee7850>

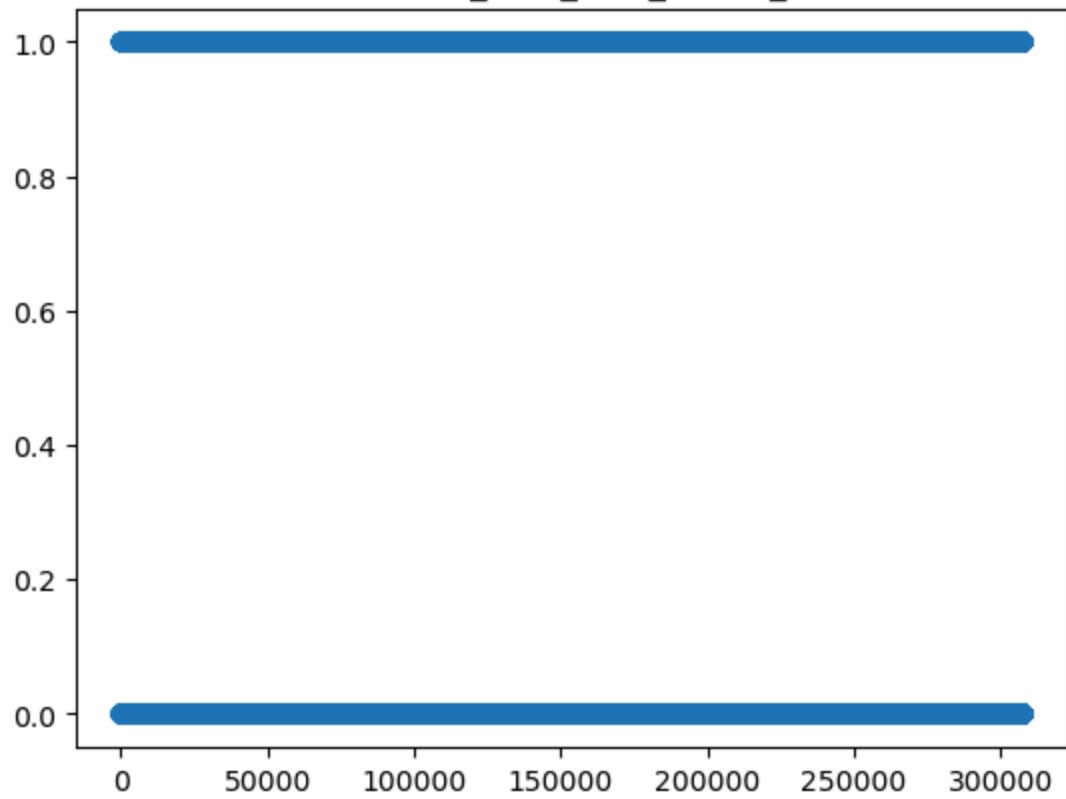
Out[82]: Text(0.5, 1.0, 'Plot of REG_CITY_NOT_WORK_CITY')



```
Out[82]: <matplotlib.collections.PathCollection at 0x22fdce14990>
```

```
Out[82]: Text(0.5, 1.0, 'Plot of LIVE_CITY_NOT_WORK_CITY')
```

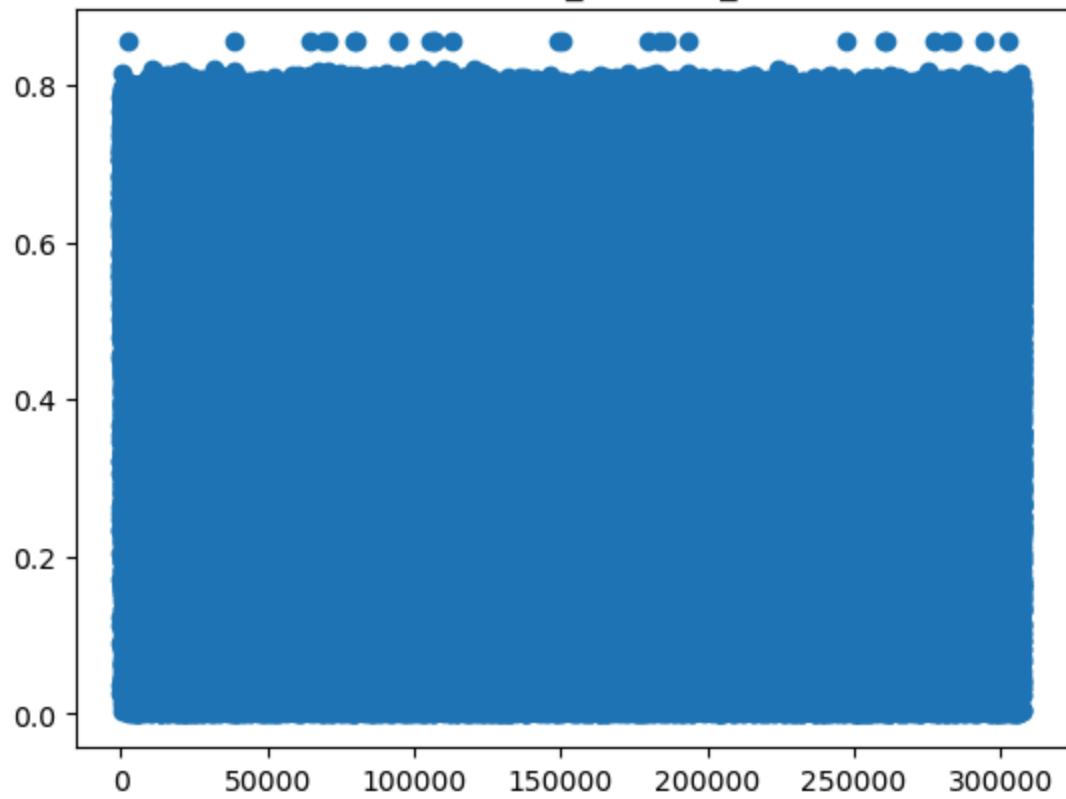
Plot of LIVE_CITY_NOT_WORK_CITY



Out[82]: <matplotlib.collections.PathCollection at 0x22fdca5c4d0>

Out[82]: Text(0.5, 1.0, 'Plot of EXT_SOURCE_2')

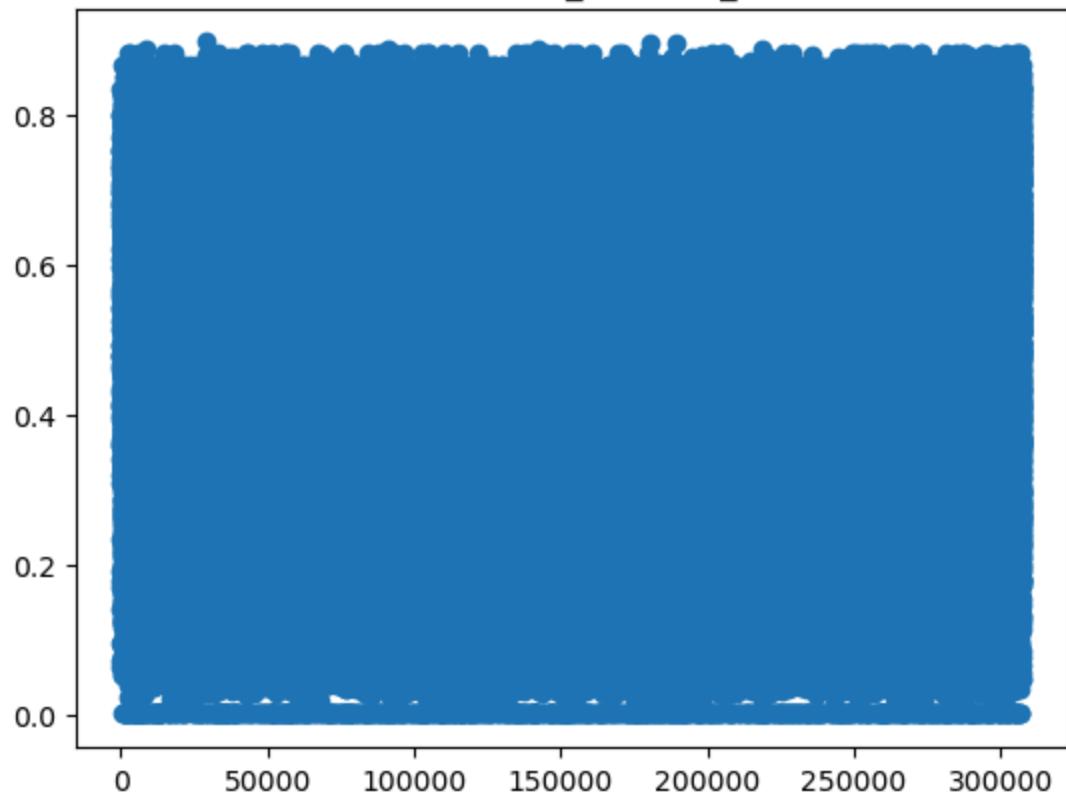
Plot of EXT_SOURCE_2



Out[82]: <matplotlib.collections.PathCollection at 0x22fdc9a3ad0>

Out[82]: Text(0.5, 1.0, 'Plot of EXT_SOURCE_3')

Plot of EXT_SOURCE_3



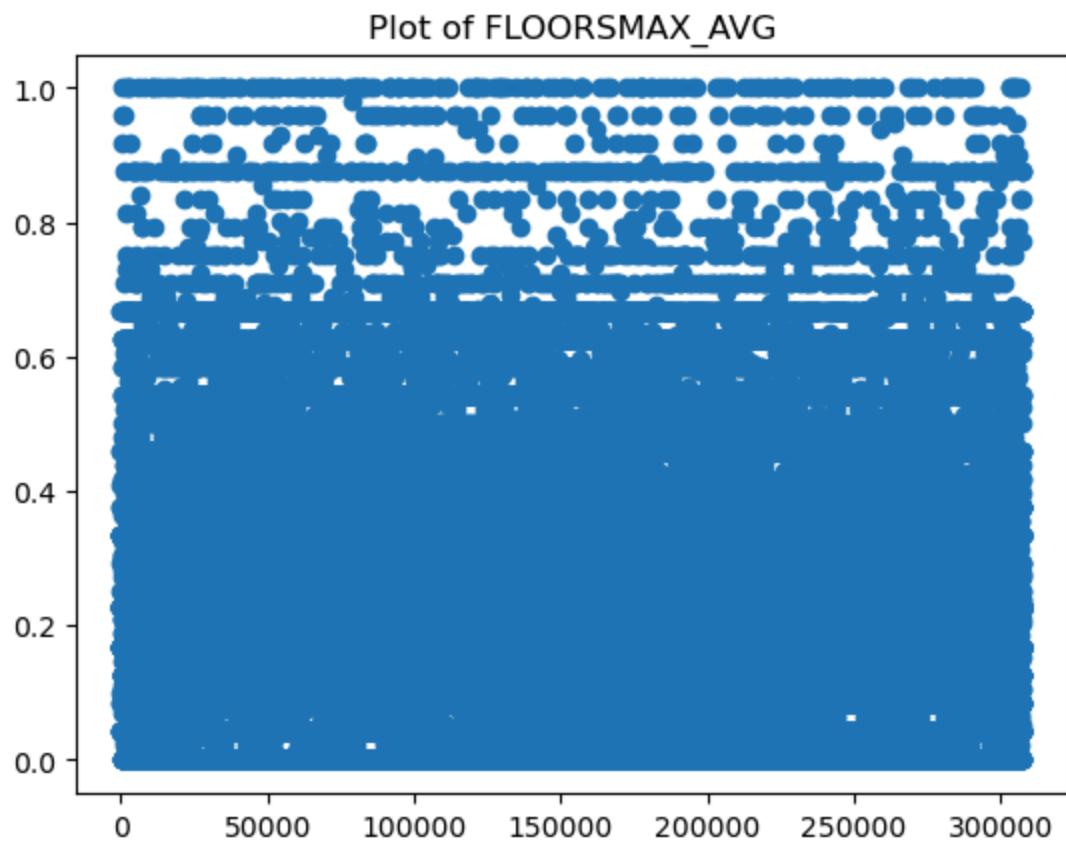
Out[82]: <matplotlib.collections.PathCollection at 0x22fdc949190>

Out[82]: Text(0.5, 1.0, 'Plot of YEARS_BEGINEXPLUATATION_AVG')



```
Out[82]: <matplotlib.collections.PathCollection at 0x22fdce82e10>
```

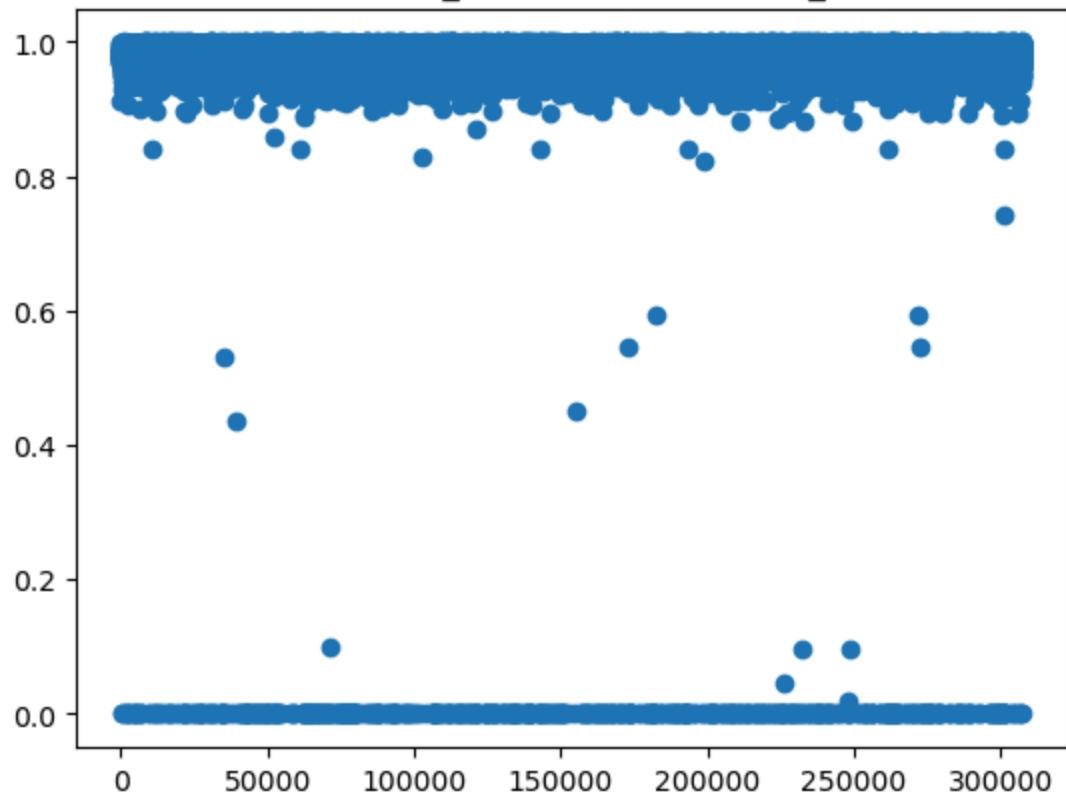
```
Out[82]: Text(0.5, 1.0, 'Plot of FLOORSMAX_AVG')
```



```
Out[82]: <matplotlib.collections.PathCollection at 0x22fdbbbe0190>
```

```
Out[82]: Text(0.5, 1.0, 'Plot of YEARS_BEGINEXPLUATATION_MODE')
```

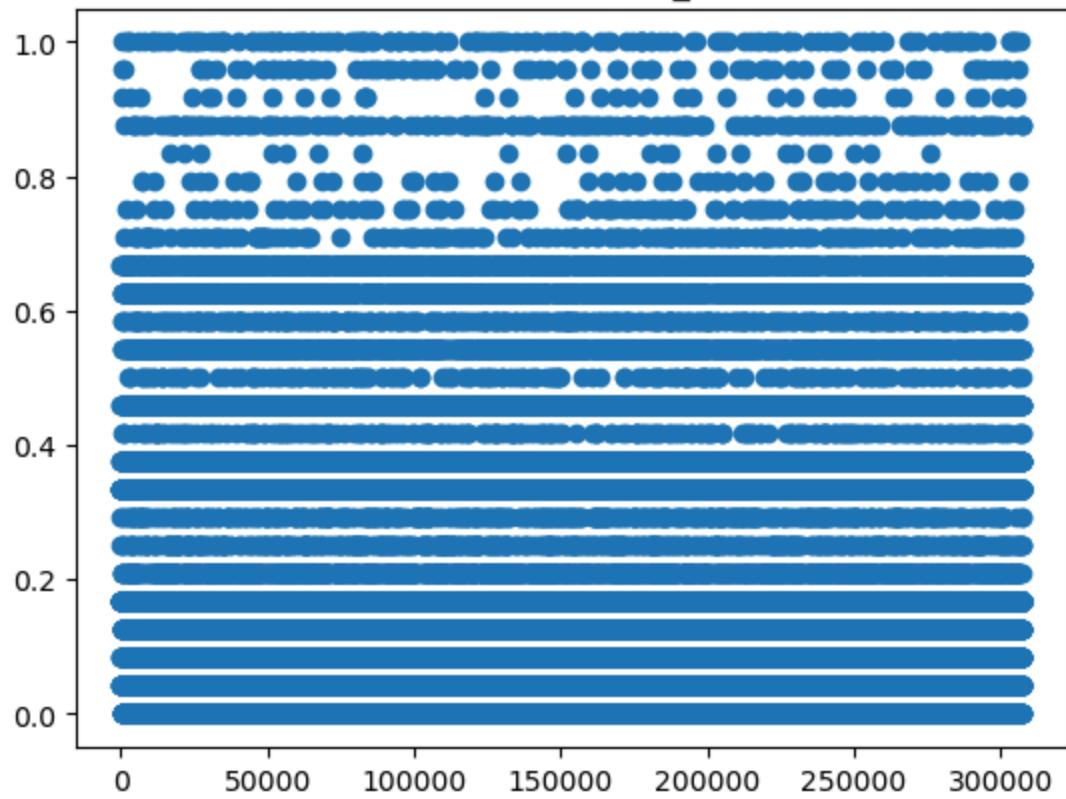
Plot of YEARS_BEGINEXPLUATATION_MODE



Out[82]: <matplotlib.collections.PathCollection at 0x22ffa0e8a90>

Out[82]: Text(0.5, 1.0, 'Plot of FLOORSMAX_MODE')

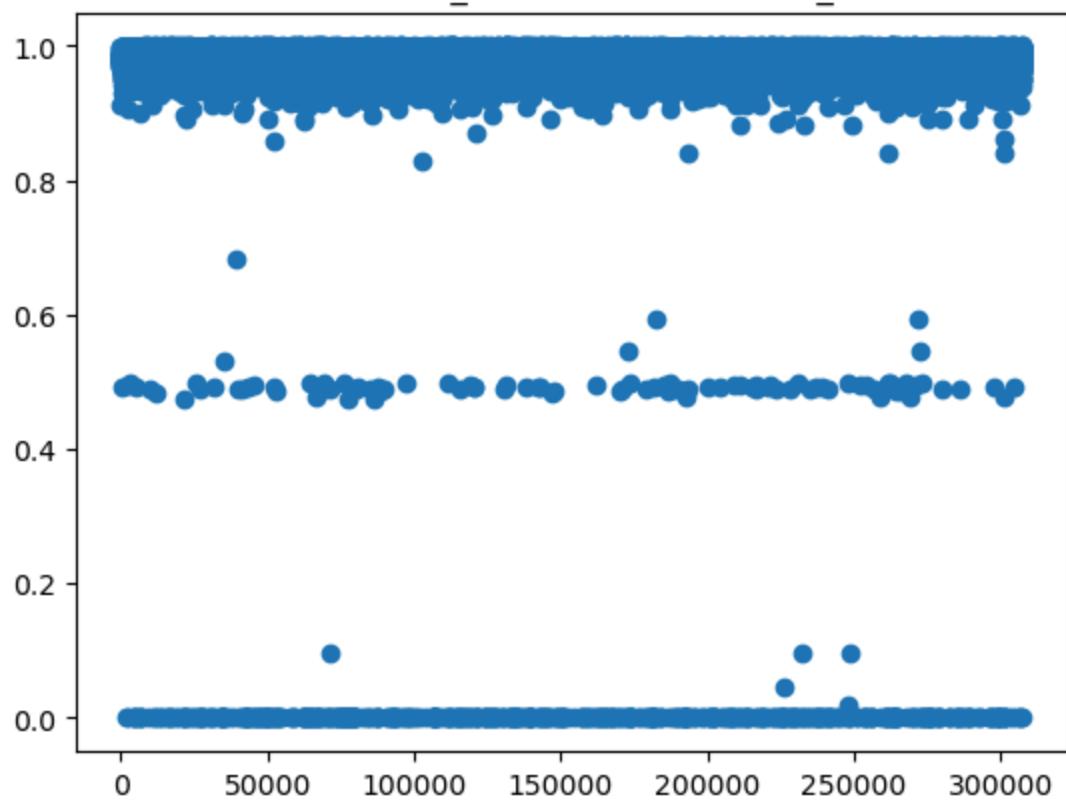
Plot of FLOORSMAX_MODE



Out[82]: <matplotlib.collections.PathCollection at 0x22fdcc97450>

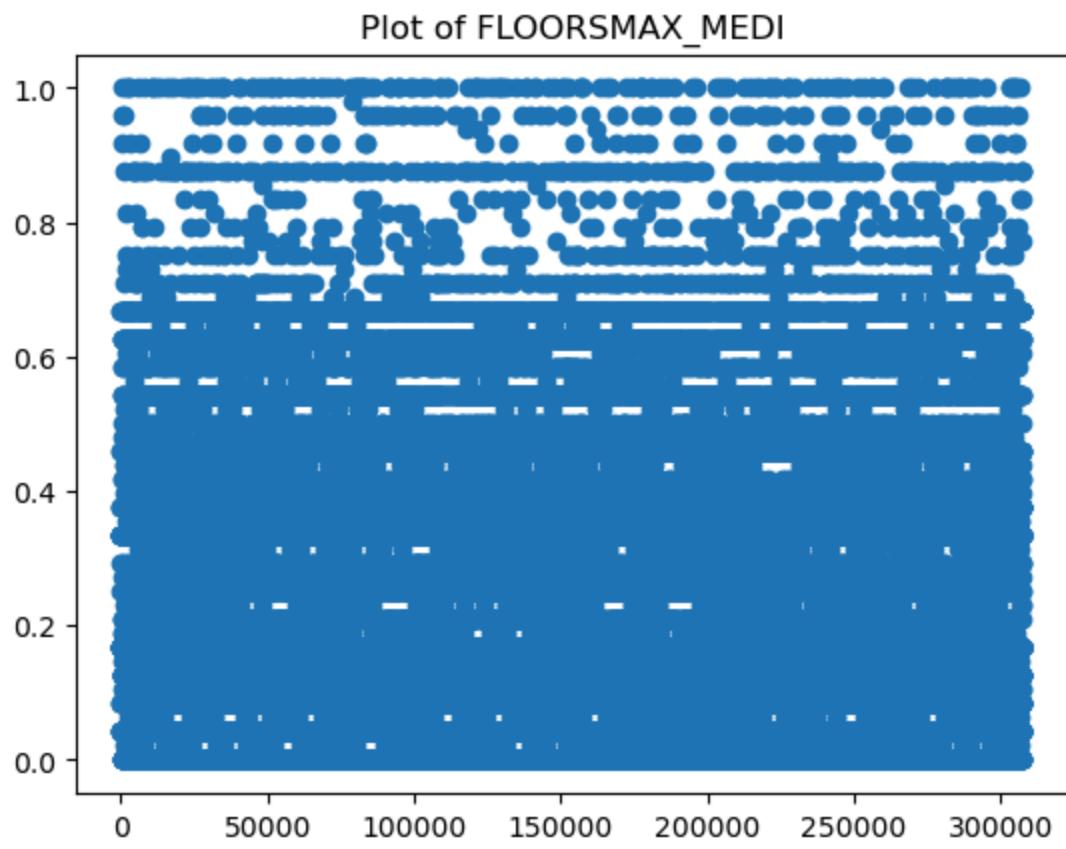
Out[82]: Text(0.5, 1.0, 'Plot of YEARS_BEGINEXPLUATATION_MEDI')

Plot of YEARS_BEGINEXPLUATATION_MEDI



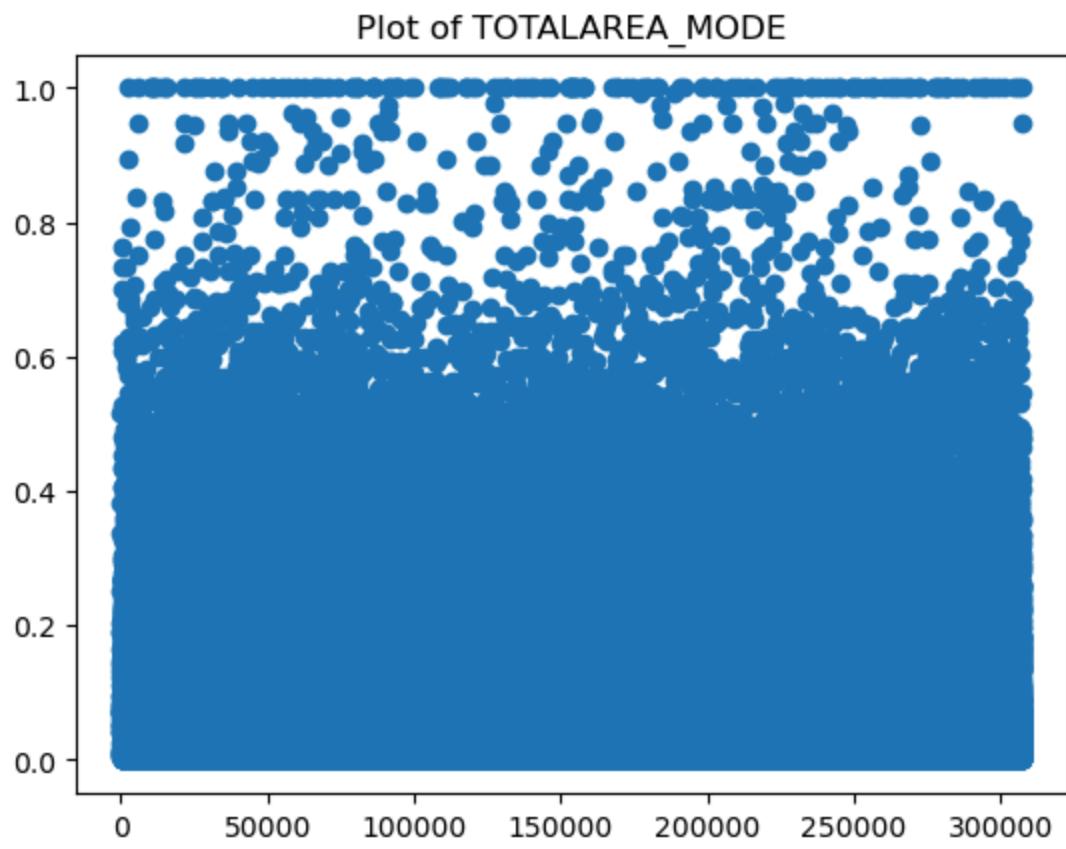
Out[82]: <matplotlib.collections.PathCollection at 0x22ffa0c1010>

Out[82]: Text(0.5, 1.0, 'Plot of FLOORSMAX_MEDI')



```
Out[82]: <matplotlib.collections.PathCollection at 0x22fdca45090>
```

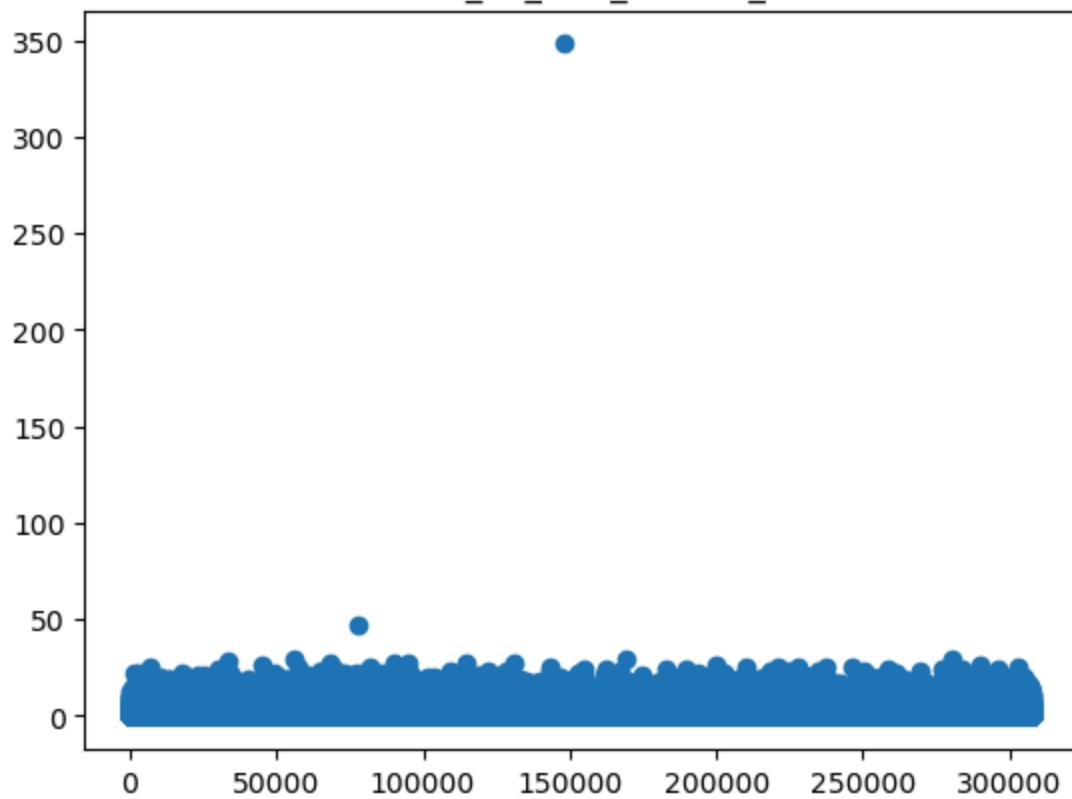
```
Out[82]: Text(0.5, 1.0, 'Plot of TOTALAREA_MODE')
```



```
Out[82]: <matplotlib.collections.PathCollection at 0x22fdcaa6390>
```

```
Out[82]: Text(0.5, 1.0, 'Plot of OBS_30_CNT_SOCIAL_CIRCLE')
```

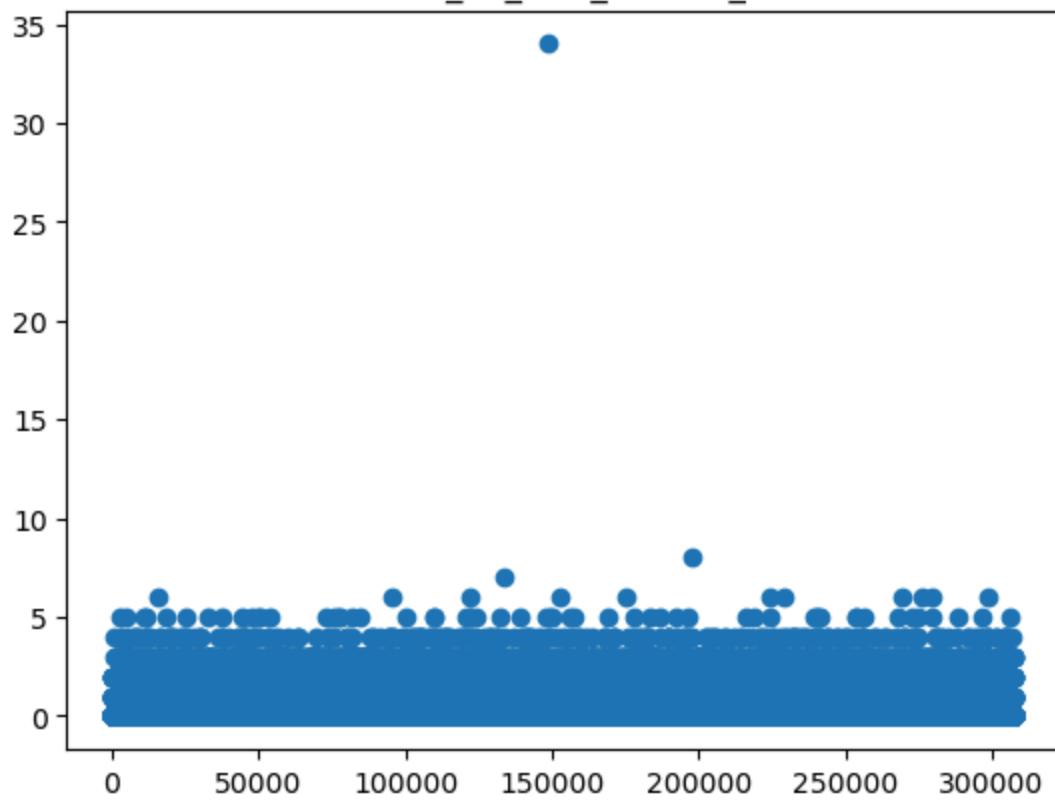
Plot of OBS_30_CNT_SOCIAL_CIRCLE



Out[82]: <matplotlib.collections.PathCollection at 0x22fdcaac850>

Out[82]: Text(0.5, 1.0, 'Plot of DEF_30_CNT_SOCIAL_CIRCLE')

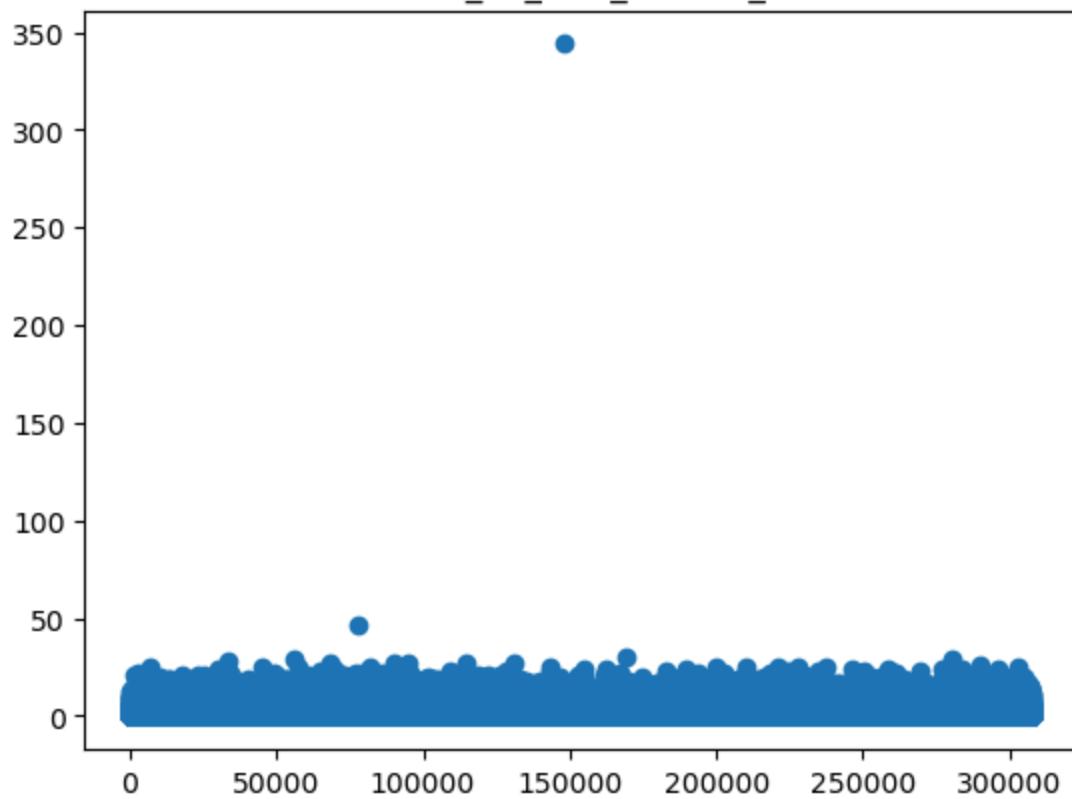
Plot of DEF_30_CNT_SOCIAL_CIRCLE



Out[82]: <matplotlib.collections.PathCollection at 0x22fdcb76790>

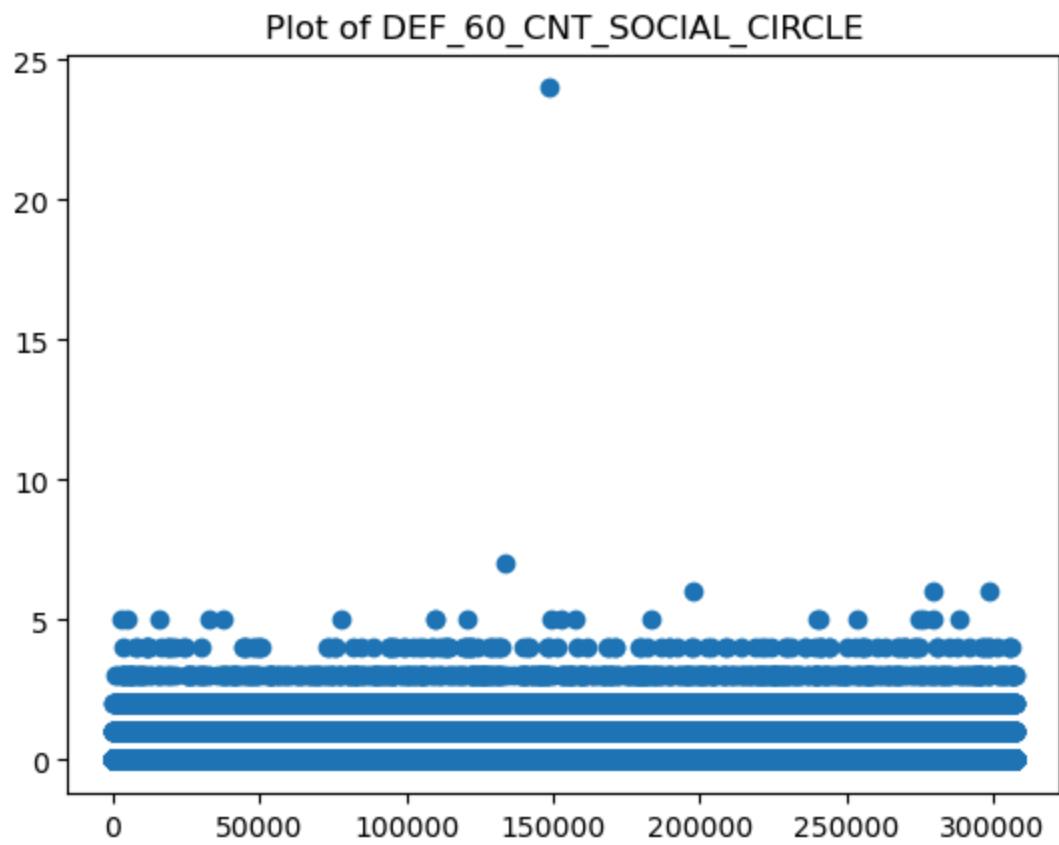
Out[82]: Text(0.5, 1.0, 'Plot of OBS_60_CNT_SOCIAL_CIRCLE')

Plot of OBS_60_CNT_SOCIAL_CIRCLE



Out[82]: <matplotlib.collections.PathCollection at 0x22fdce063d0>

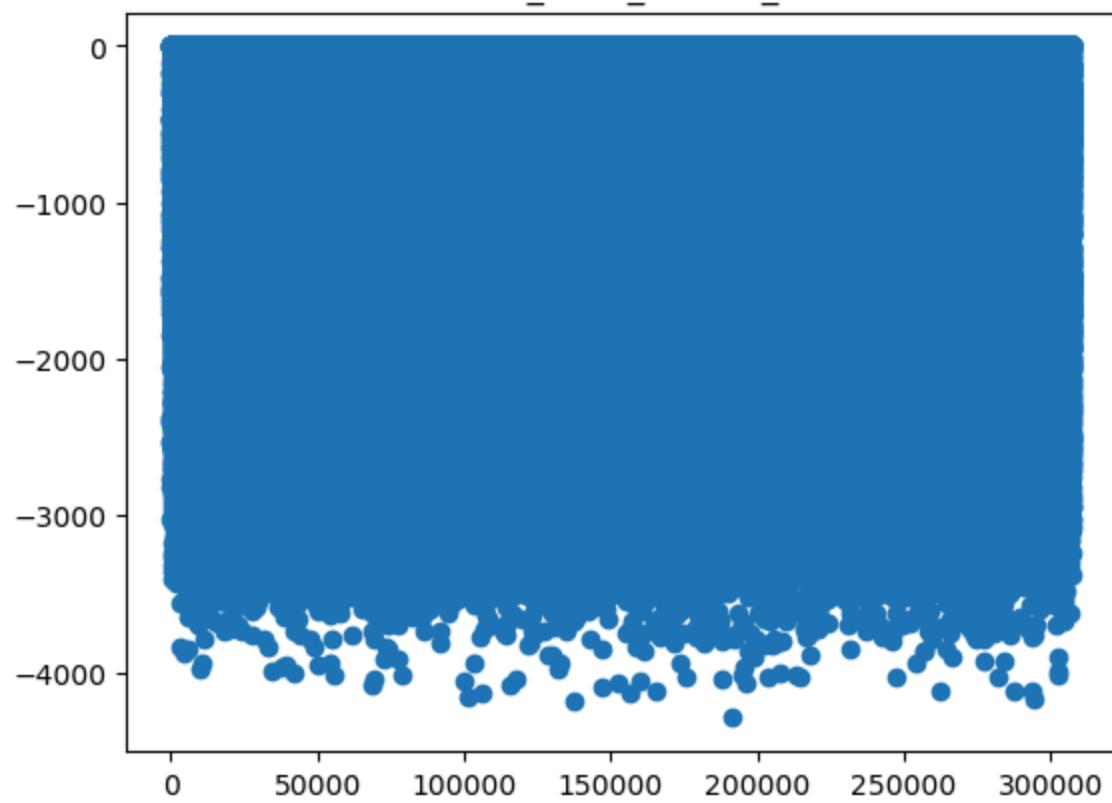
Out[82]: Text(0.5, 1.0, 'Plot of DEF_60_CNT_SOCIAL_CIRCLE')



```
Out[82]: <matplotlib.collections.PathCollection at 0x22fdadb2fd0>
```

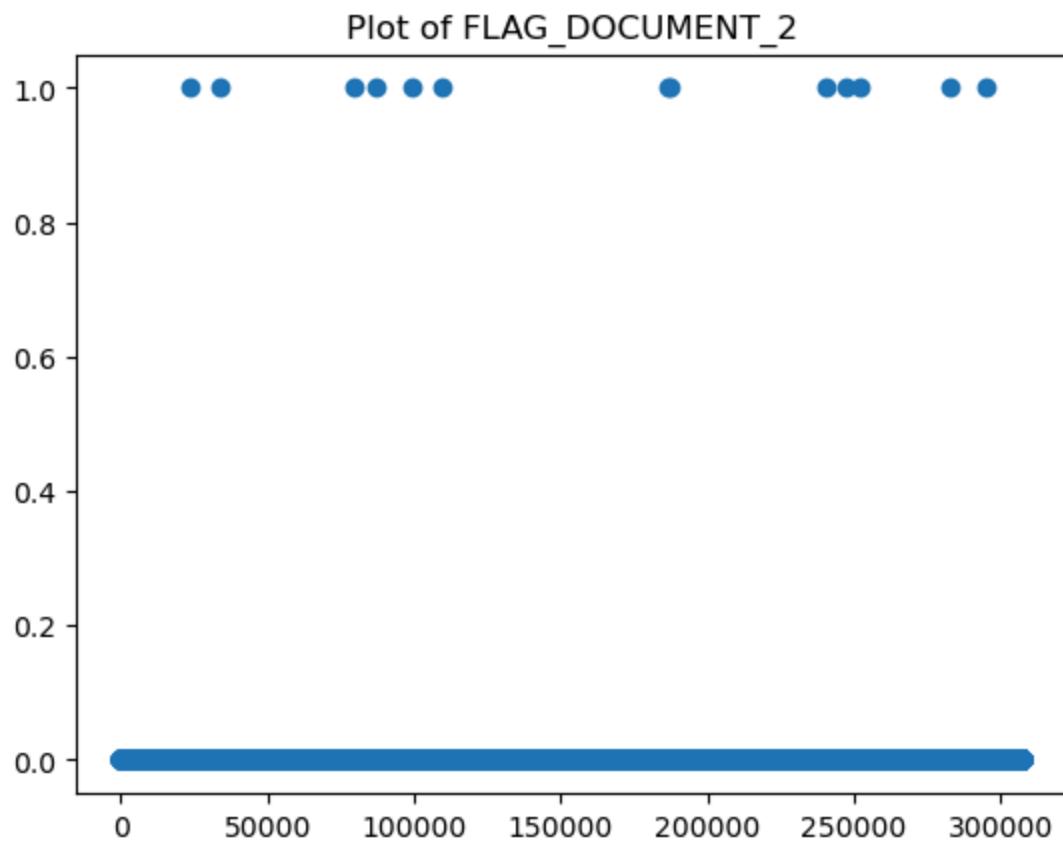
```
Out[82]: Text(0.5, 1.0, 'Plot of DAYS_LAST_PHONE_CHANGE')
```

Plot of DAYS_LAST_PHONE_CHANGE



Out[82]: <matplotlib.collections.PathCollection at 0x22fceb18b50>

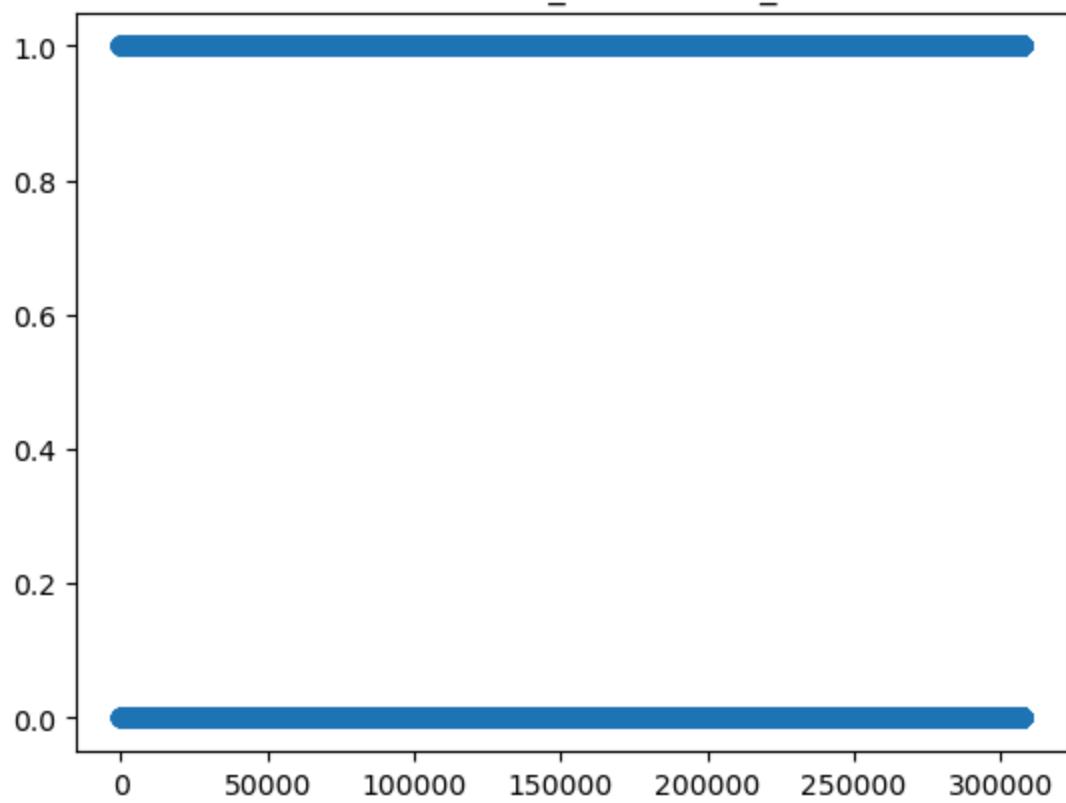
Out[82]: Text(0.5, 1.0, 'Plot of FLAG_DOCUMENT_2')



```
Out[82]: <matplotlib.collections.PathCollection at 0x22fcec19890>
```

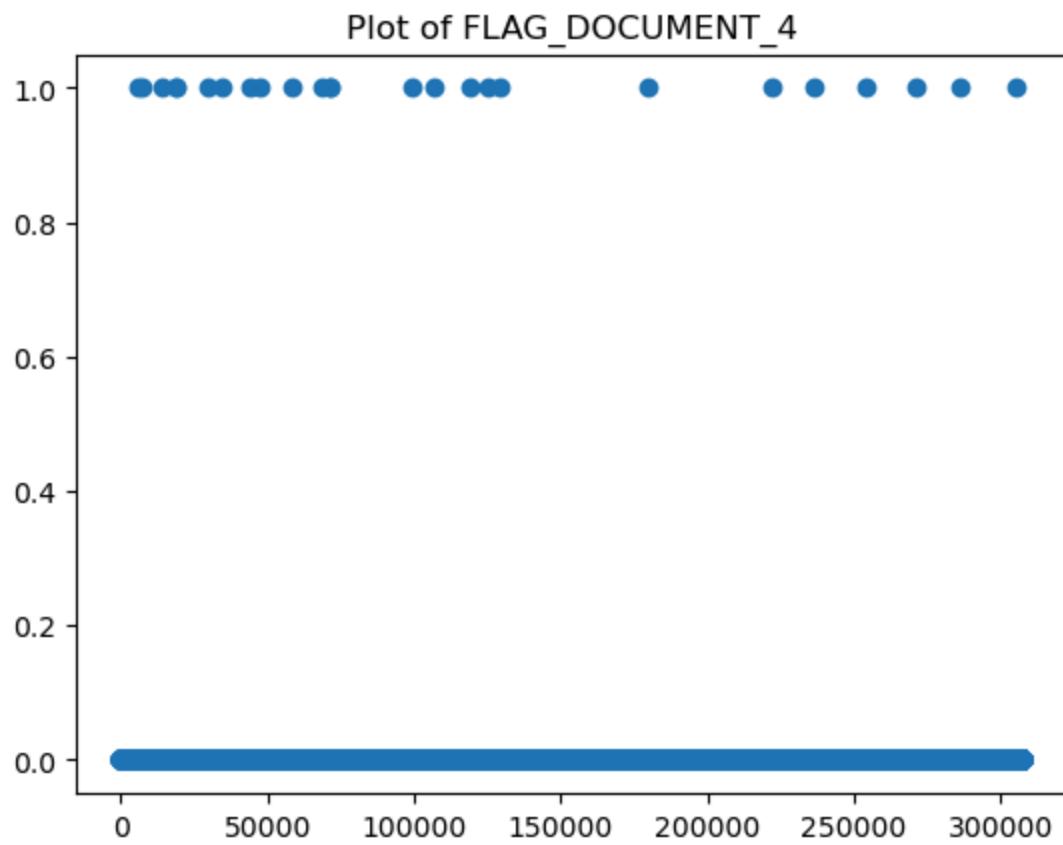
```
Out[82]: Text(0.5, 1.0, 'Plot of FLAG_DOCUMENT_3')
```

Plot of FLAG_DOCUMENT_3



Out[82]: <matplotlib.collections.PathCollection at 0x22fda7275d0>

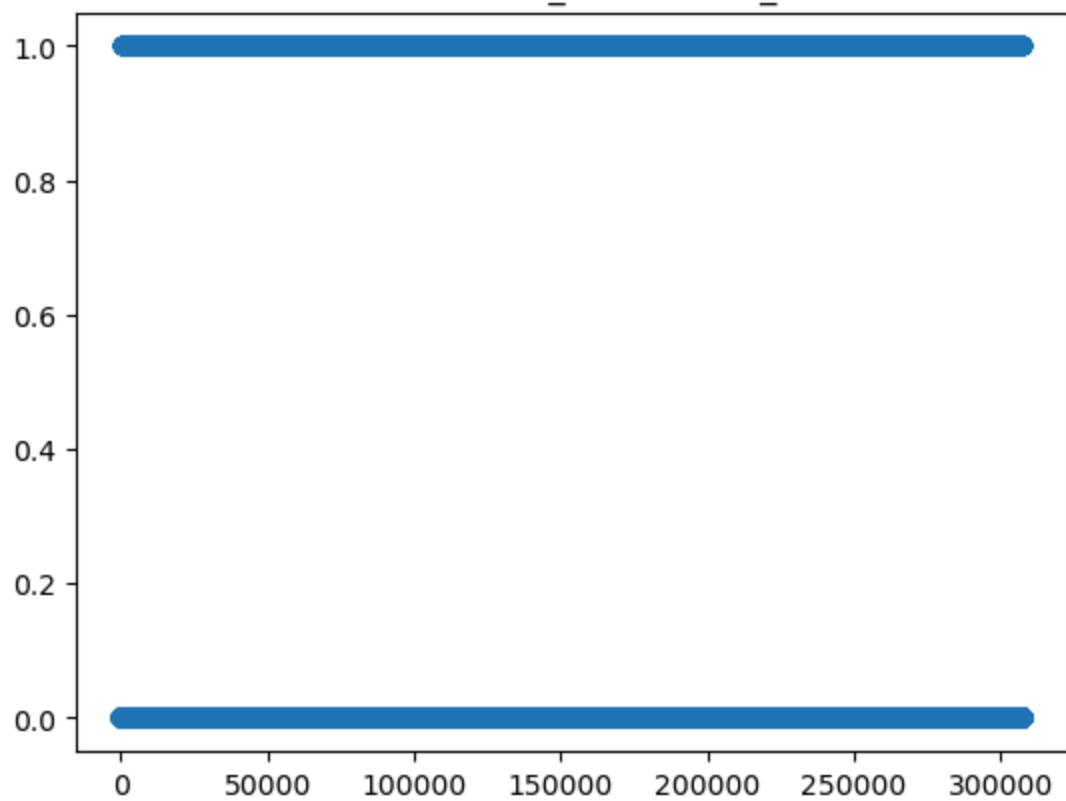
Out[82]: Text(0.5, 1.0, 'Plot of FLAG_DOCUMENT_4')



```
Out[82]: <matplotlib.collections.PathCollection at 0x22fceb7c510>
```

```
Out[82]: Text(0.5, 1.0, 'Plot of FLAG_DOCUMENT_5')
```

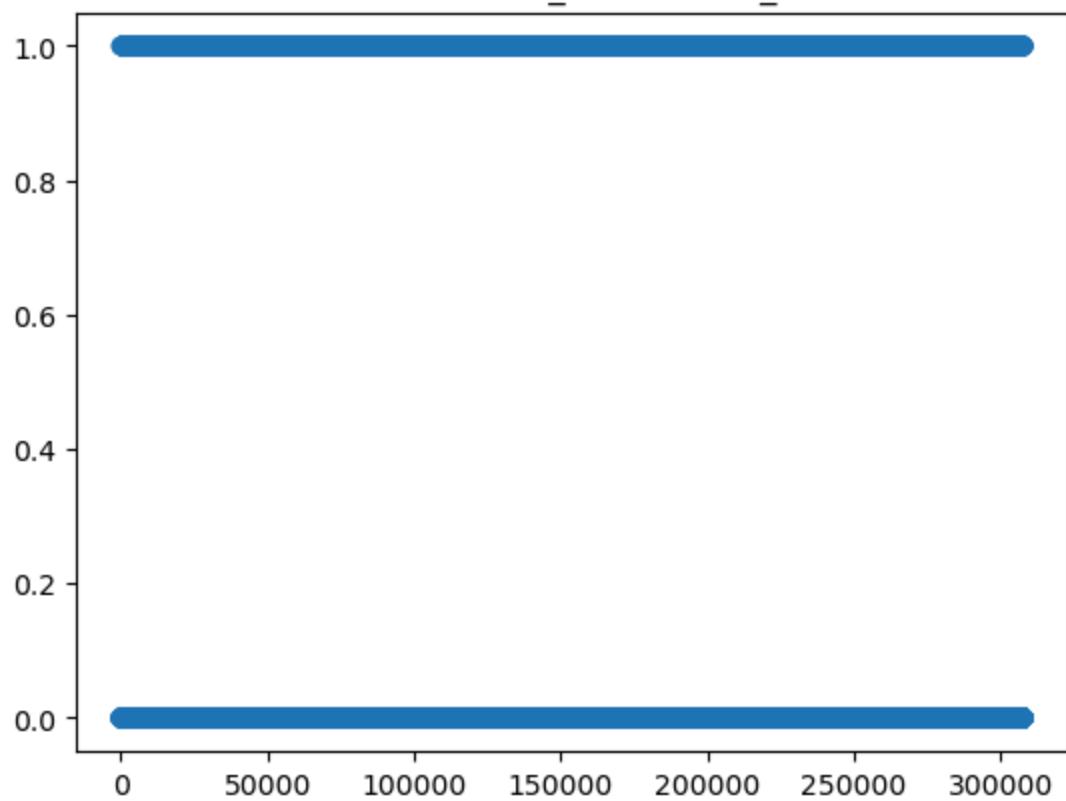
Plot of FLAG_DOCUMENT_5



Out[82]: <matplotlib.collections.PathCollection at 0x22fdc919ed0>

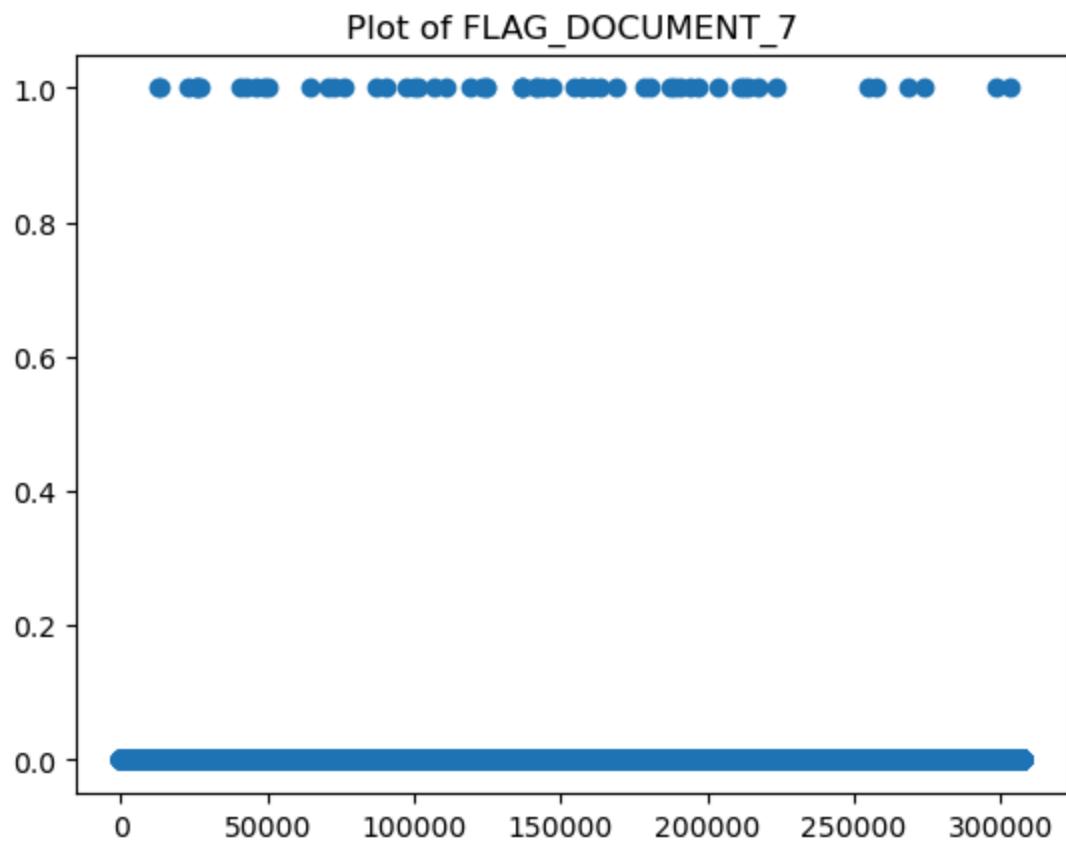
Out[82]: Text(0.5, 1.0, 'Plot of FLAG_DOCUMENT_6')

Plot of FLAG_DOCUMENT_6



Out[82]: <matplotlib.collections.PathCollection at 0x22fcebef5d0>

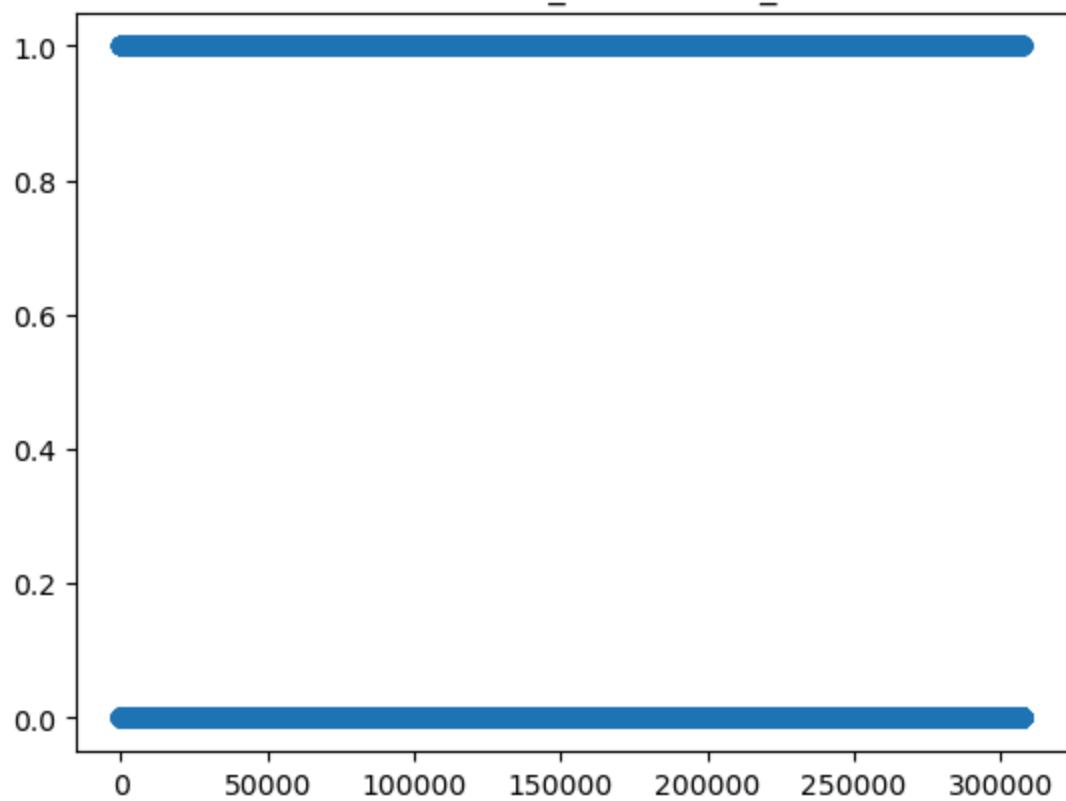
Out[82]: Text(0.5, 1.0, 'Plot of FLAG_DOCUMENT_7')



```
Out[82]: <matplotlib.collections.PathCollection at 0x22fdbb2ba10>
```

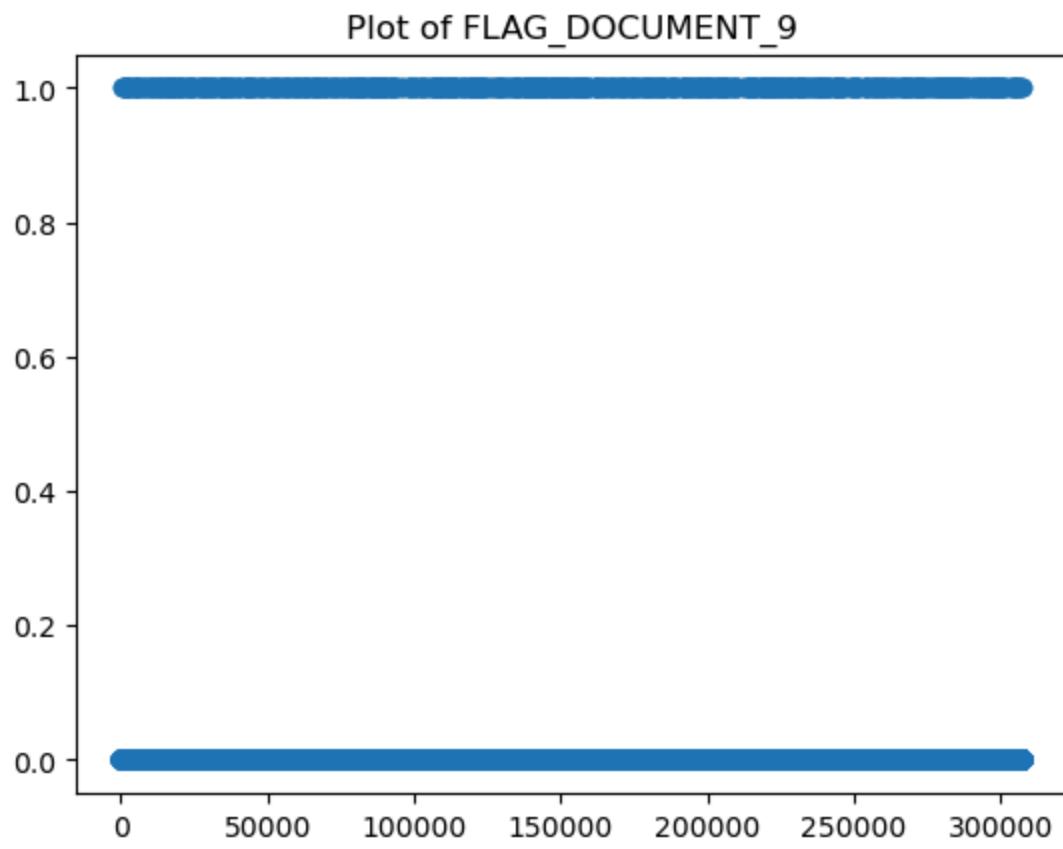
```
Out[82]: Text(0.5, 1.0, 'Plot of FLAG_DOCUMENT_8')
```

Plot of FLAG_DOCUMENT_8



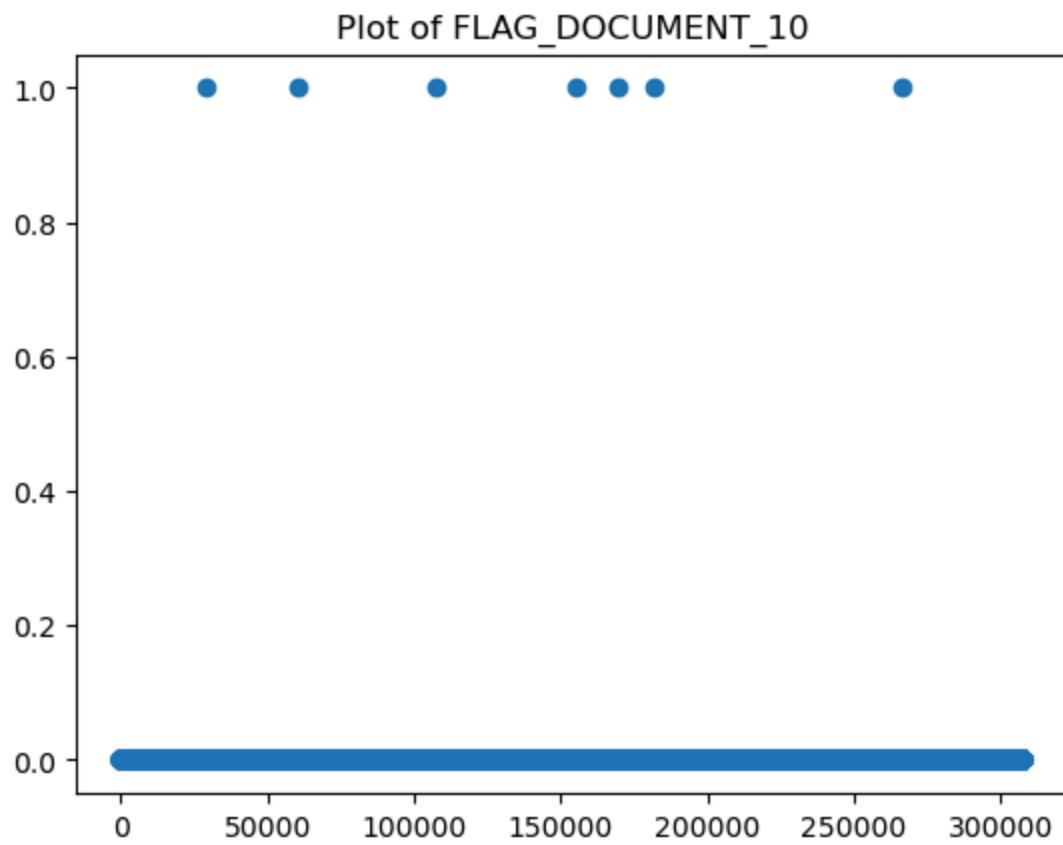
Out[82]: <matplotlib.collections.PathCollection at 0x22fdbc0bf10>

Out[82]: Text(0.5, 1.0, 'Plot of FLAG_DOCUMENT_9')



```
Out[82]: <matplotlib.collections.PathCollection at 0x22ffa05c210>
```

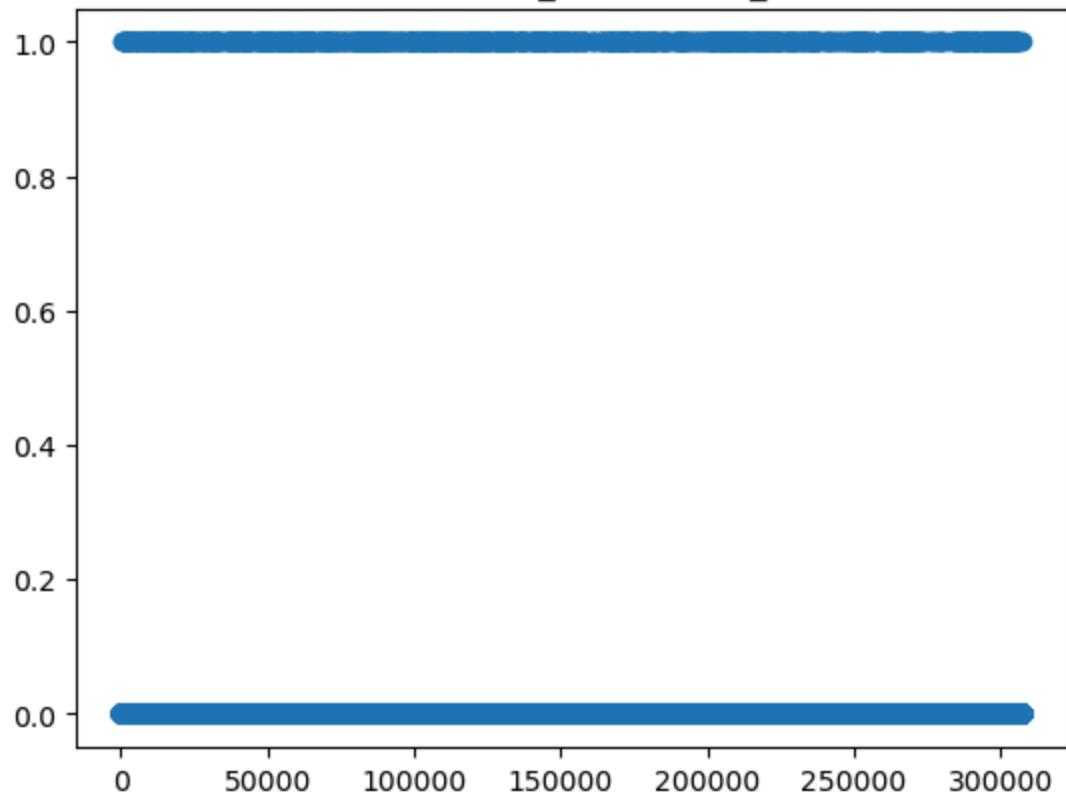
```
Out[82]: Text(0.5, 1.0, 'Plot of FLAG_DOCUMENT_10')
```



```
Out[82]: <matplotlib.collections.PathCollection at 0x22ffa02d110>
```

```
Out[82]: Text(0.5, 1.0, 'Plot of FLAG_DOCUMENT_11')
```

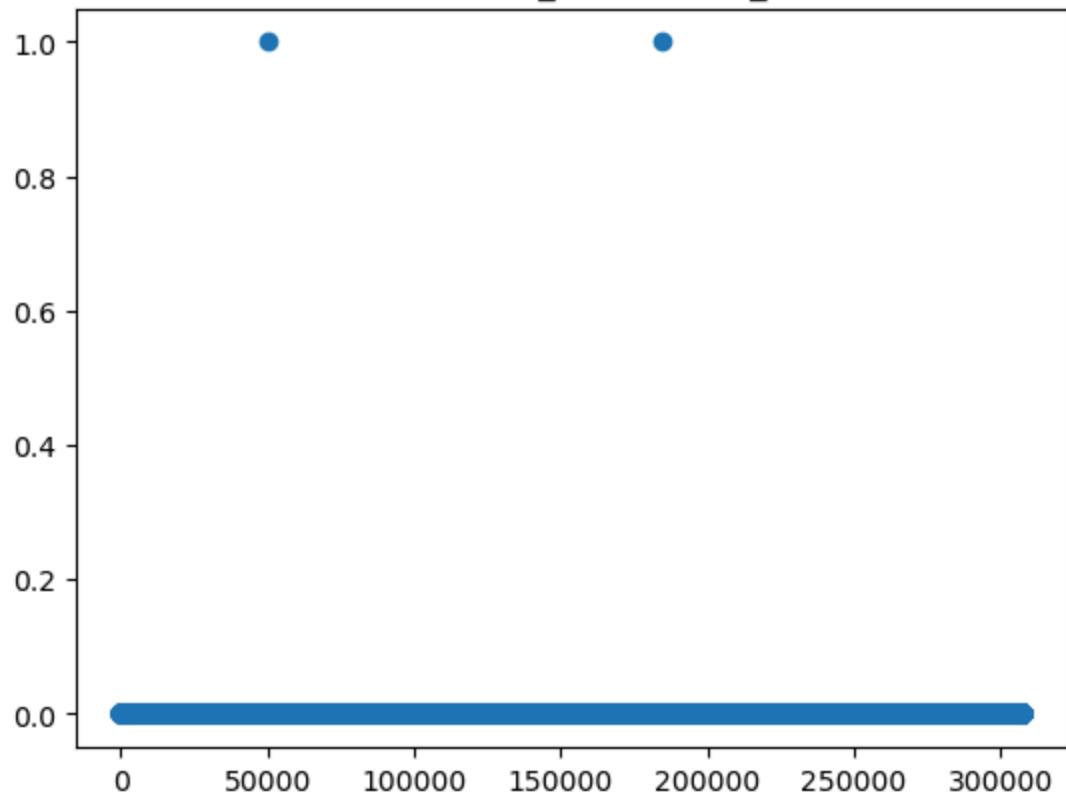
Plot of FLAG_DOCUMENT_11



Out[82]: <matplotlib.collections.PathCollection at 0x22fdbcbc210>

Out[82]: Text(0.5, 1.0, 'Plot of FLAG_DOCUMENT_12')

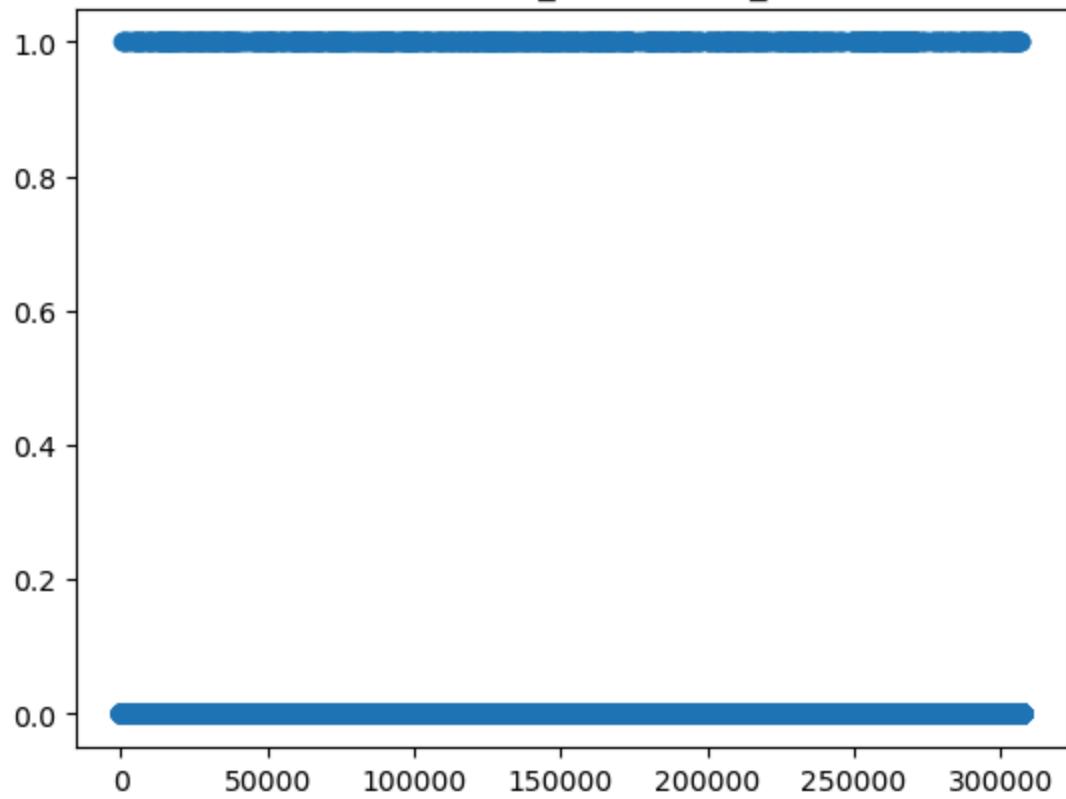
Plot of FLAG_DOCUMENT_12



Out[82]: <matplotlib.collections.PathCollection at 0x22fdbba48d0>

Out[82]: Text(0.5, 1.0, 'Plot of FLAG_DOCUMENT_13')

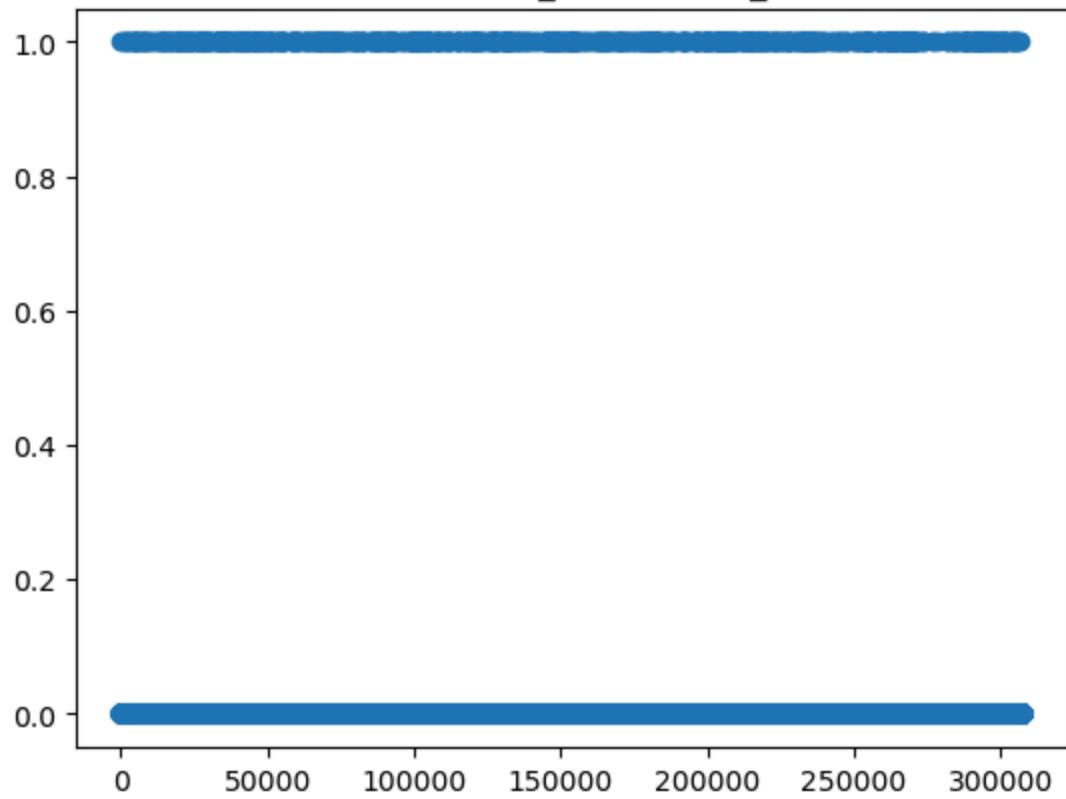
Plot of FLAG_DOCUMENT_13



Out[82]: <matplotlib.collections.PathCollection at 0x22fdb5a110>

Out[82]: Text(0.5, 1.0, 'Plot of FLAG_DOCUMENT_14')

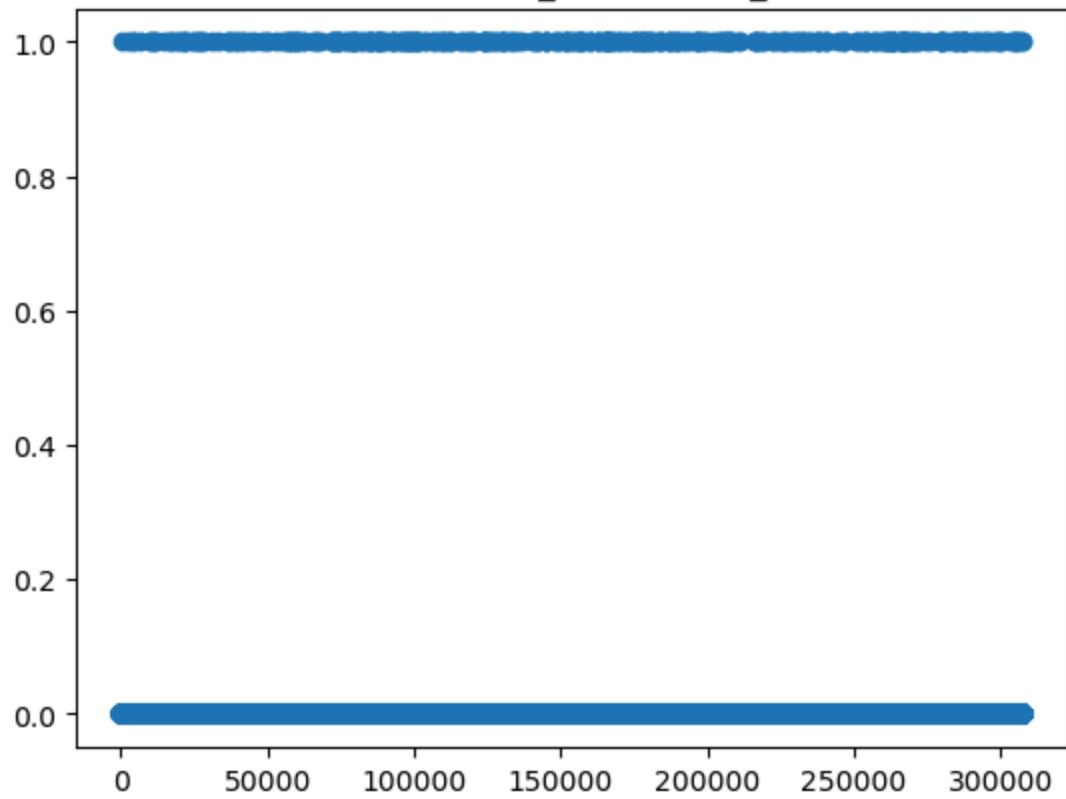
Plot of FLAG_DOCUMENT_14



Out[82]: <matplotlib.collections.PathCollection at 0x22fdbdab510>

Out[82]: Text(0.5, 1.0, 'Plot of FLAG_DOCUMENT_15')

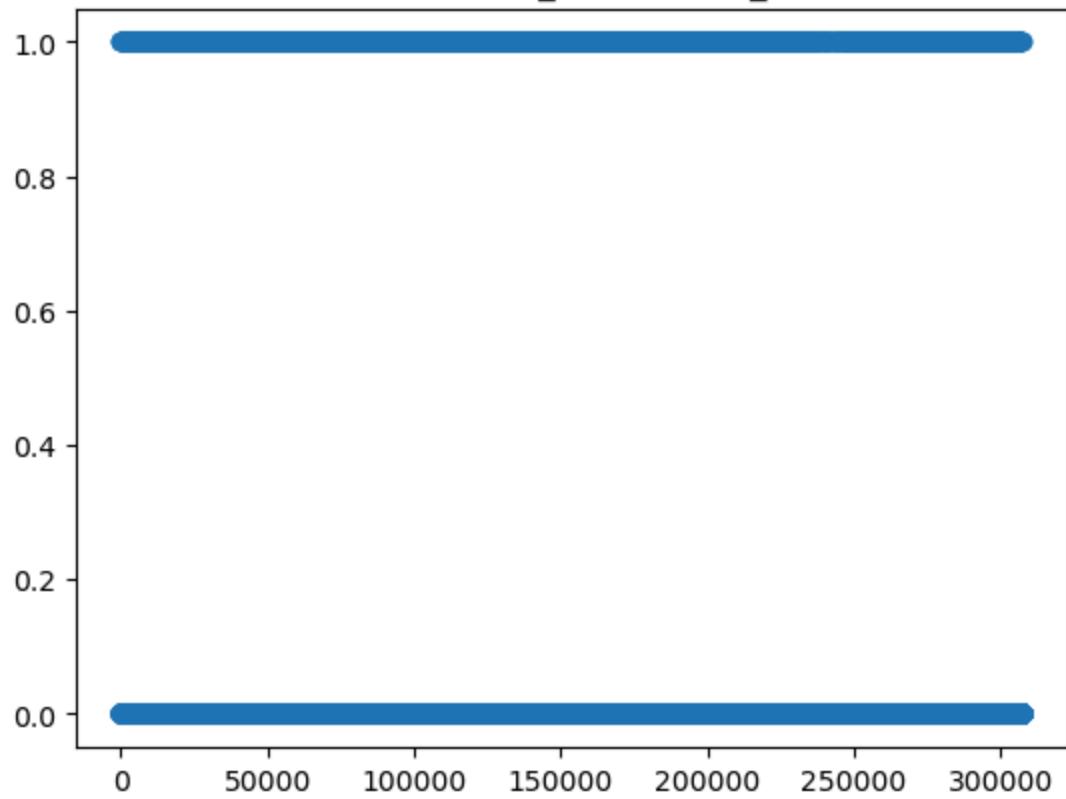
Plot of FLAG_DOCUMENT_15



Out[82]: <matplotlib.collections.PathCollection at 0x22fdbda80d0>

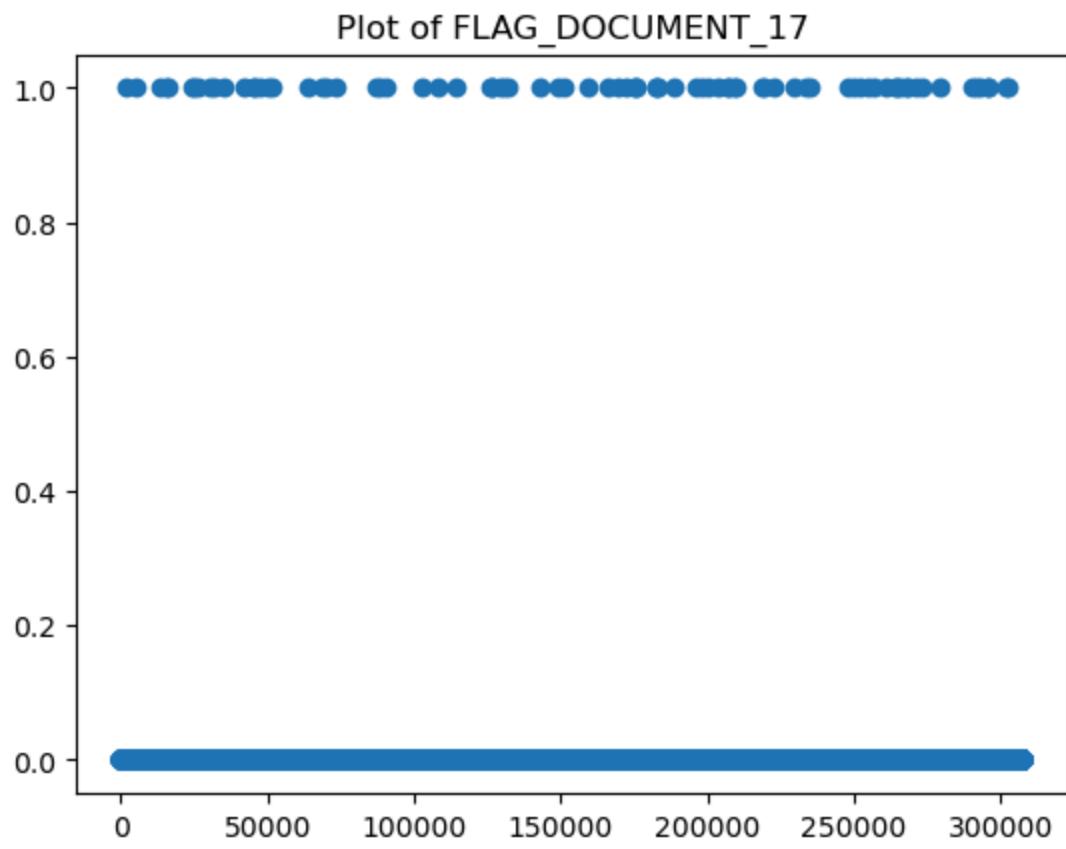
Out[82]: Text(0.5, 1.0, 'Plot of FLAG_DOCUMENT_16')

Plot of FLAG_DOCUMENT_16



Out[82]: <matplotlib.collections.PathCollection at 0x22fdccc15d0>

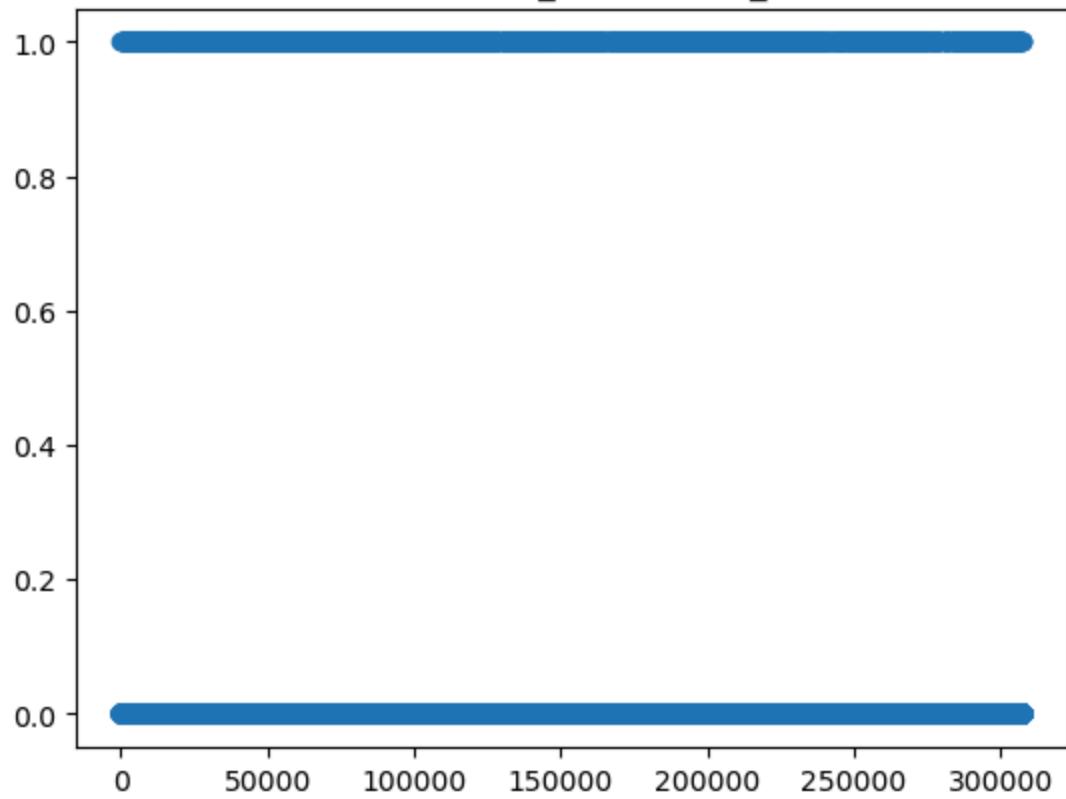
Out[82]: Text(0.5, 1.0, 'Plot of FLAG_DOCUMENT_17')



```
Out[82]: <matplotlib.collections.PathCollection at 0x22fdcccd1f50>
```

```
Out[82]: Text(0.5, 1.0, 'Plot of FLAG_DOCUMENT_18')
```

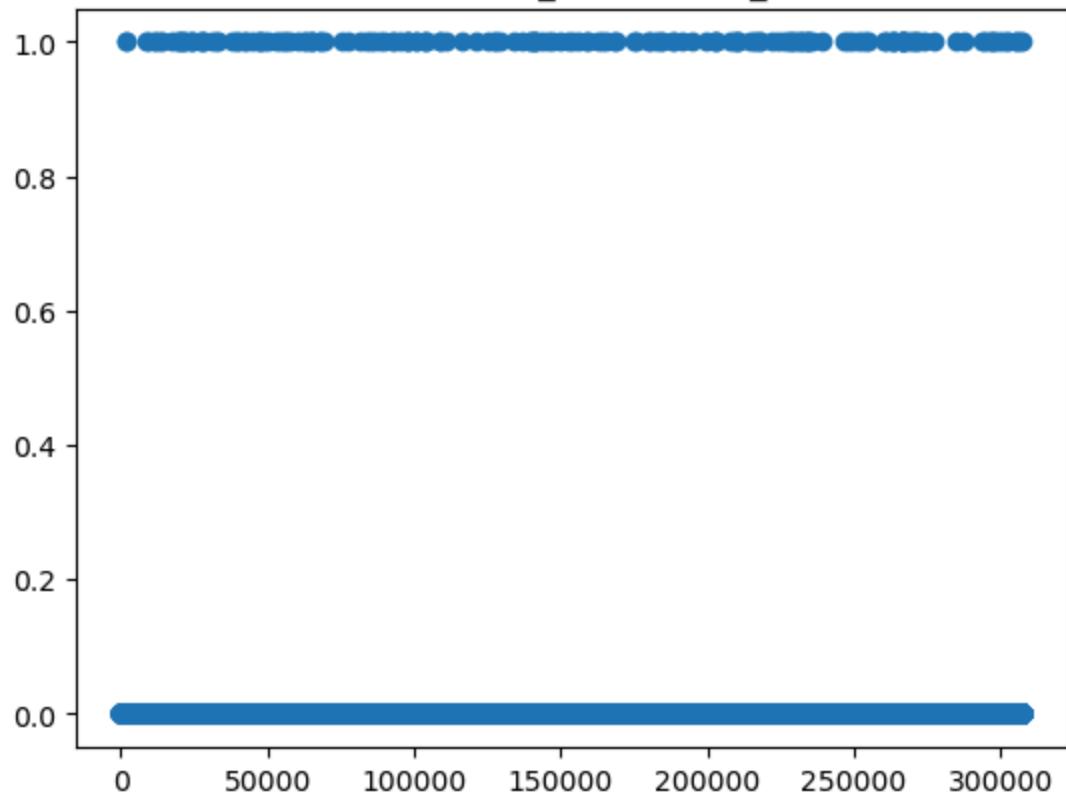
Plot of FLAG_DOCUMENT_18



Out[82]: <matplotlib.collections.PathCollection at 0x22fdbda7390>

Out[82]: Text(0.5, 1.0, 'Plot of FLAG_DOCUMENT_19')

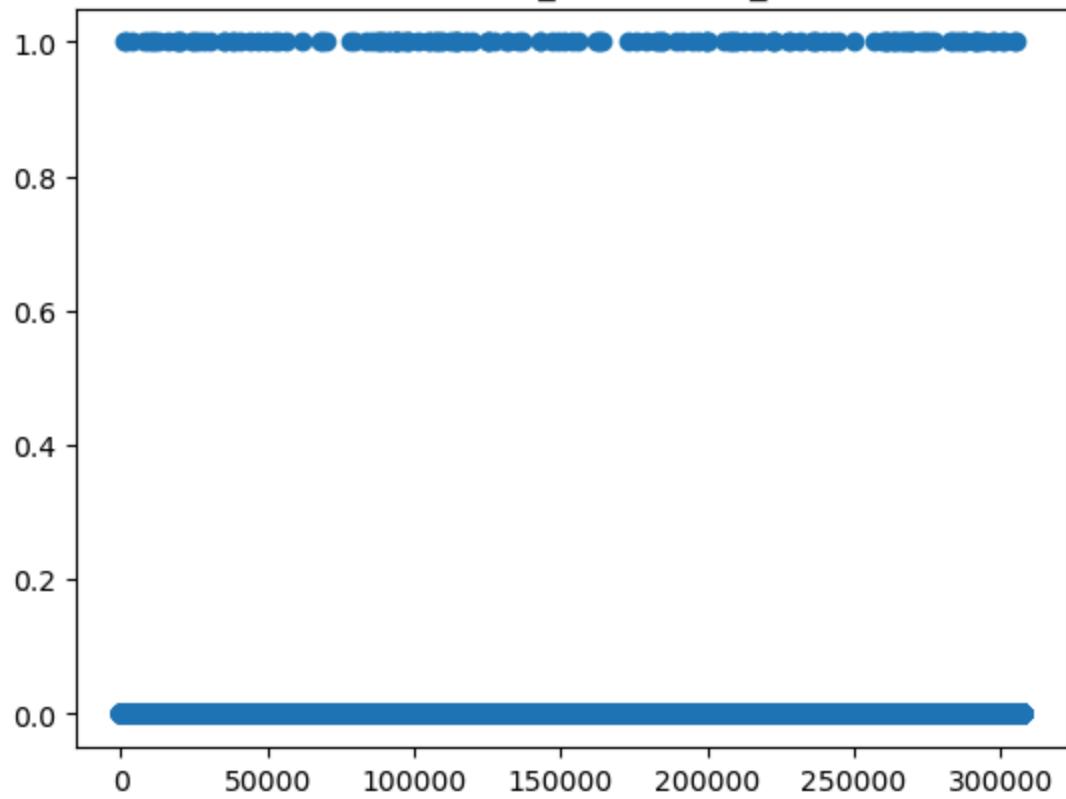
Plot of FLAG_DOCUMENT_19



Out[82]: <matplotlib.collections.PathCollection at 0x22fdbd0c210>

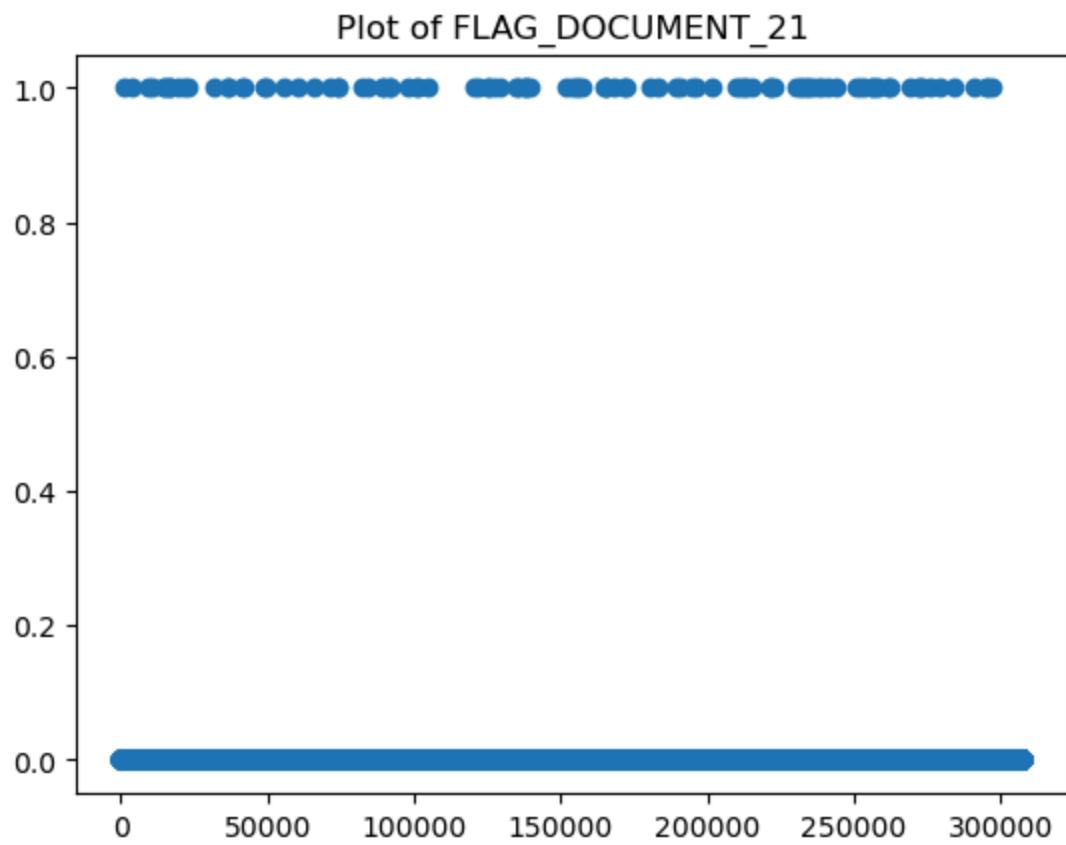
Out[82]: Text(0.5, 1.0, 'Plot of FLAG_DOCUMENT_20')

Plot of FLAG_DOCUMENT_20



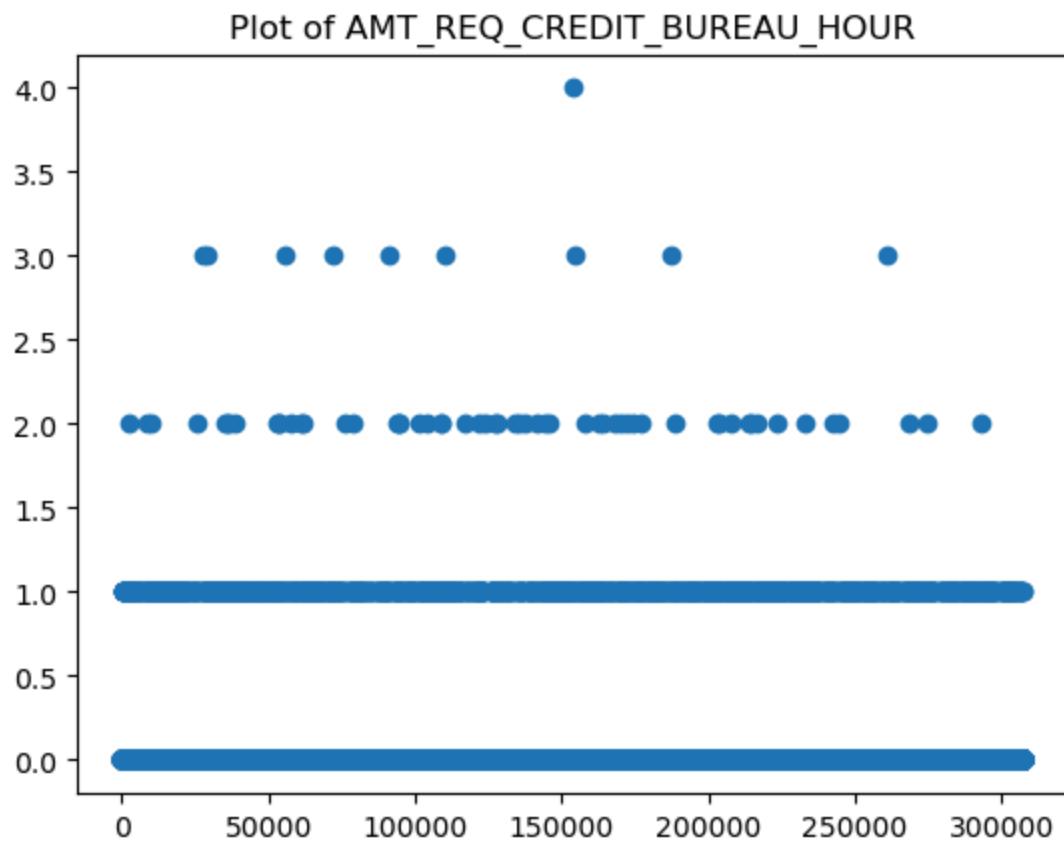
Out[82]: <matplotlib.collections.PathCollection at 0x22fdbc84210>

Out[82]: Text(0.5, 1.0, 'Plot of FLAG_DOCUMENT_21')



```
Out[82]: <matplotlib.collections.PathCollection at 0x22fdae13f10>
```

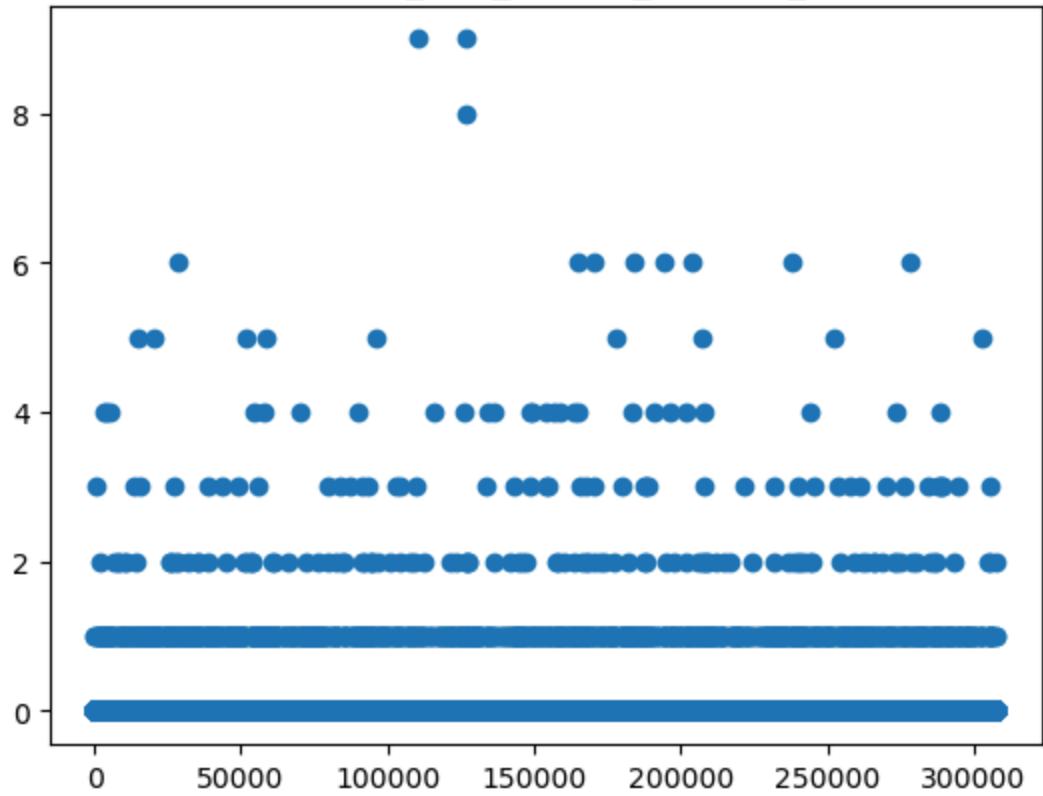
```
Out[82]: Text(0.5, 1.0, 'Plot of AMT_REQ_CREDIT_BUREAU_HOUR')
```



```
Out[82]: <matplotlib.collections.PathCollection at 0x22fdbc4c510>
```

```
Out[82]: Text(0.5, 1.0, 'Plot of AMT_REQ_CREDIT_BUREAU_DAY')
```

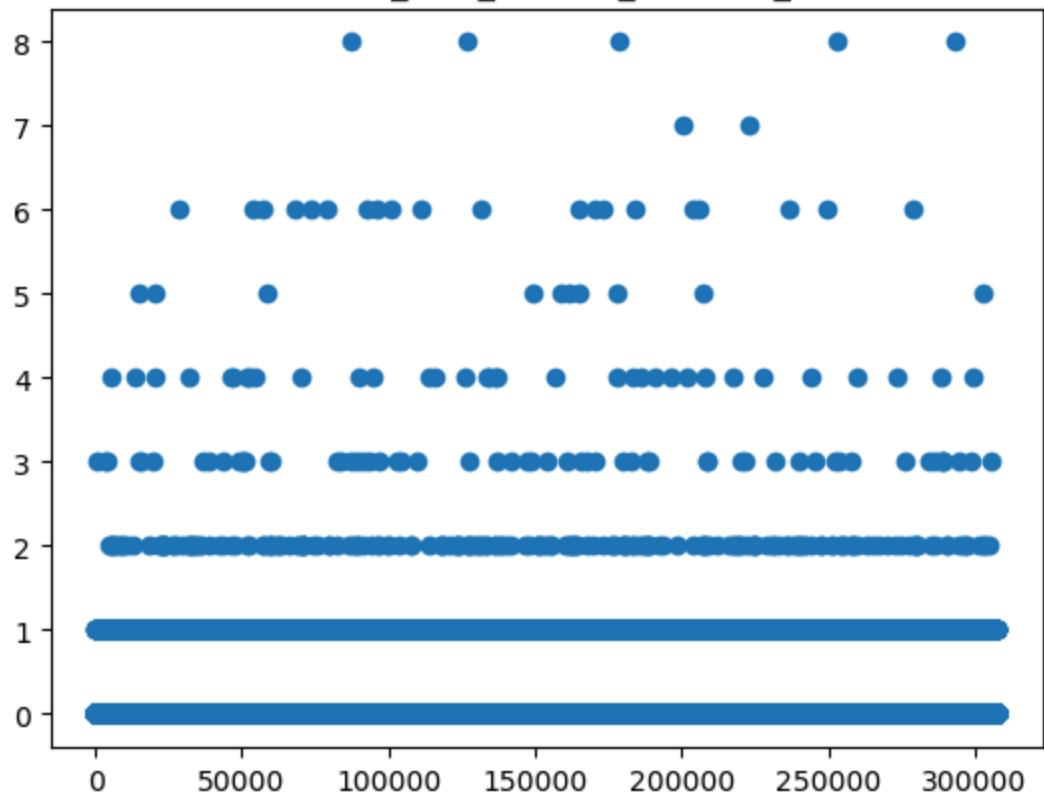
Plot of AMT_REQ_CREDIT_BUREAU_DAY



Out[82]: <matplotlib.collections.PathCollection at 0x22fdbb2abd0>

Out[82]: Text(0.5, 1.0, 'Plot of AMT_REQ_CREDIT_BUREAU_WEEK')

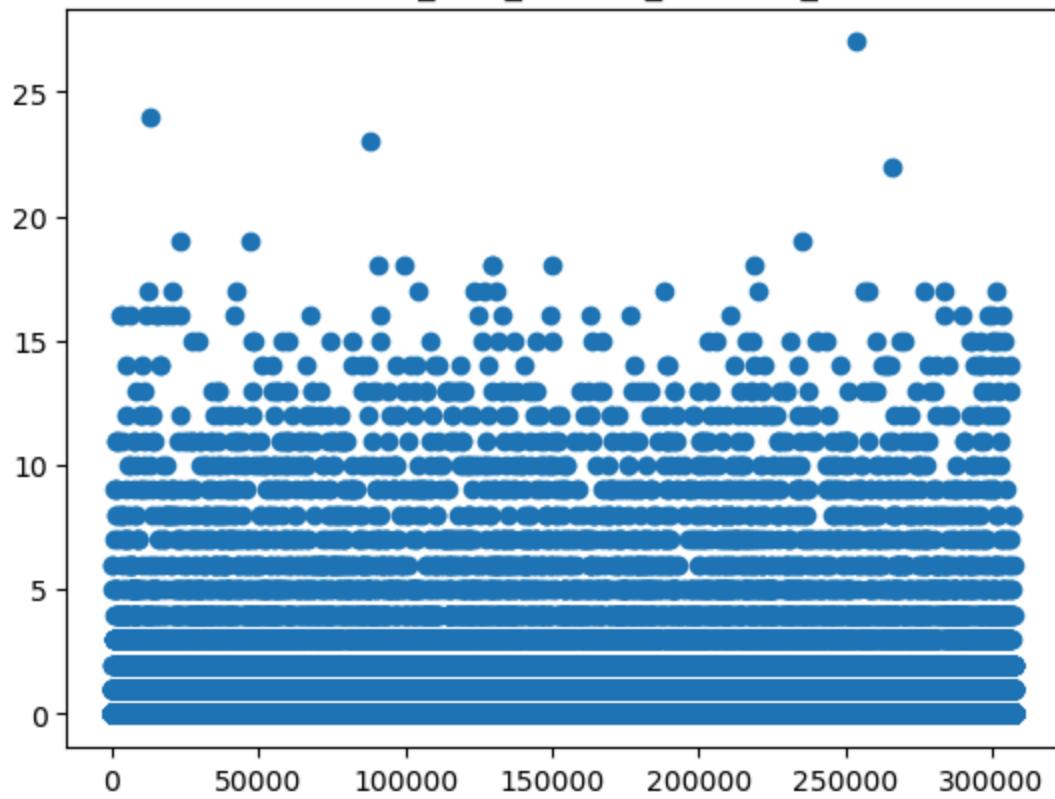
Plot of AMT_REQ_CREDIT_BUREAU_WEEK



Out[82]: <matplotlib.collections.PathCollection at 0x22fdbb0ed10>

Out[82]: Text(0.5, 1.0, 'Plot of AMT_REQ_CREDIT_BUREAU_MON')

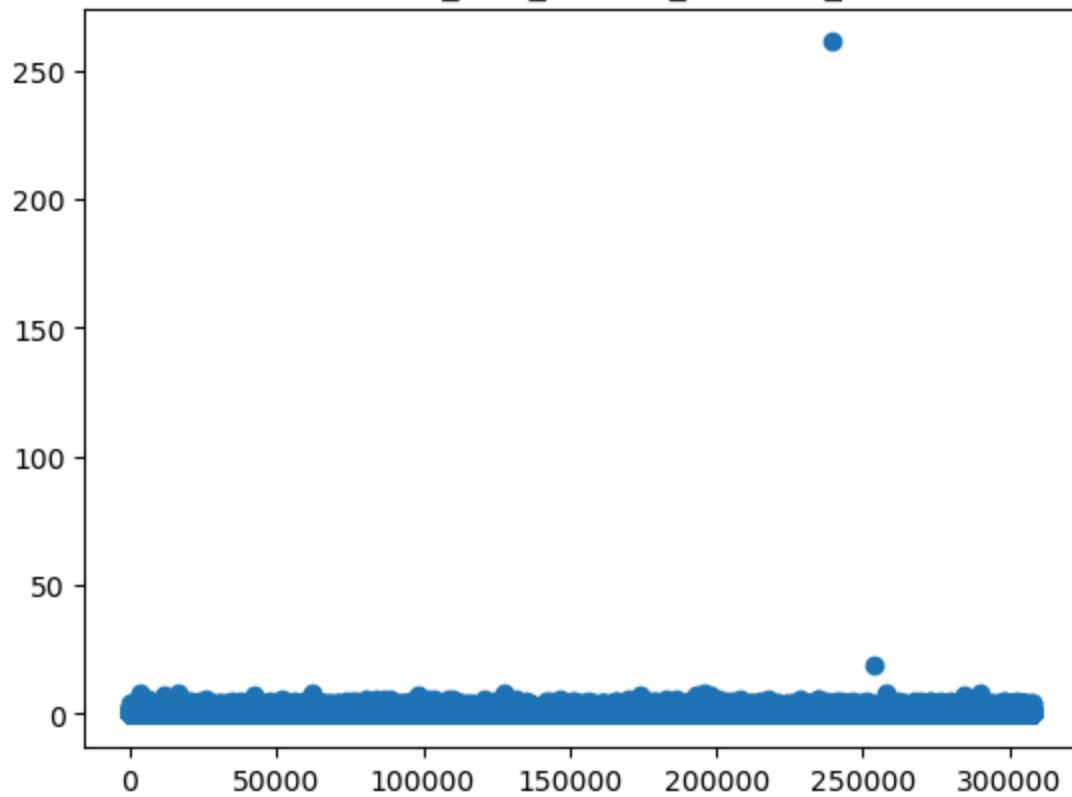
Plot of AMT_REQ_CREDIT_BUREAU_MON



Out[82]: <matplotlib.collections.PathCollection at 0x22fdc8ffb50>

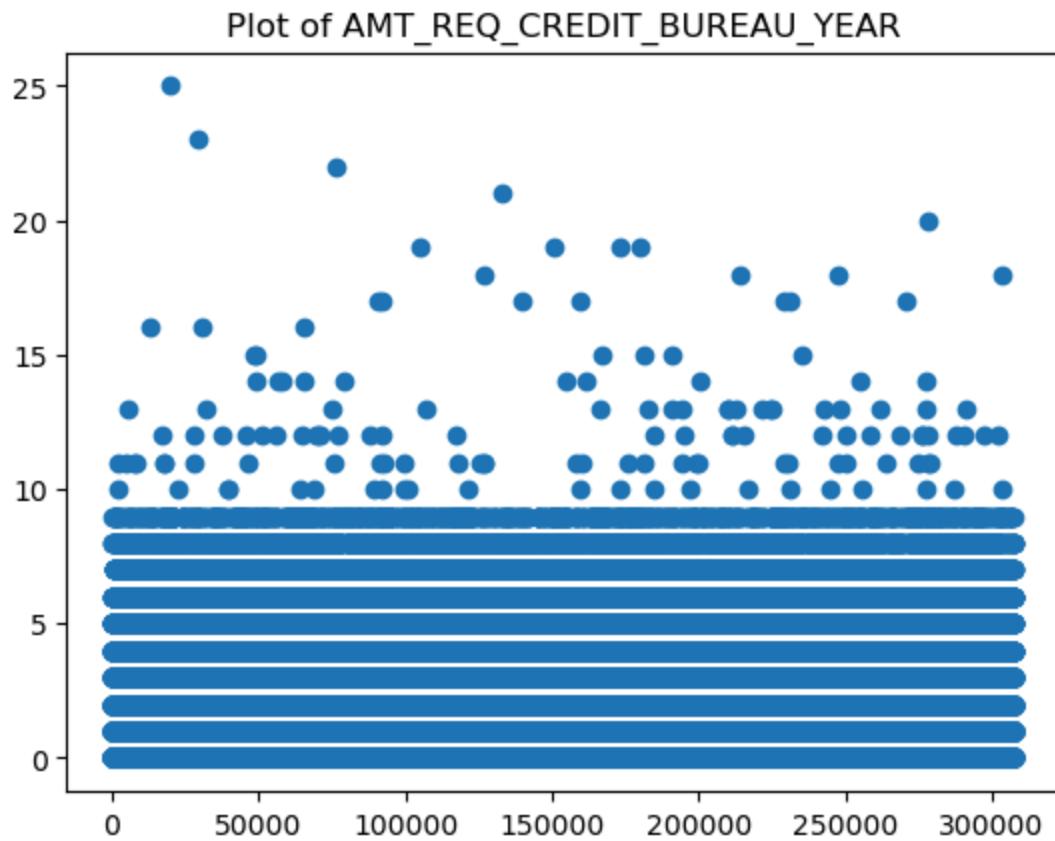
Out[82]: Text(0.5, 1.0, 'Plot of AMT_REQ_CREDIT_BUREAU_QRT')

Plot of AMT_REQ_CREDIT_BUREAU_QRT



Out[82]: <matplotlib.collections.PathCollection at 0x22fcec26850>

Out[82]: Text(0.5, 1.0, 'Plot of AMT_REQ_CREDIT_BUREAU_YEAR')



```
In [86]: for column in train_categorical:
    title = "Plot of "+column
    print(title)
    plt.hist(train_0[column], alpha=0.5, label='0', bins=8)
    plt.hist(train_1[column], alpha=0.5, label='1', bins=8)
    plt.show()

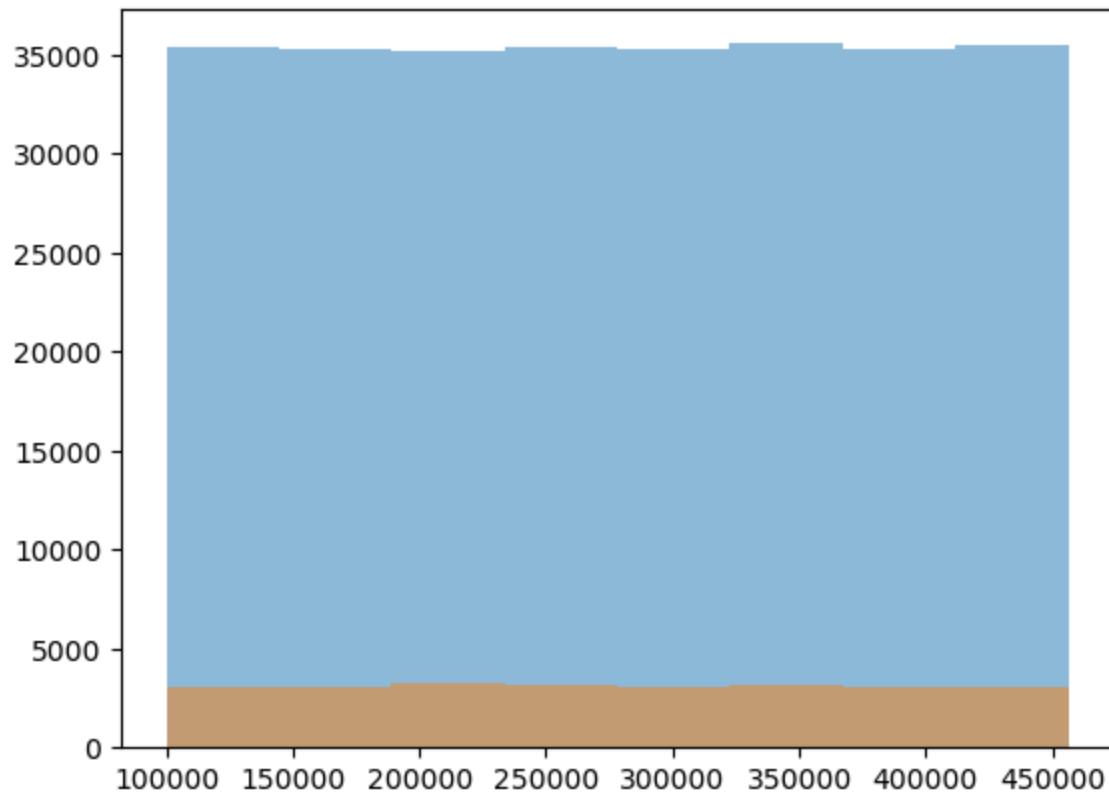
    sns.distplot(train_0[column].dropna(), label='0')
    sns.distplot(train_1[column].dropna(), label='1')

    plt.show()
    #box_plot(train_0, train_1, column)
    print("-----")
```

Plot of SK_ID_CURR

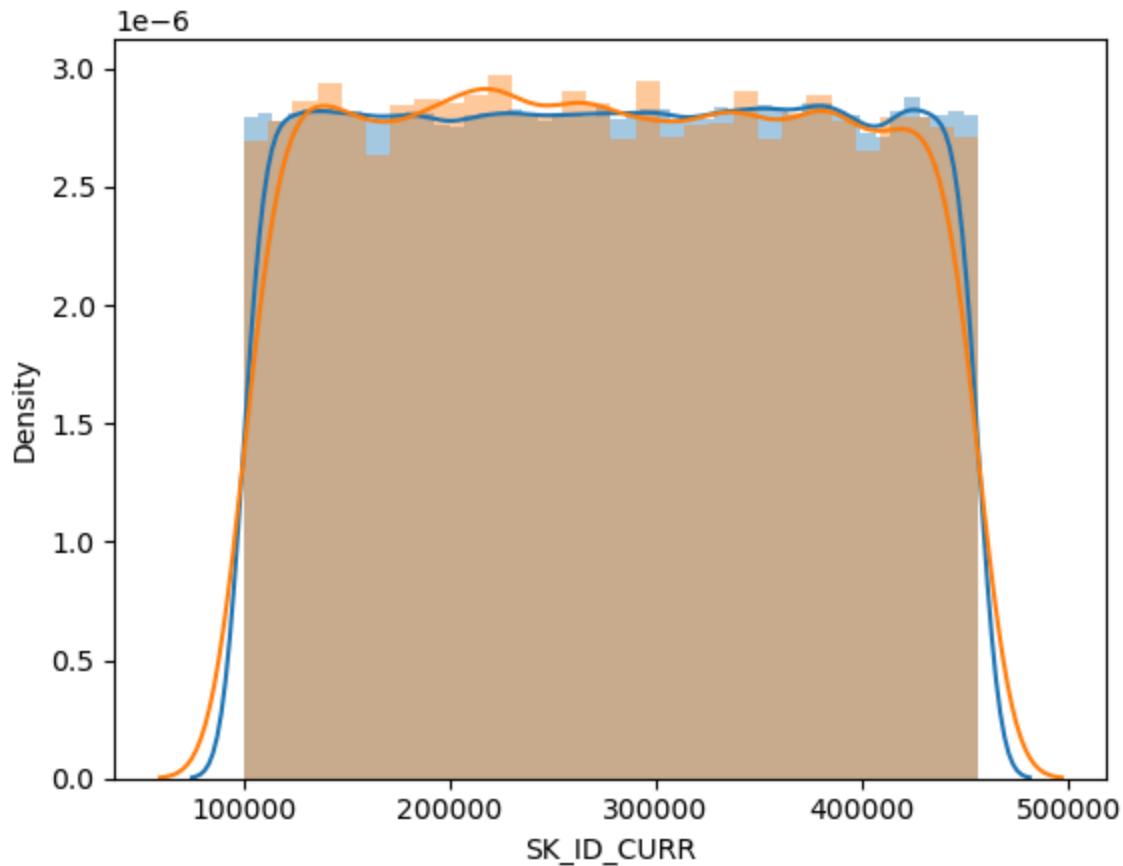
```
Out[86]: (array([35359., 35281., 35179., 35334., 35273., 35548., 35251., 35461.]),
array([100003., 144534.5, 189066., 233597.5, 278129., 322660.5,
       367192., 411723.5, 456255.]),
<BarContainer object of 8 artists>)
```

```
Out[86]: (array([3102., 3062., 3221., 3152., 3055., 3111., 3065., 3057.]),
array([100002., 144533.5, 189065., 233596.5, 278128., 322659.5,
       367191., 411722.5, 456254.]),
<BarContainer object of 8 artists>)
```



```
Out[86]: <Axes: xlabel='SK_ID_CURR', ylabel='Density'>
```

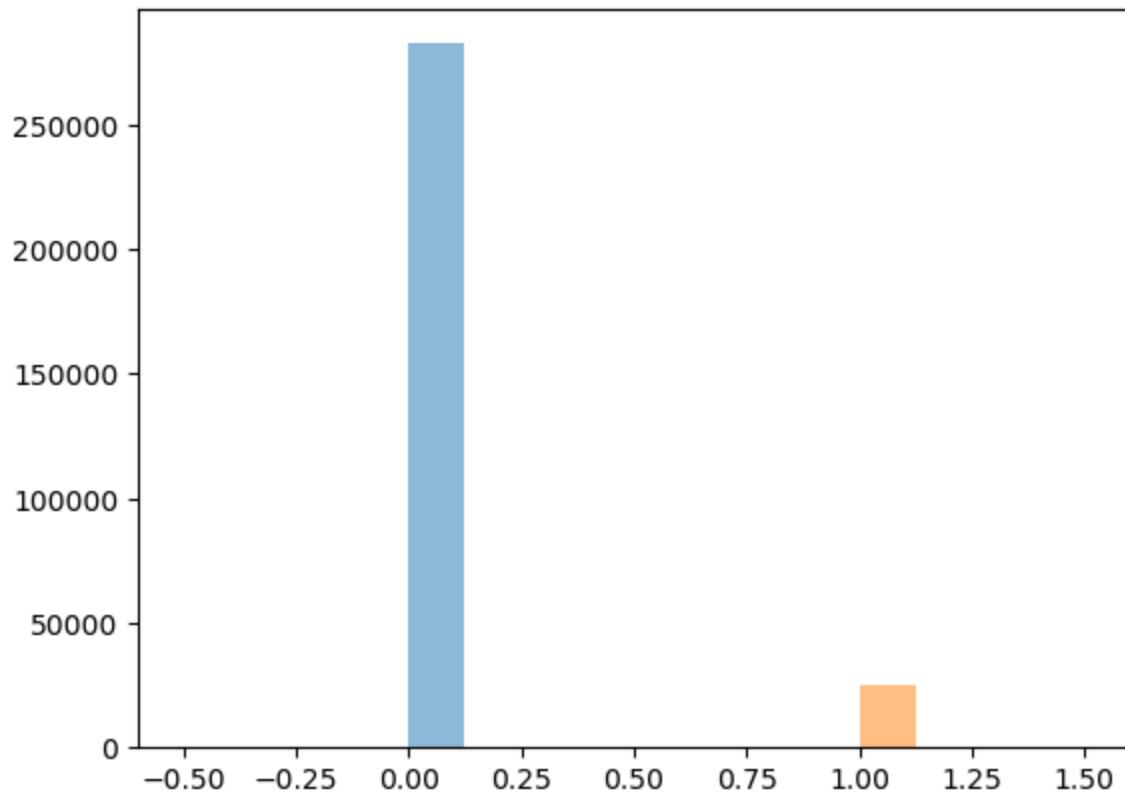
```
Out[86]: <Axes: xlabel='SK_ID_CURR', ylabel='Density'>
```



Plot of TARGET

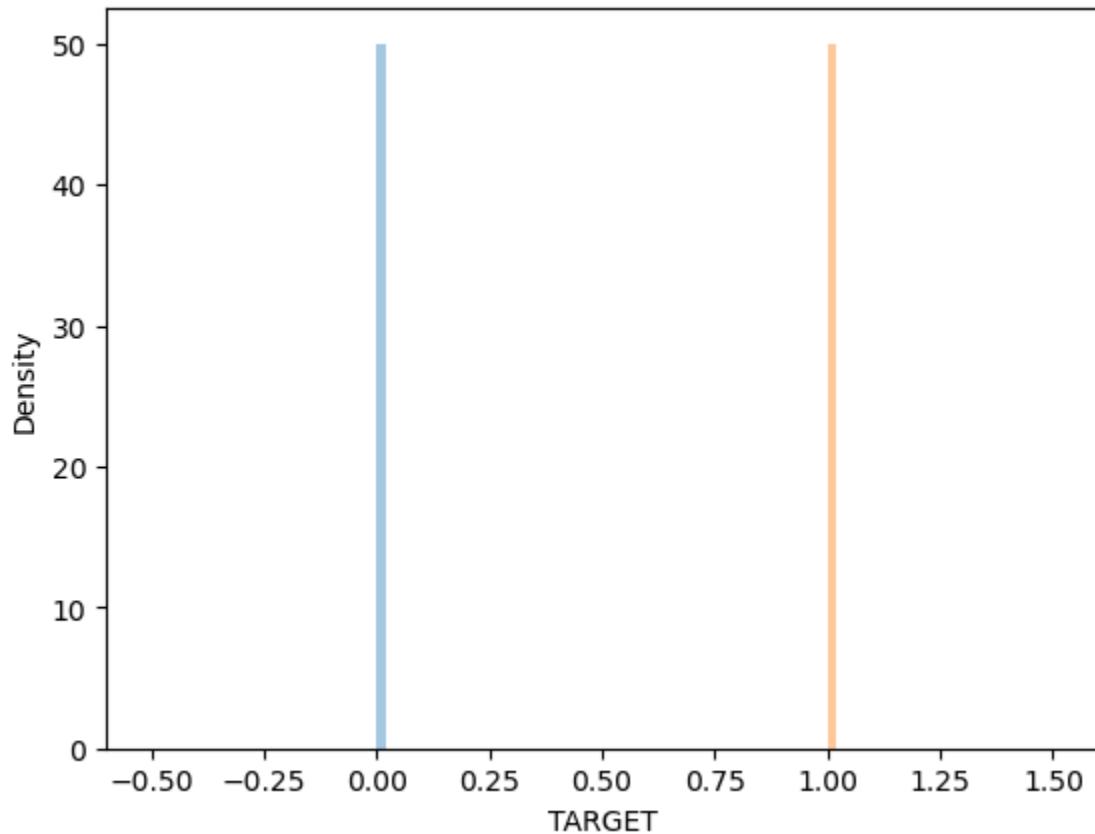
```
Out[86]: (array([ 0., 0., 0., 0., 282686., 0., 0.,
                   0.]),
           array([-0.5 , -0.375, -0.25 , -0.125, 0. , 0.125, 0.25 , 0.375,
                  0.5 ]),
           <BarContainer object of 8 artists>)

Out[86]: (array([ 0., 0., 0., 0., 24825., 0., 0., 0.]),
           array([0.5 , 0.625, 0.75 , 0.875, 1. , 1.125, 1.25 , 1.375, 1.5 ]),
           <BarContainer object of 8 artists>)
```



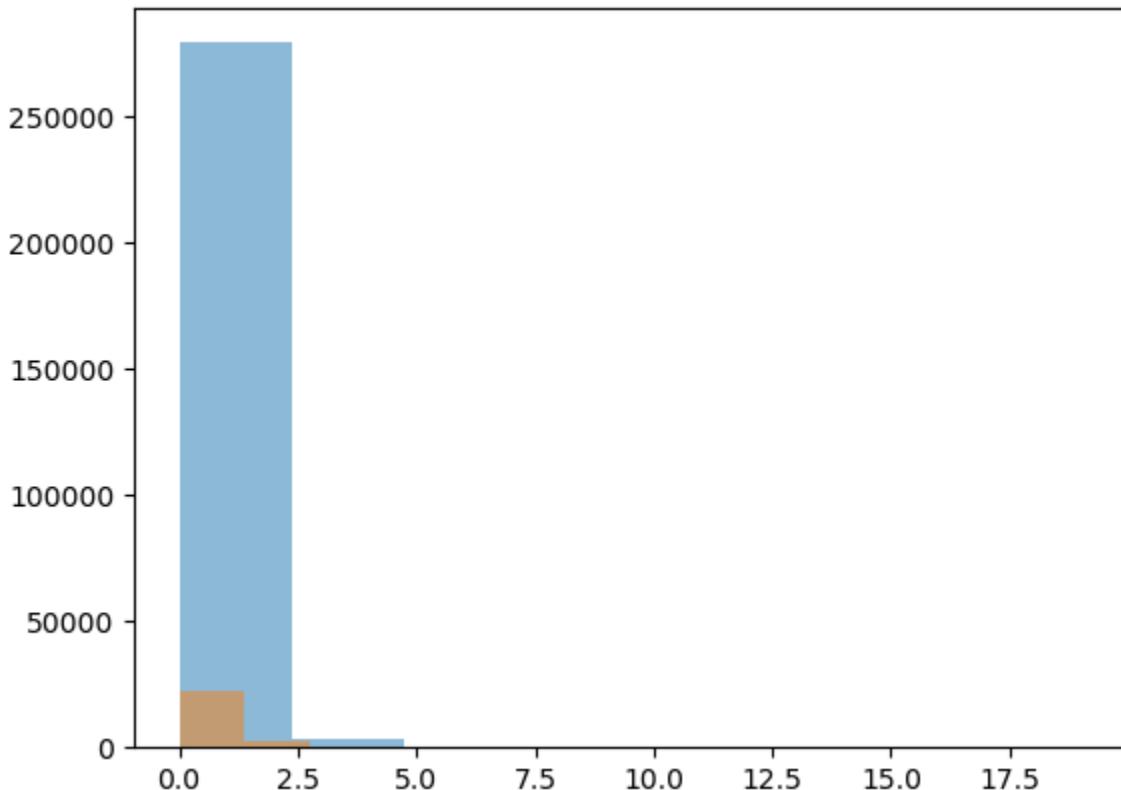
```
Out[86]: <Axes: xlabel='TARGET', ylabel='Density'>
```

```
Out[86]: <Axes: xlabel='TARGET', ylabel='Density'>
```



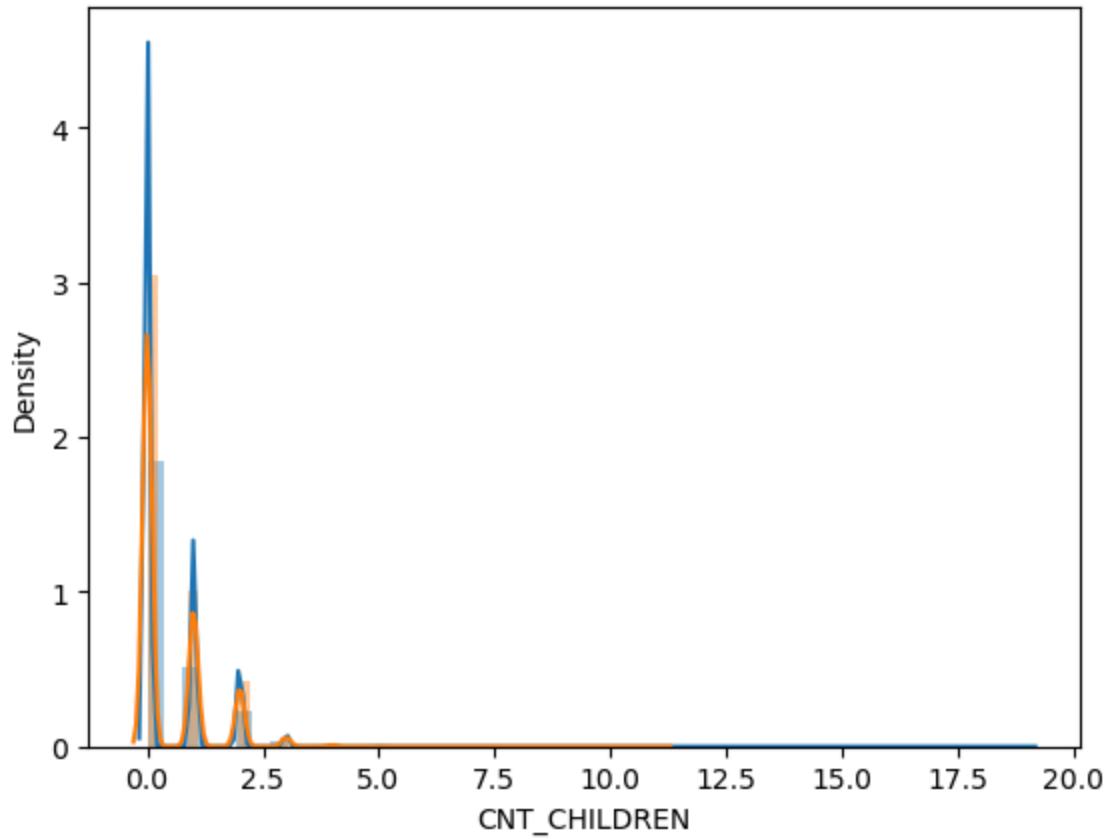
Plot of CNT_CHILDREN

```
Out[86]: (array([2.78843e+05, 3.73300e+03, 9.90000e+01, 2.00000e+00, 2.00000e+00,
      5.00000e+00, 0.00000e+00, 2.00000e+00]),  
 array([ 0. ,  2.375,  4.75 ,  7.125,  9.5 , 11.875, 14.25 , 16.625,  
    19. ]),  
 <BarContainer object of 8 artists>)  
  
Out[86]: (array([2.2063e+04, 2.3330e+03, 4.1300e+02, 7.0000e+00, 6.0000e+00,  
      0.0000e+00, 2.0000e+00, 1.0000e+00]),  
 array([ 0. ,  1.375,  2.75 ,  4.125,  5.5 ,  6.875,  8.25 ,  9.625,  
    11. ]),  
 <BarContainer object of 8 artists>)
```



```
Out[86]: <Axes: xlabel='CNT_CHILDREN', ylabel='Density'>
```

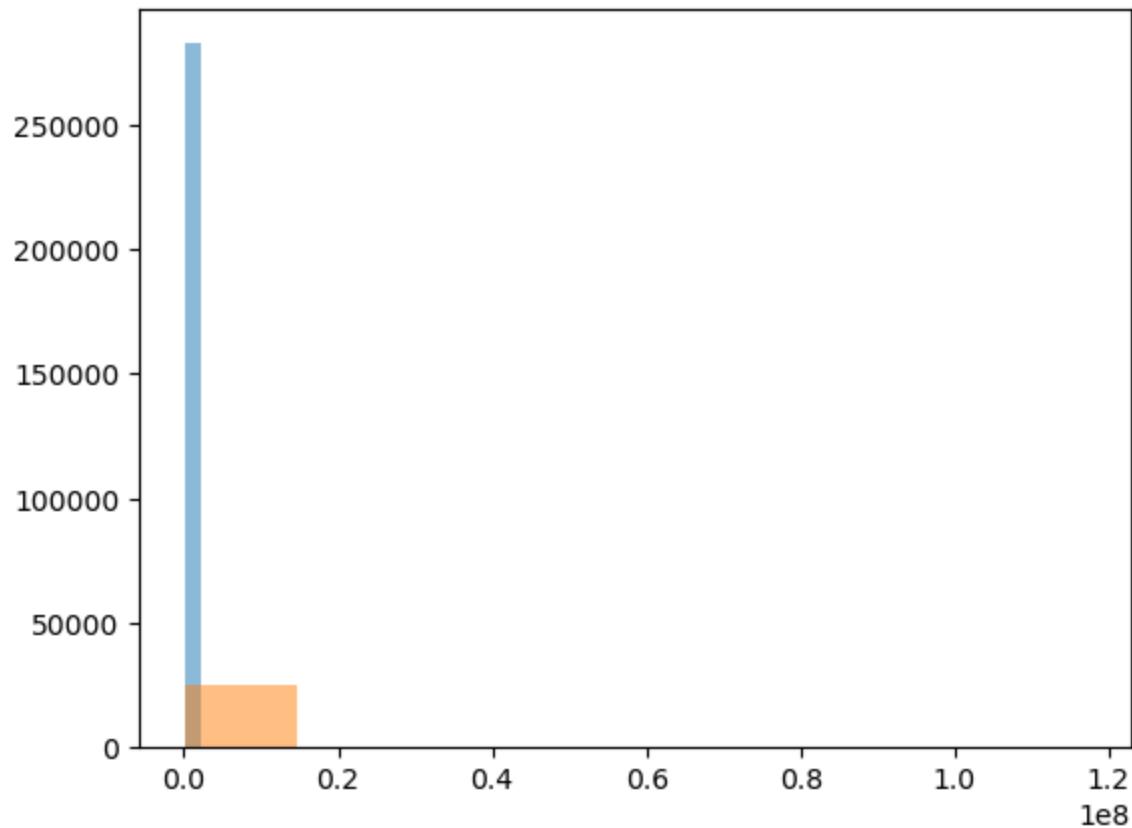
```
Out[86]: <Axes: xlabel='CNT_CHILDREN', ylabel='Density'>
```



Plot of `AMT_INCOME_TOTAL`

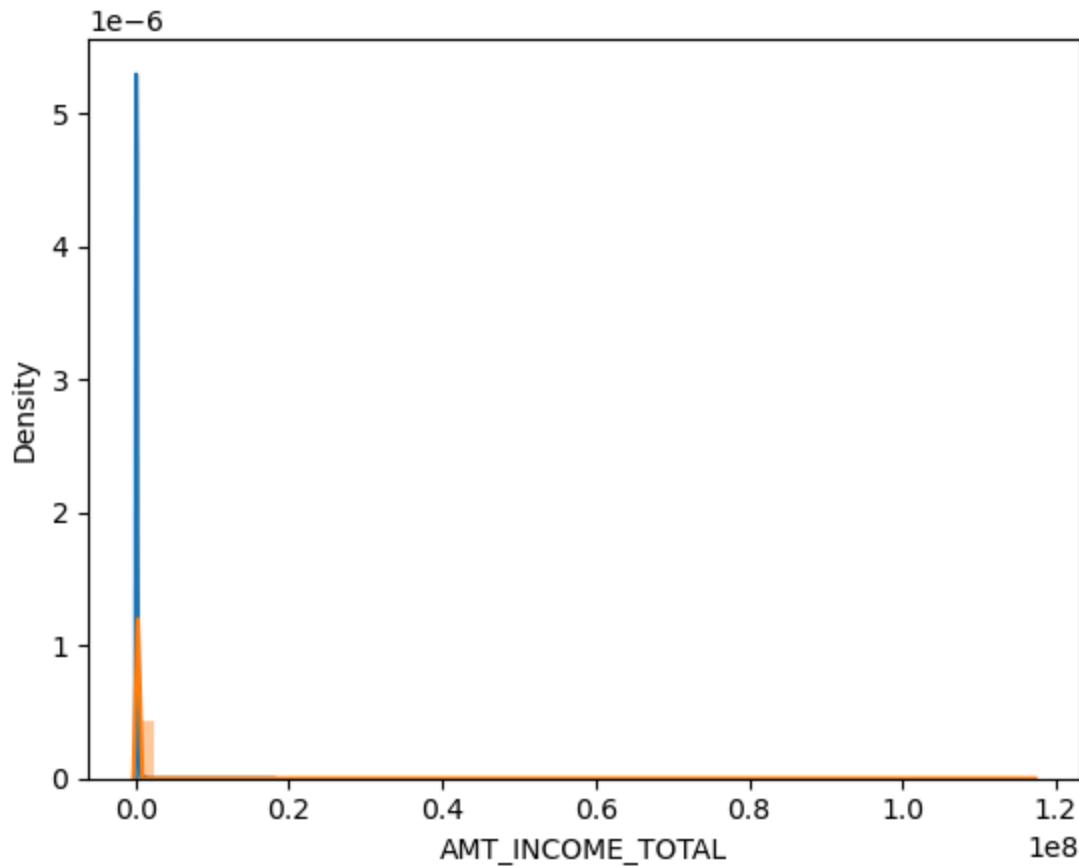
```
Out[86]: (array([2.82668e+05, 1.40000e+01, 1.00000e+00, 1.00000e+00, 0.00000e+00,
       1.00000e+00, 0.00000e+00, 1.00000e+00]),
       array([ 25650.,  2272455.,  4519260.,  6766065.,  9012870., 11259675.,
              13506480., 15753285., 18000090.]),
       <BarContainer object of 8 artists>)

Out[86]: (array([2.4824e+04, 0.0000e+00, 0.0000e+00, 0.0000e+00, 0.0000e+00,
       0.0000e+00, 0.0000e+00, 1.0000e+00]),
       array([2.56500000e+04, 1.46474438e+07, 2.92692375e+07, 4.38910312e+07,
              5.85128250e+07, 7.31346188e+07, 8.77564125e+07, 1.02378206e+08,
              1.17000000e+08]),
       <BarContainer object of 8 artists>)
```



```
Out[86]: <Axes: xlabel='AMT_INCOME_TOTAL', ylabel='Density'>
```

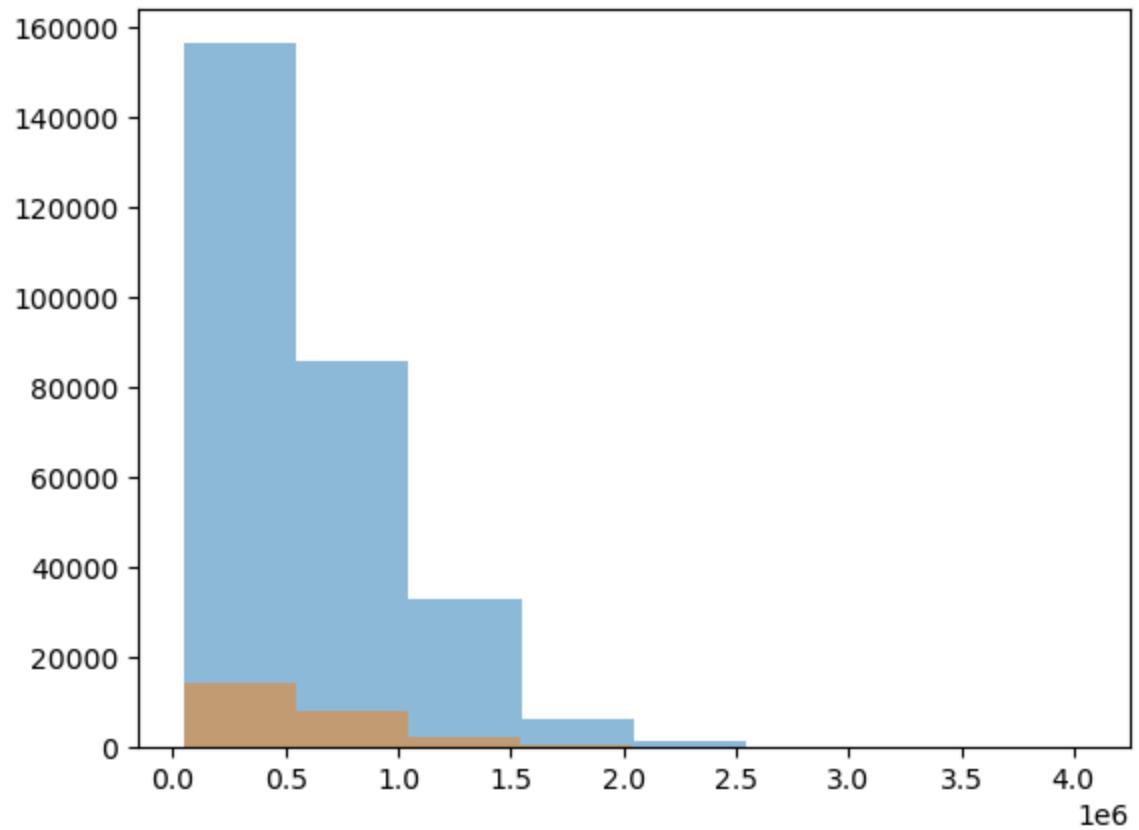
```
Out[86]: <Axes: xlabel='AMT_INCOME_TOTAL', ylabel='Density'>
```



Plot of `AMT_CREDIT`

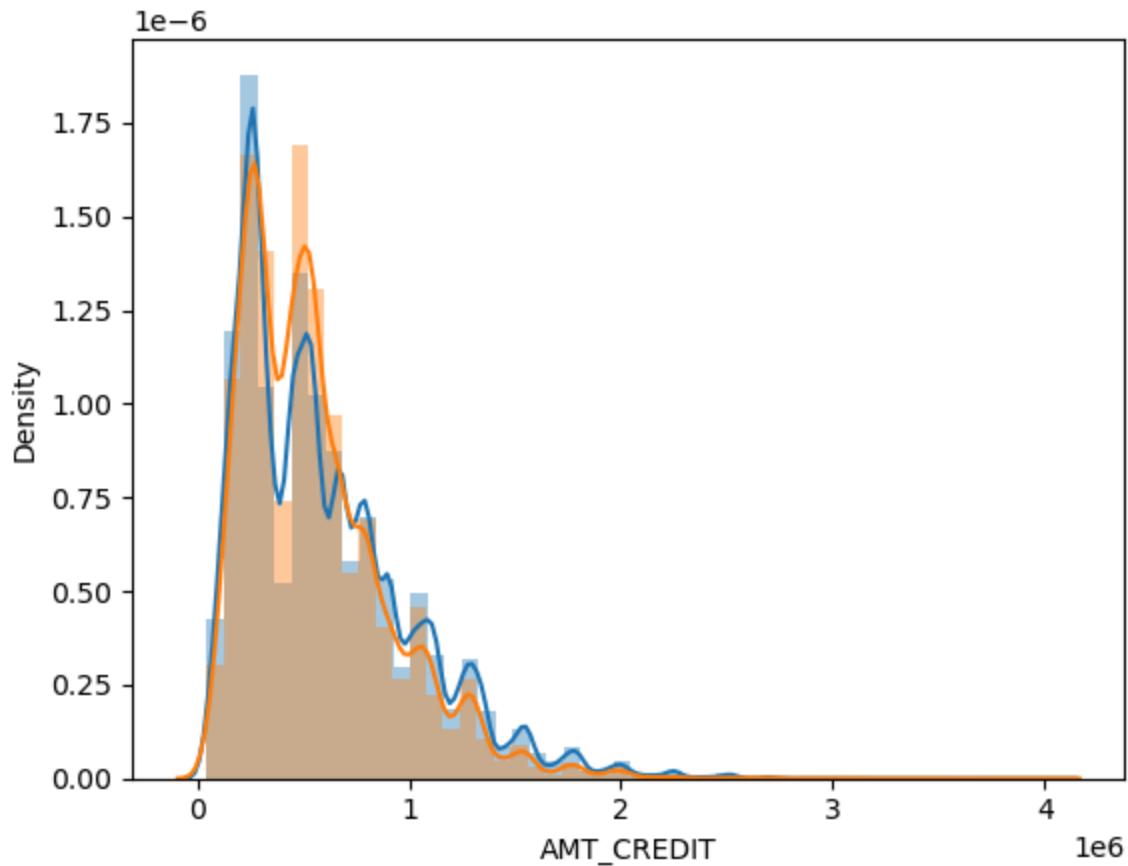
```
Out[86]: (array([1.56304e+05, 8.57420e+04, 3.29170e+04, 6.31800e+03, 1.28300e+03,
   9.20000e+01, 1.70000e+01, 1.30000e+01]),
 array([ 45000.,  545625., 1046250., 1546875., 2047500., 2548125.,
  3048750., 3549375., 4050000.]),
 <BarContainer object of 8 artists>)

Out[86]: (array([1.4077e+04, 8.2130e+03, 2.1190e+03, 3.7400e+02, 3.2000e+01,
   9.0000e+00, 0.0000e+00, 1.0000e+00]),
 array([ 45000.,  542835., 1040670., 1538505., 2036340., 2534175.,
  3032010., 3529845., 4027680.]),
 <BarContainer object of 8 artists>)
```



```
Out[86]: <Axes: xlabel='AMT_CREDIT', ylabel='Density'>
```

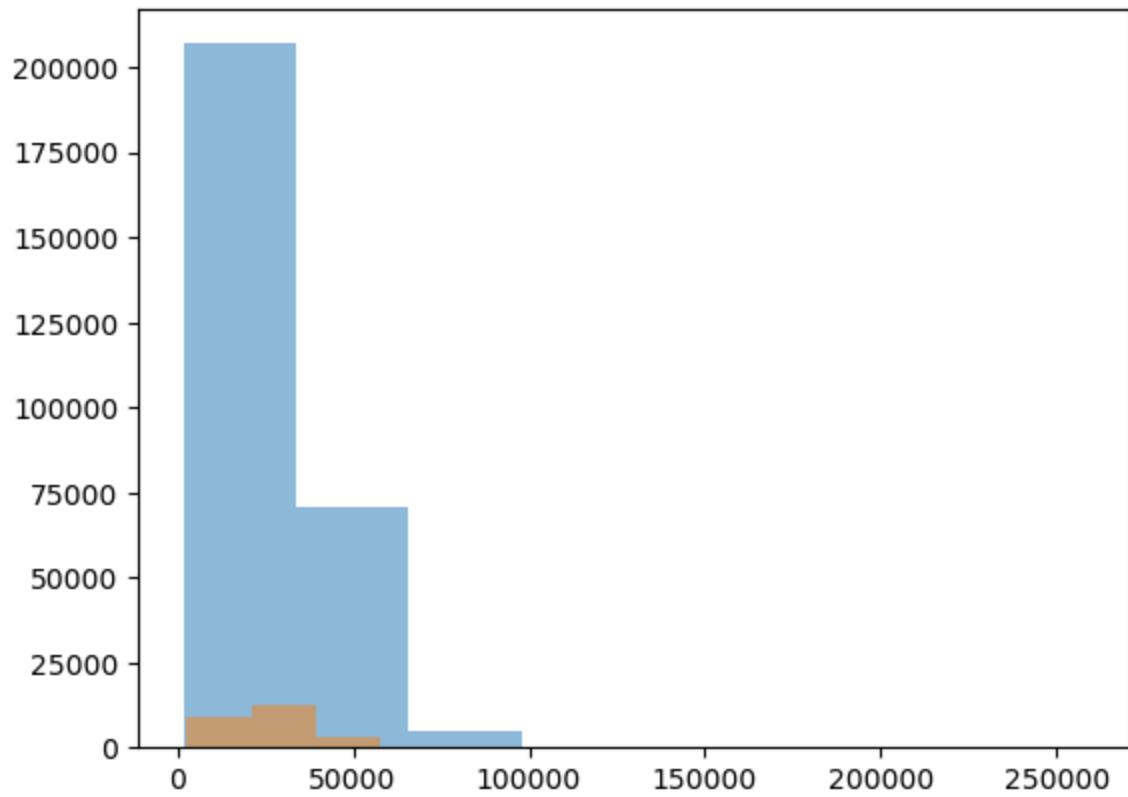
```
Out[86]: <Axes: xlabel='AMT_CREDIT', ylabel='Density'>
```



Plot of `AMT_ANNUITY`

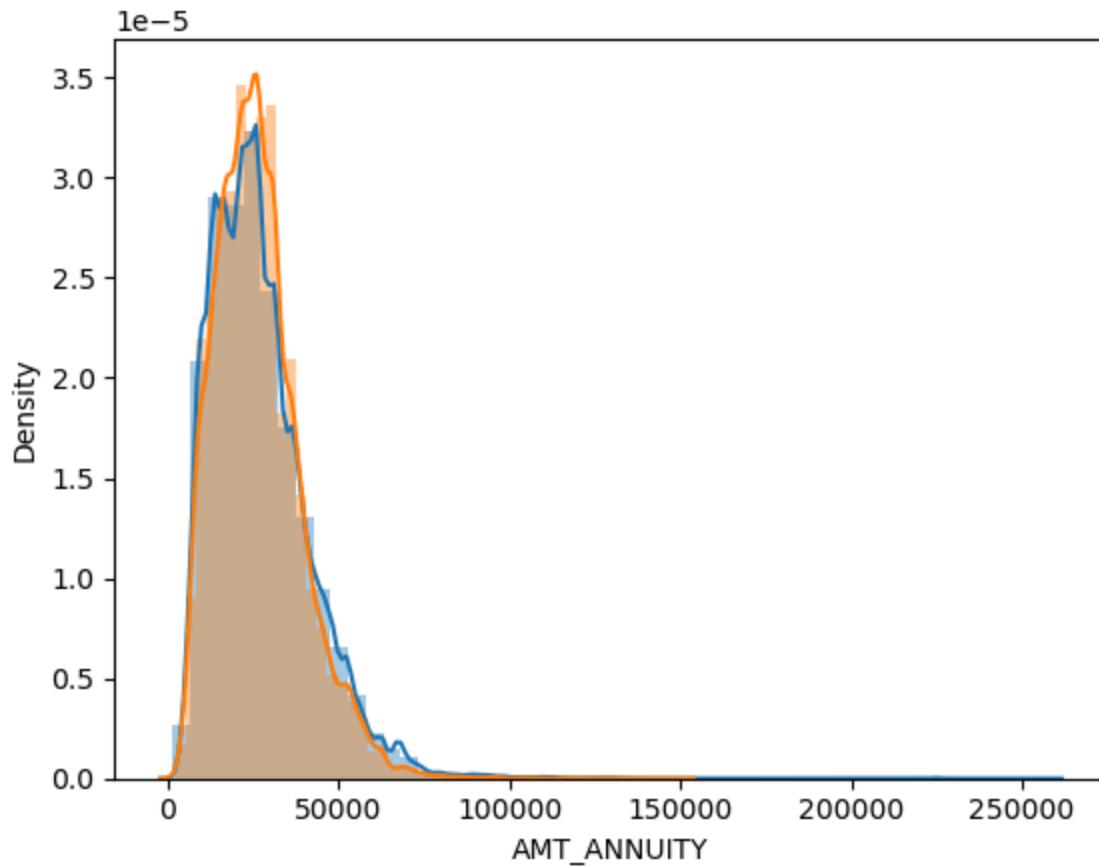
```
Out[86]: (array([2.06974e+05, 7.05530e+04, 4.57600e+03, 4.50000e+02, 6.30000e+01,
   2.10000e+01, 3.50000e+01, 2.00000e+00]),
 array([ 1615.5 , 33666.75, 65718. , 97769.25, 129820.5 , 161871.75,
  193923. , 225974.25, 258025.5 ]),
 <BarContainer object of 8 artists>)

Out[86]: (array([8.8130e+03, 1.2597e+04, 2.9450e+03, 4.0900e+02, 4.3000e+01,
   1.3000e+01, 4.0000e+00, 1.0000e+00]),
 array([ 2722.5 , 21033.5625, 39344.625 , 57655.6875, 75966.75 ,
  94277.8125, 112588.875 , 130899.9375, 149211.     ]),
 <BarContainer object of 8 artists>)
```



```
Out[86]: <Axes: xlabel='AMT_ANNUITY', ylabel='Density'>
```

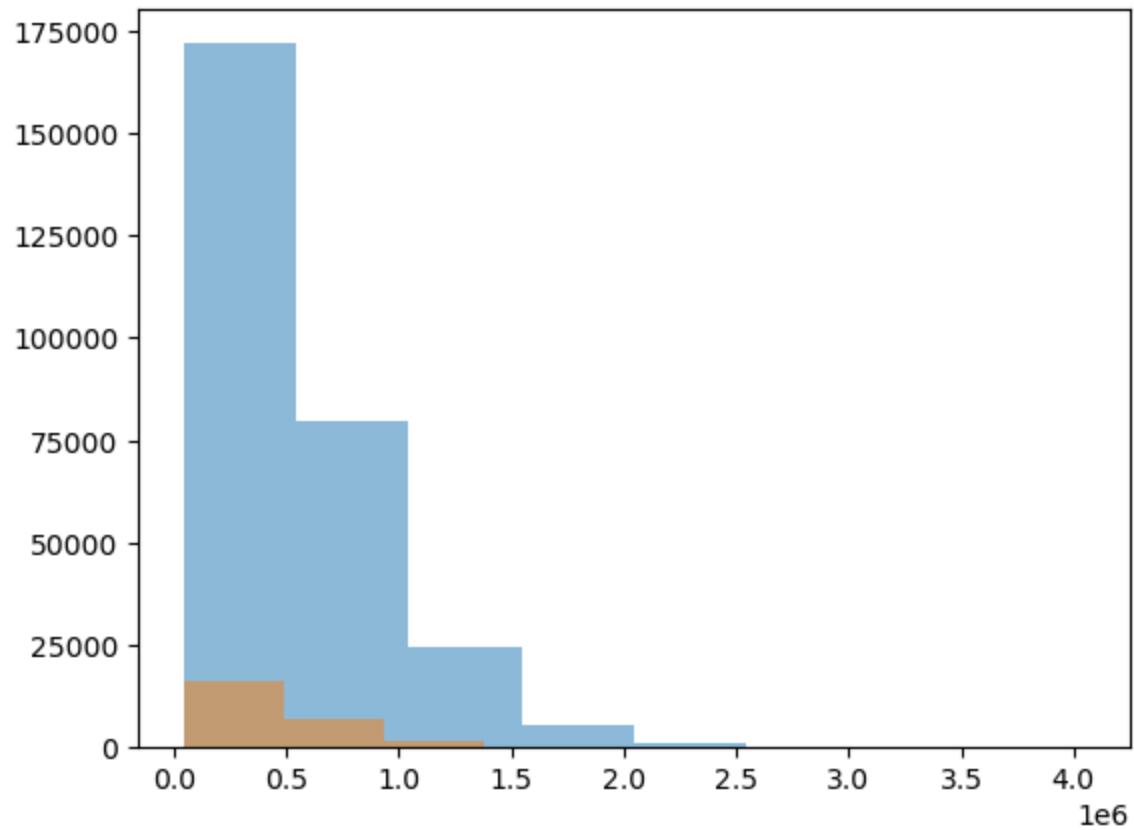
```
Out[86]: <Axes: xlabel='AMT_ANNUITY', ylabel='Density'>
```



Plot of AMT_GOODS_PRICE

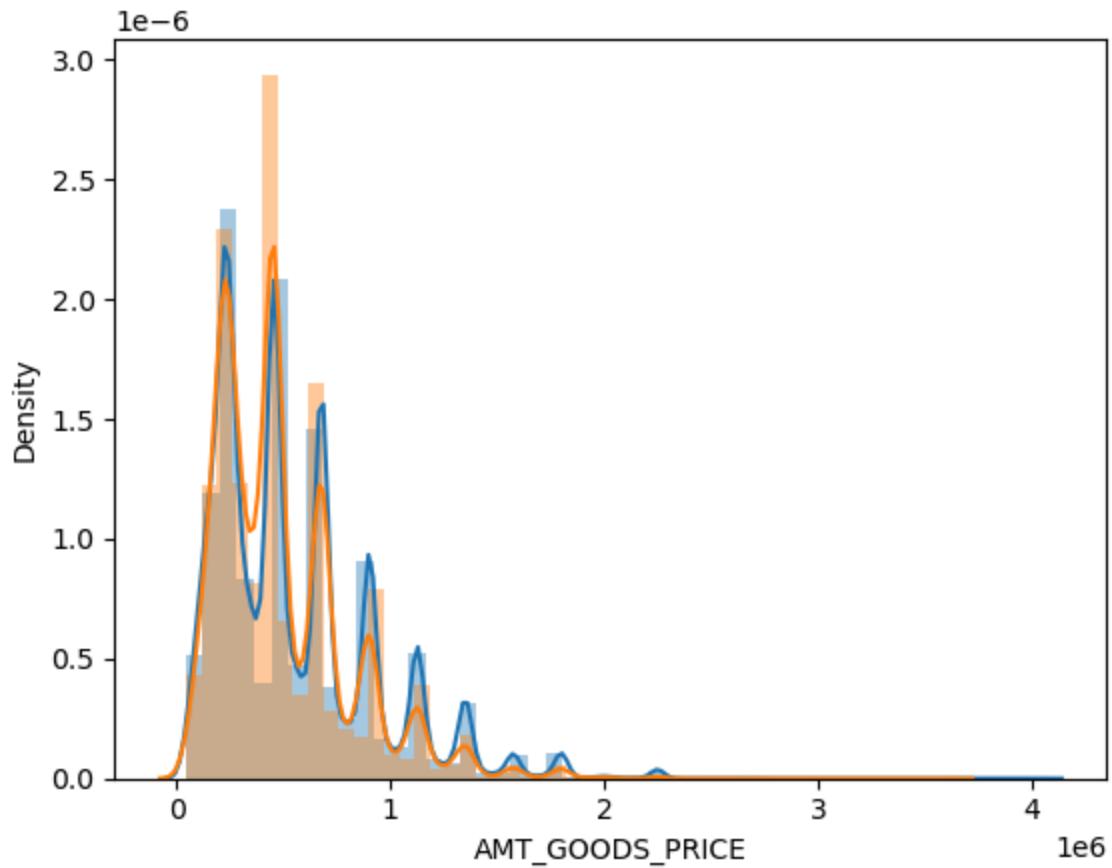
```
Out[86]: (array([1.72006e+05, 7.94740e+04, 2.46400e+04, 5.33700e+03, 9.21000e+02,
   2.20000e+01, 1.60000e+01, 1.30000e+01]),
 array([ 40500.,  541687.5, 1042875., 1544062.5, 2045250., 2546437.5,
   3047625., 3548812.5, 4050000.]),
 <BarContainer object of 8 artists>)

Out[86]: (array([1.6187e+04, 6.7090e+03, 1.5890e+03, 2.7100e+02, 4.3000e+01,
   3.0000e+00, 1.0000e+00, 1.0000e+00]),
 array([ 45000.,  489375.,  933750., 1378125., 1822500., 2266875.,
   2711250., 3155625., 3600000.]),
 <BarContainer object of 8 artists>)
```



Out[86]: <Axes: xlabel='AMT_GOODS_PRICE', ylabel='Density'>

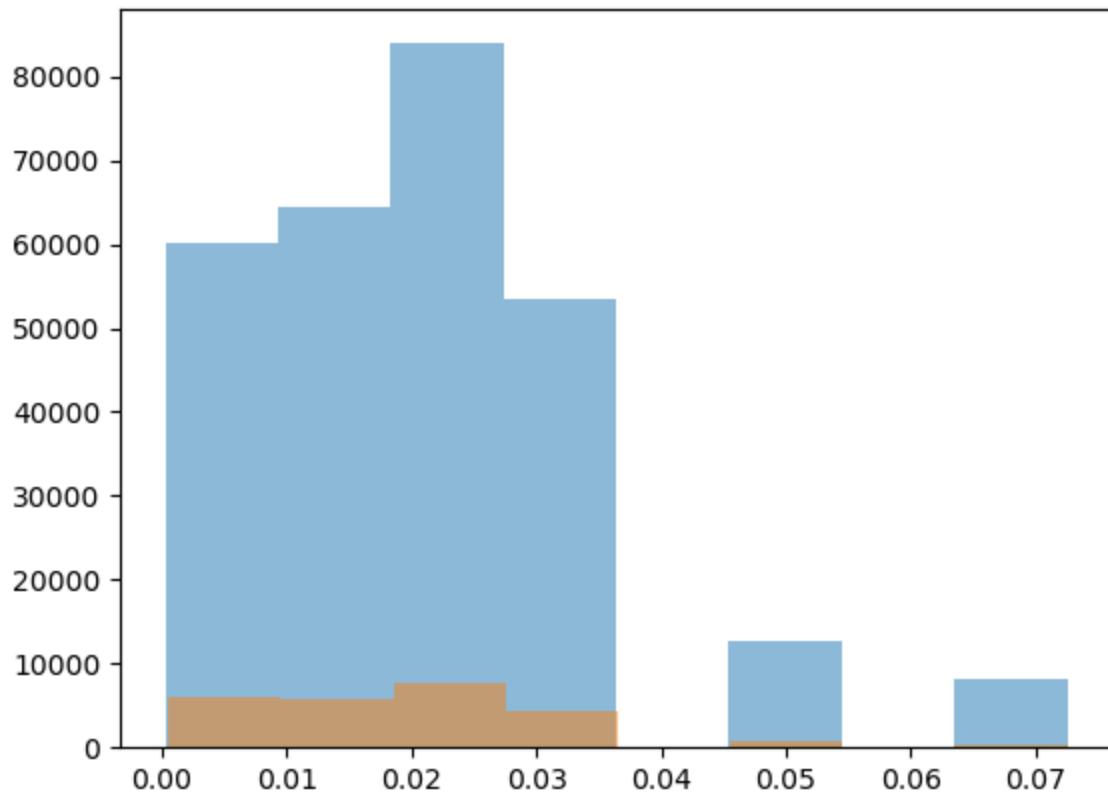
Out[86]: <Axes: xlabel='AMT_GOODS_PRICE', ylabel='Density'>



Plot of `REGION_POPULATION_RELATIVE`

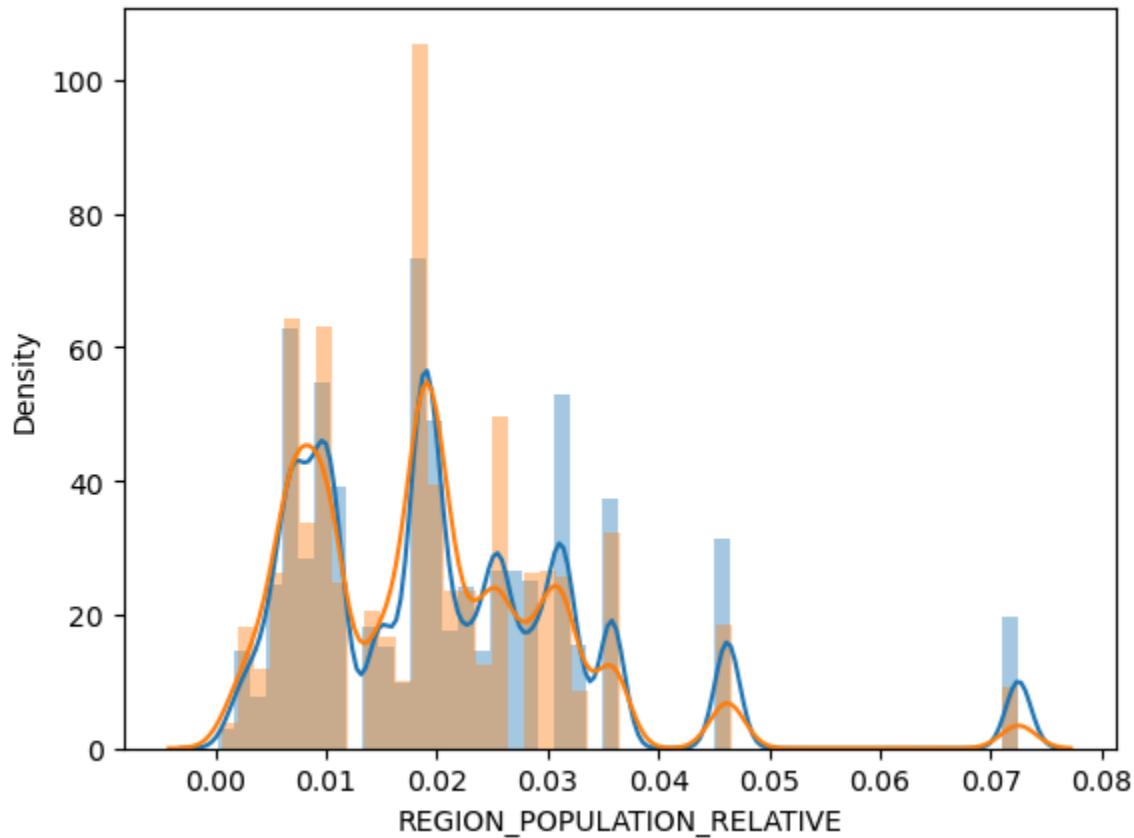
```
Out[86]: (array([60126., 64370., 83909., 53424.,      0., 12779.,      0., 8078.]),
          array([0.00029    , 0.00931725, 0.0183445 , 0.02737175, 0.036399    ,
                 0.04542625, 0.0544535 , 0.06348075, 0.072508    ]),
          <BarContainer object of 8 artists>)

Out[86]: (array([5952., 5895., 7720., 4261.,      0., 663.,      0., 334.]),
          array([0.000533   , 0.00952988, 0.01852675, 0.02752362, 0.0365205   ,
                 0.04551737, 0.05451425, 0.06351113, 0.072508    ]),
          <BarContainer object of 8 artists>)
```



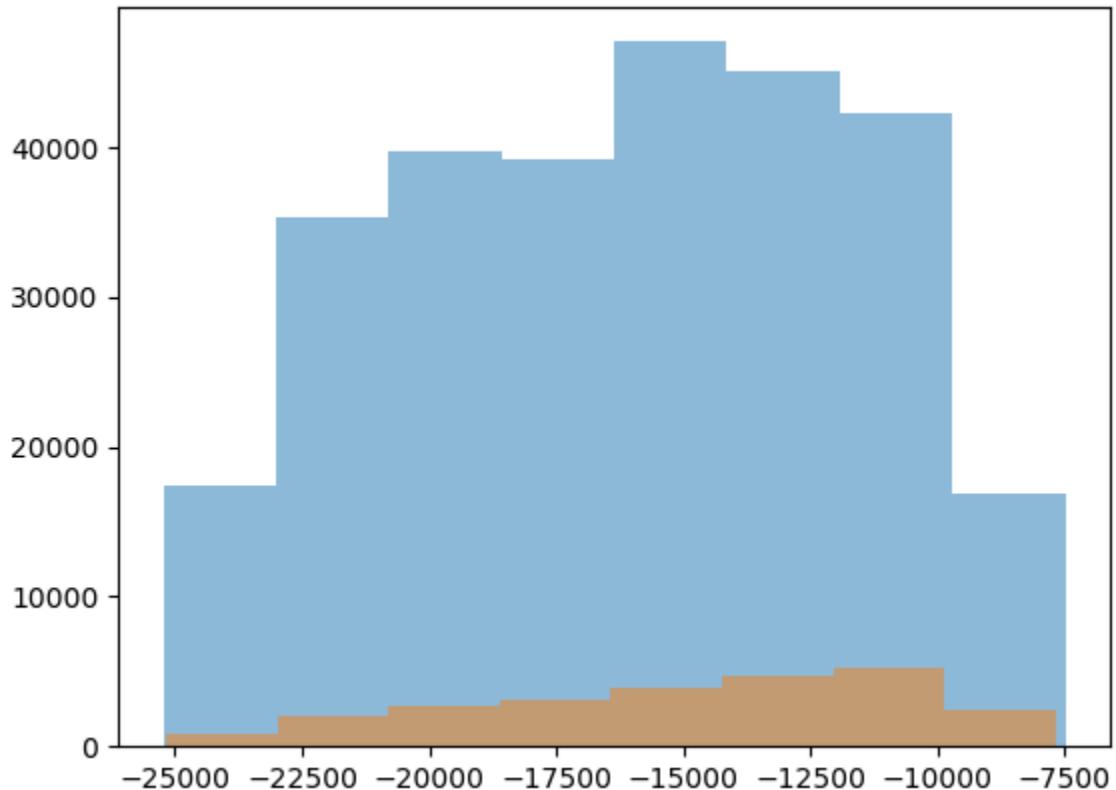
```
Out[86]: <Axes: xlabel='REGION_POPULATION_RELATIVE', ylabel='Density'>
```

```
Out[86]: <Axes: xlabel='REGION_POPULATION_RELATIVE', ylabel='Density'>
```



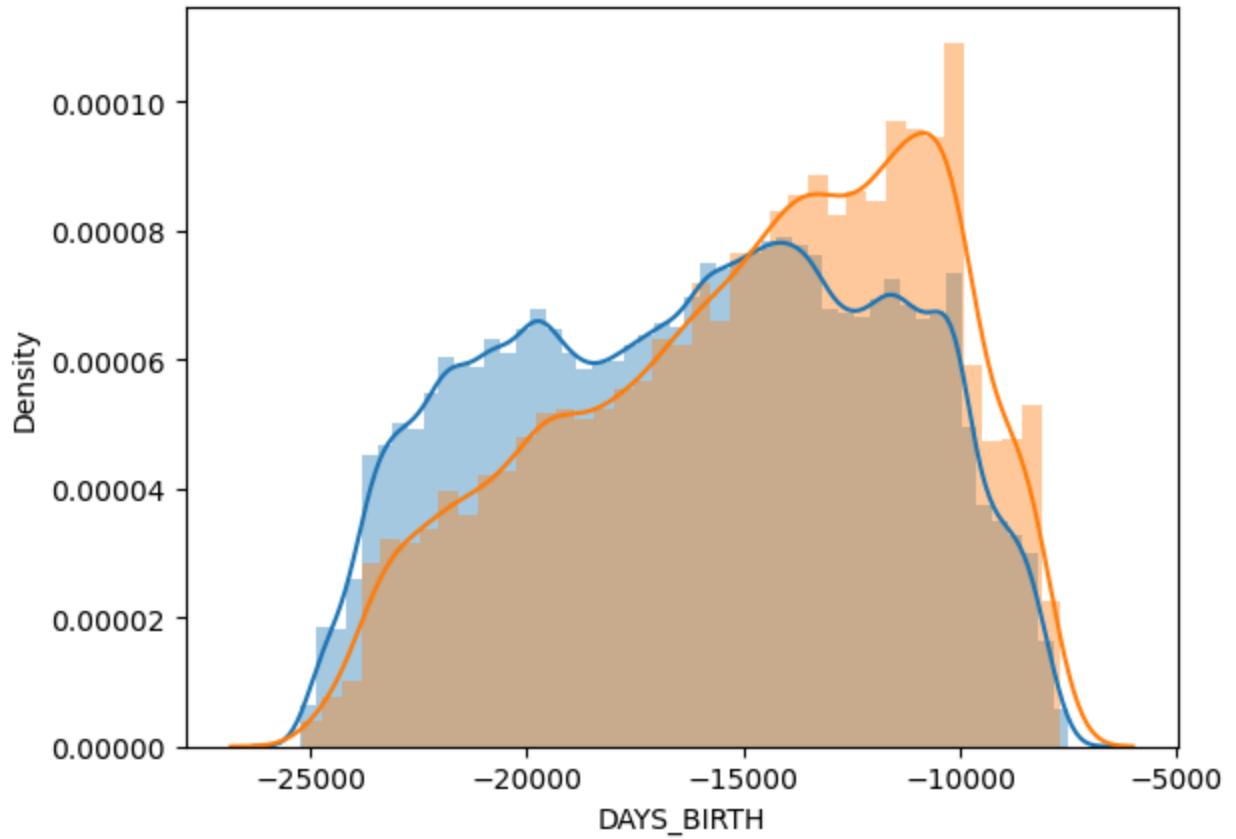
Plot of DAYS_BIRTH

```
Out[86]: (array([17373., 35279., 39653., 39235., 47012., 45019., 42291., 16824.]),  
          array([-25229. , -23011.5, -20794. , -18576.5, -16359. , -14141.5,  
                 -11924. , -9706.5, -7489. ]),  
          <BarContainer object of 8 artists>)  
  
Out[86]: (array([ 854., 1975., 2647., 3055., 3971., 4650., 5267., 2406.]),  
          array([-25168. , -22981.75, -20795.5 , -18609.25, -16423. , -14236.75,  
                 -12050.5 , -9864.25, -7678. ]),  
          <BarContainer object of 8 artists>)
```



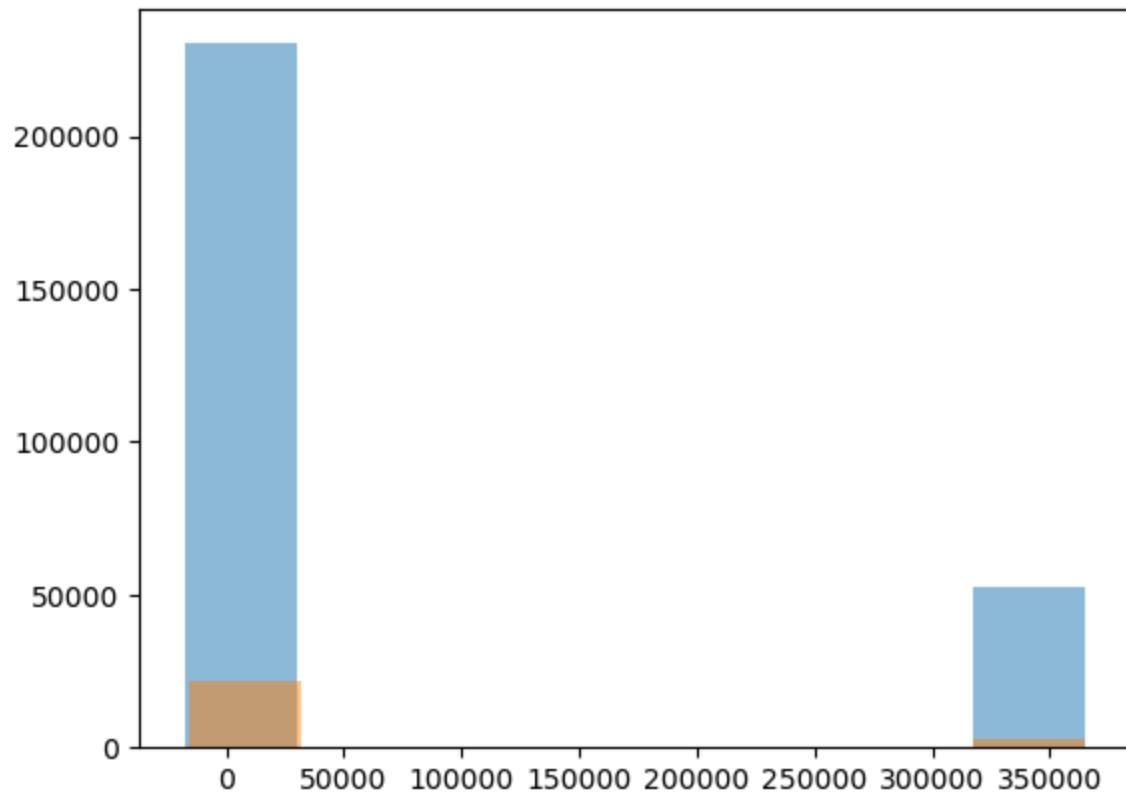
```
Out[86]: <Axes: xlabel='DAYS_BIRTH', ylabel='Density'>
```

```
Out[86]: <Axes: xlabel='DAYS_BIRTH', ylabel='Density'>
```



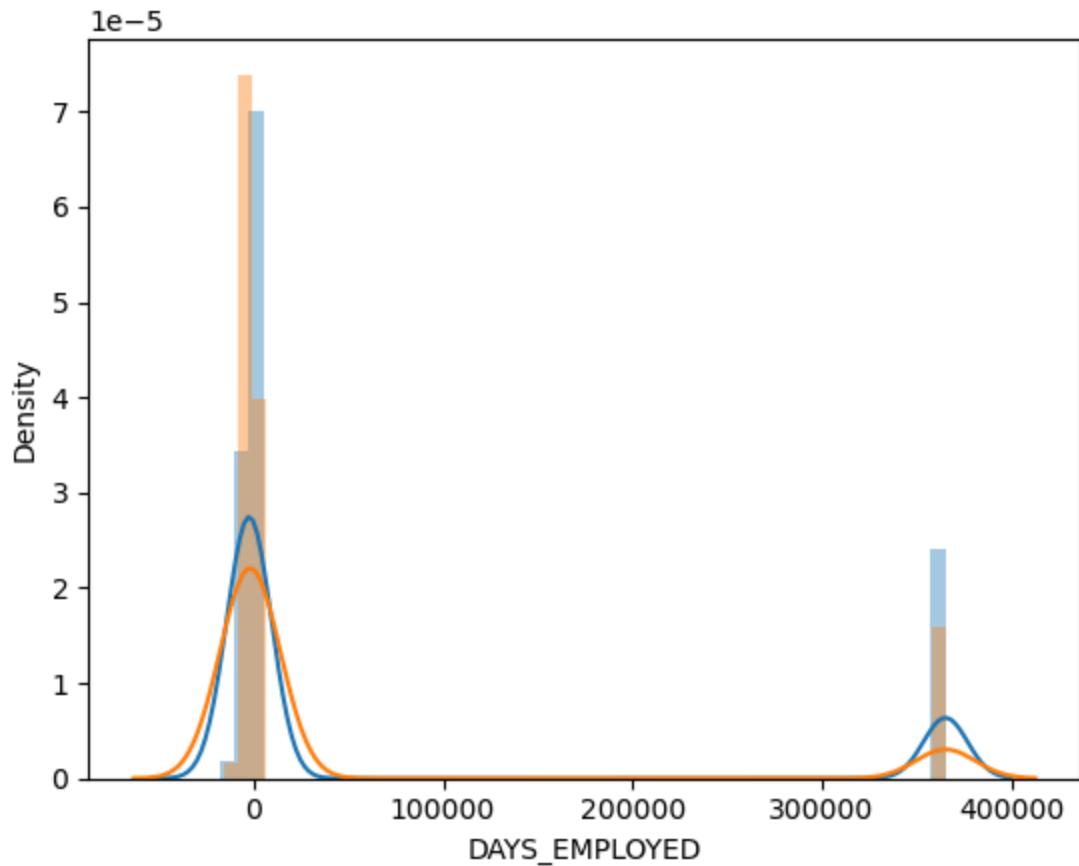
Plot of DAYS_EMPLOYED

```
Out[86]: (array([230302.,      0.,      0.,      0.,      0.,      0.,      0.,      0.,
   52384.]),
 array([-17912. ,  29982.375,  77876.75 , 125771.125, 173665.5 ,
 221559.875, 269454.25 , 317348.625, 365243. ]),
 <BarContainer object of 8 artists>
)
Out[86]: (array([21835.,      0.,      0.,      0.,      0.,      0.,      0.,      2990.]),
 array([-16069.,  31595.,  79259., 126923., 174587., 222251., 269915.,
 317579., 365243.]),
 <BarContainer object of 8 artists>
)
```



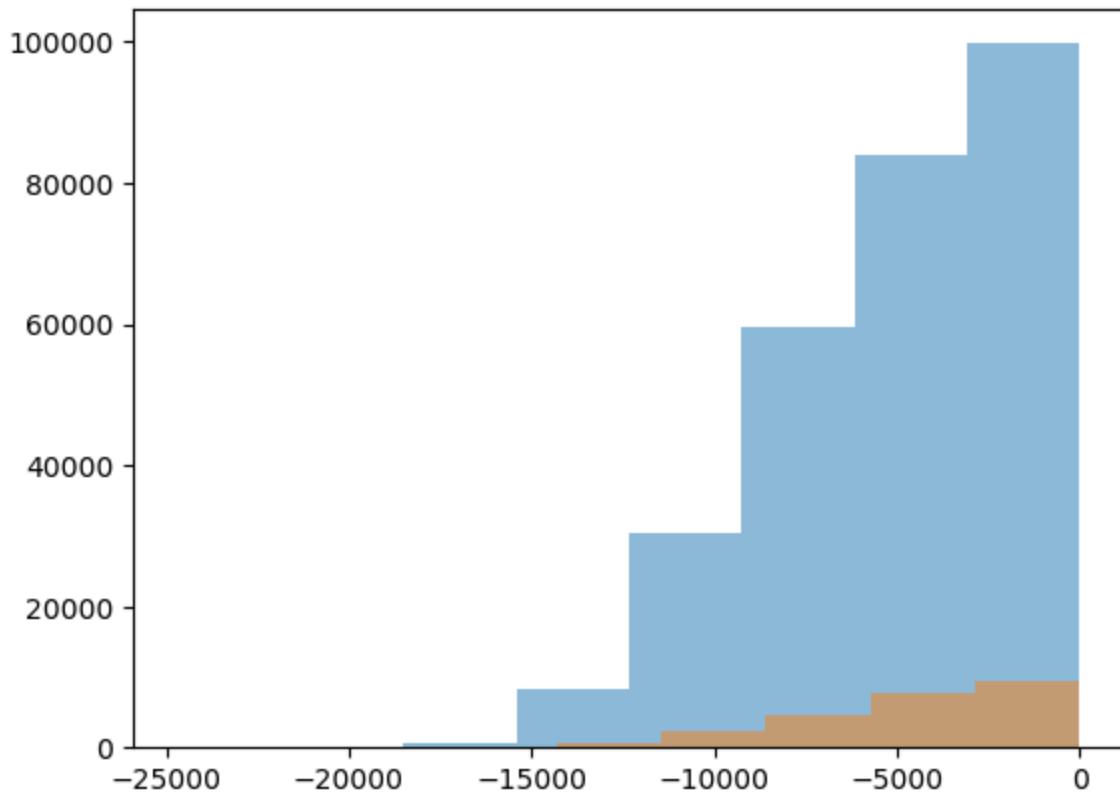
```
Out[86]: <Axes: xlabel='DAYS_EMPLOYED', ylabel='Density'>
```

```
Out[86]: <Axes: xlabel='DAYS_EMPLOYED', ylabel='Density'>
```



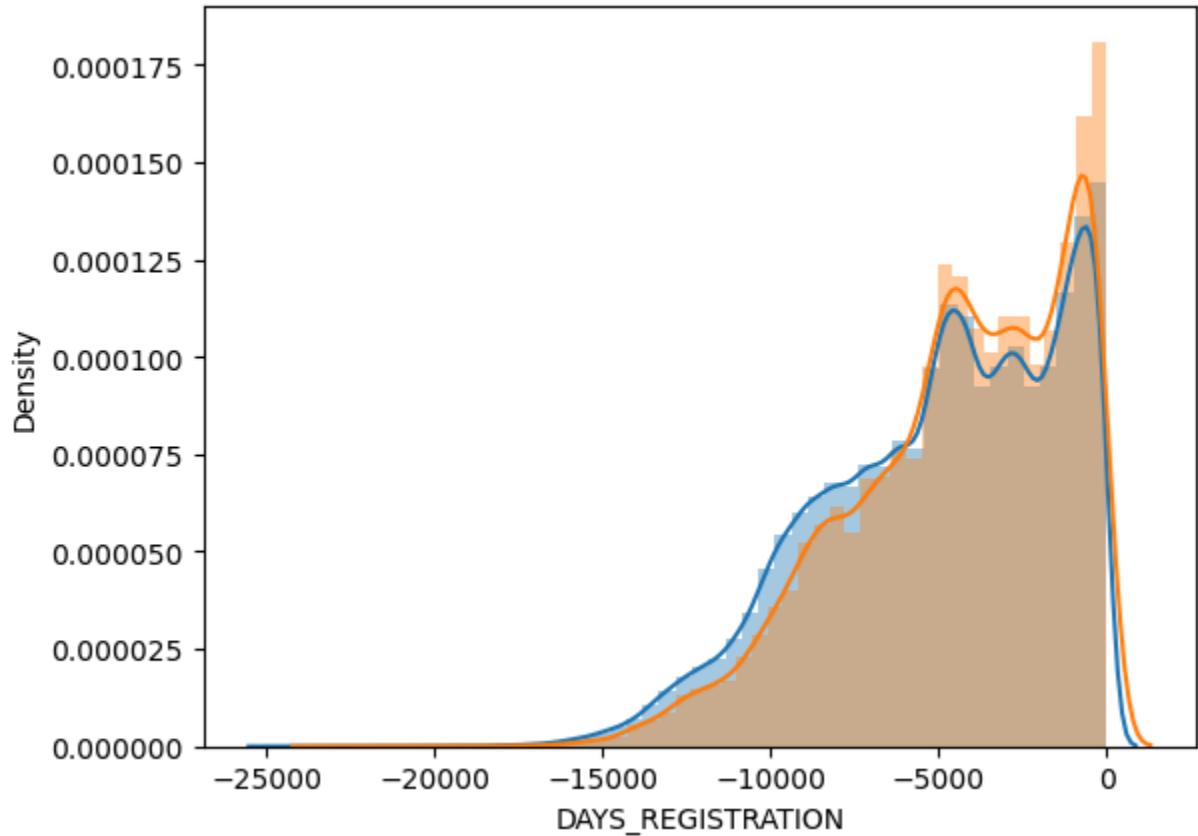
Plot of DAYS_REGISTRATION

```
Out[86]: (array([1.0000e+01, 5.5000e+01, 7.4700e+02, 8.2060e+03, 3.0445e+04,
      5.9532e+04, 8.3959e+04, 9.9732e+04]),  
array([-24672., -21588., -18504., -15420., -12336., -9252., -6168.,
      -3084., 0.]),  
<BarContainer object of 8 artists>)  
Out[86]: (array([1.000e+00, 8.000e+00, 6.900e+01, 7.450e+02, 2.420e+03, 4.613e+03,
      7.671e+03, 9.298e+03]),  
array([-22928., -20062., -17196., -14330., -11464., -8598., -5732.,
      -2866., 0.]),  
<BarContainer object of 8 artists>)
```



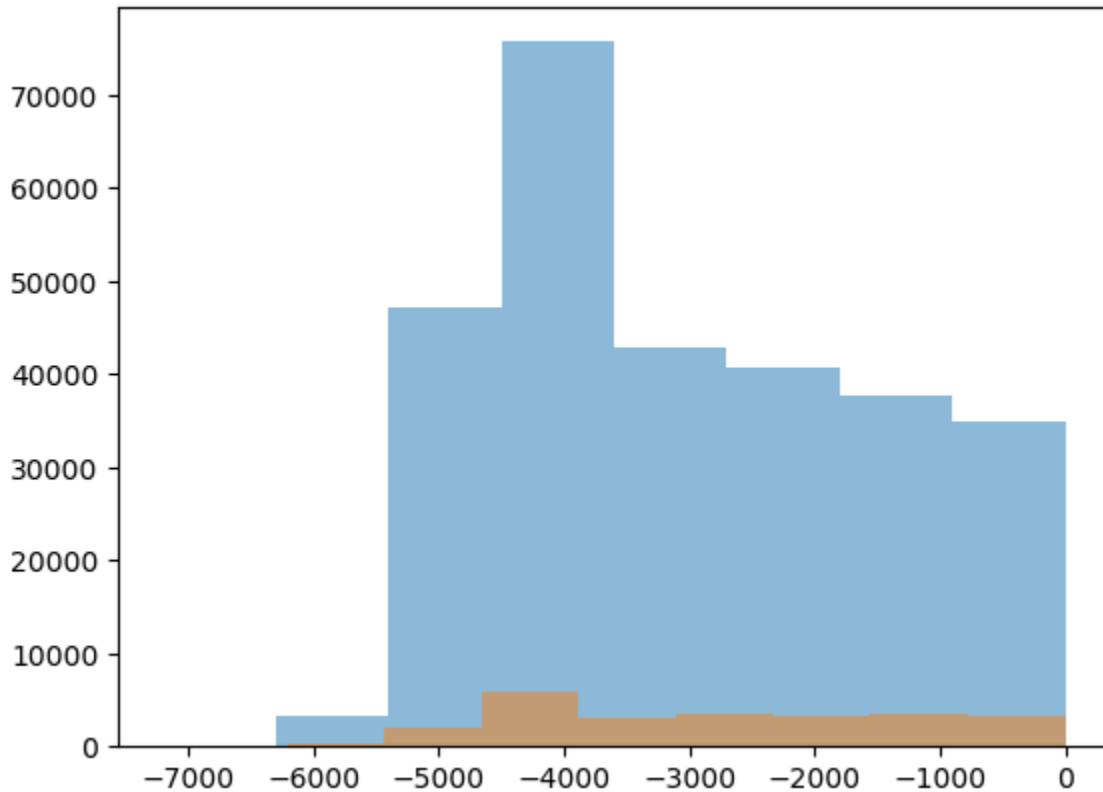
```
Out[86]: <Axes: xlabel='DAYS_REGISTRATION', ylabel='Density'>
```

```
Out[86]: <Axes: xlabel='DAYS_REGISTRATION', ylabel='Density'>
```



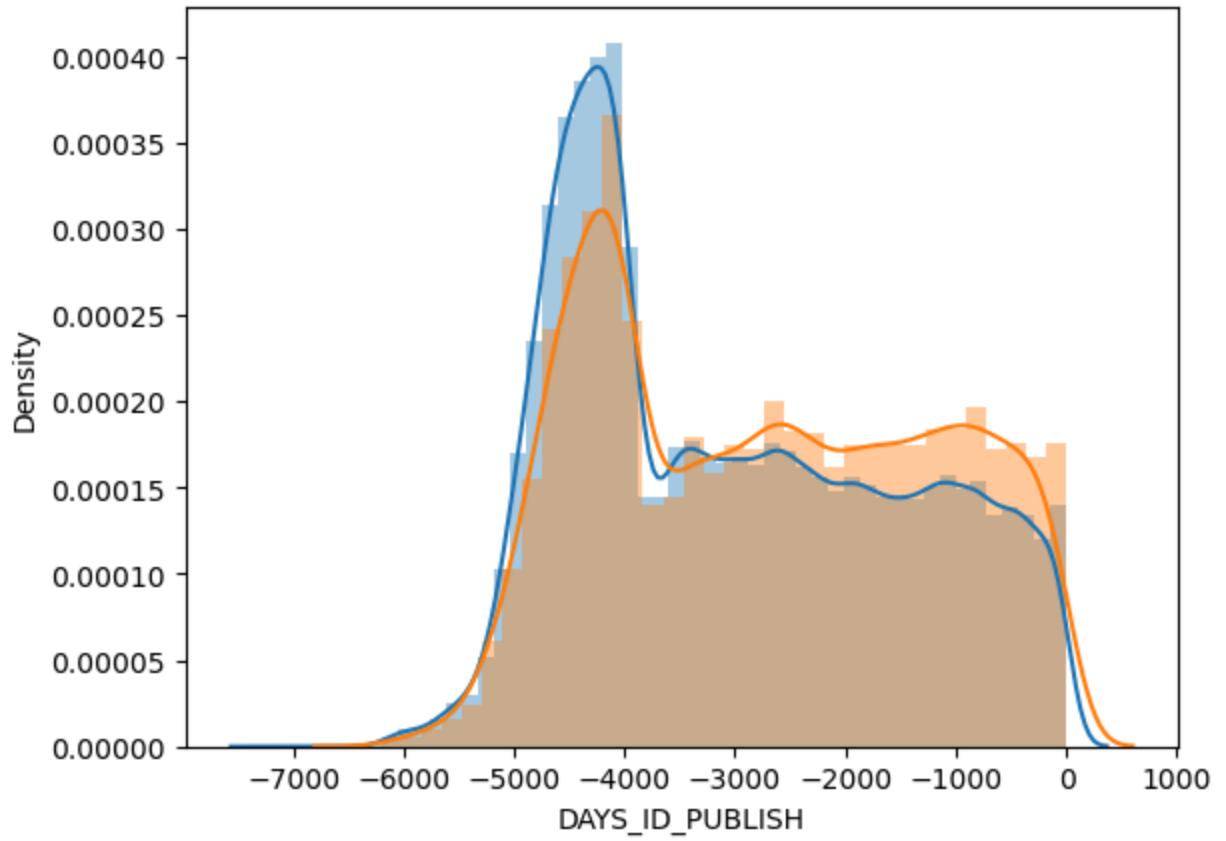
Plot of DAYS_ID_PUBLISH

```
Out[86]: (array([4.0000e+00, 3.2880e+03, 4.7288e+04, 7.5758e+04, 4.2894e+04,
   4.0810e+04, 3.7662e+04, 3.4982e+04]),  
 array([-7197. , -6297.375, -5397.75 , -4498.125, -3598.5 , -2698.875,  
 -1799.25 , -899.625, 0. ]),  
<BarContainer object of 8 artists>)  
Out[86]: (array([ 185., 2061., 5852., 2994., 3520., 3315., 3535., 3363.]),  
 array([-6207. , -5431.125, -4655.25 , -3879.375, -3103.5 , -2327.625,  
 -1551.75 , -775.875, 0. ]),  
<BarContainer object of 8 artists>)
```



```
Out[86]: <Axes: xlabel='DAYS_ID_PUBLISH', ylabel='Density'>
```

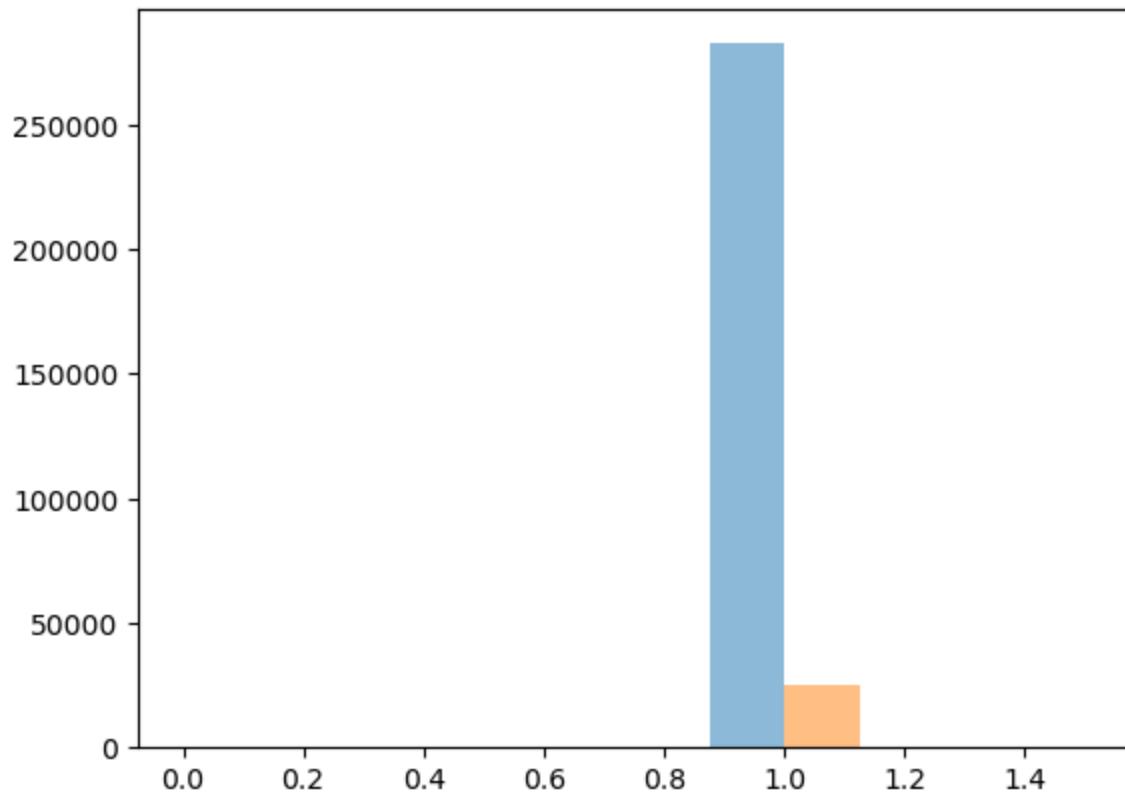
```
Out[86]: <Axes: xlabel='DAYS_ID_PUBLISH', ylabel='Density'>
```



Plot of FLAG_MOBIL

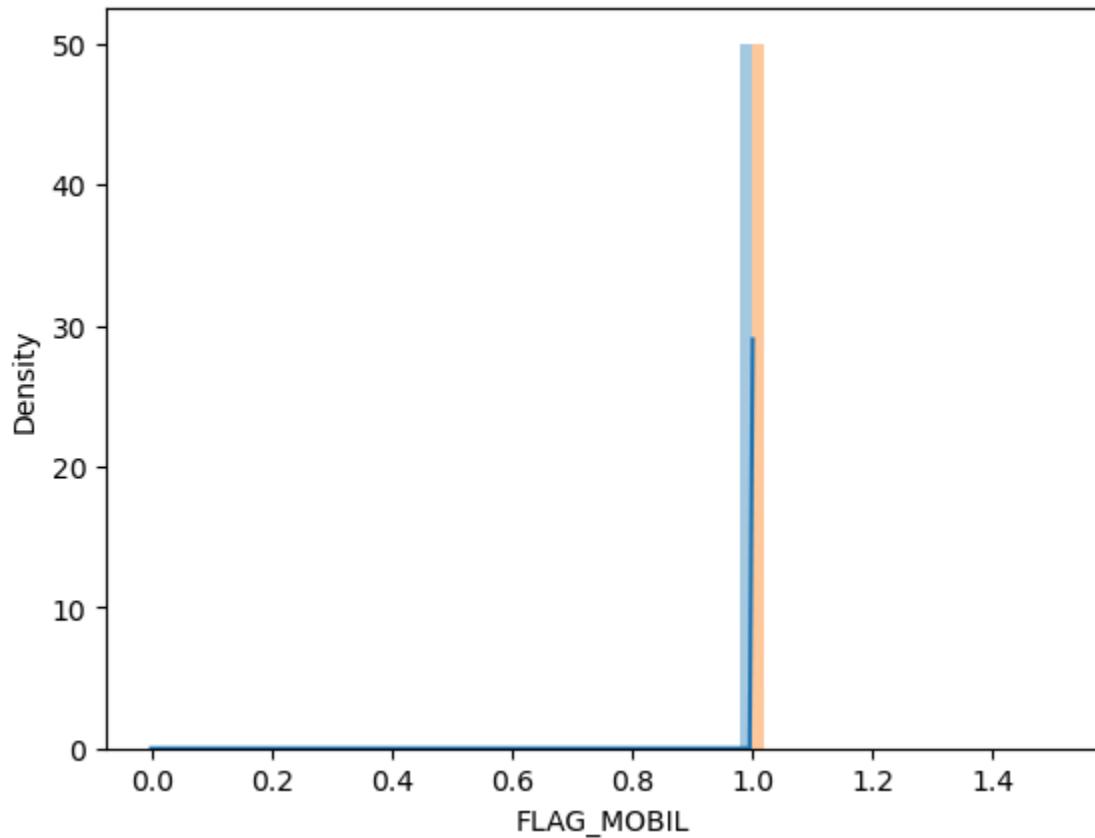
```
Out[86]: (array([1.00000e+00, 0.00000e+00, 0.00000e+00, 0.00000e+00, 0.00000e+00,
       0.00000e+00, 0.00000e+00, 2.82685e+05]),
       array([0.    , 0.125, 0.25 , 0.375, 0.5   , 0.625, 0.75 , 0.875, 1.    ]),
       <BarContainer object of 8 artists>)

Out[86]: (array([ 0.,  0.,  0.,  0., 24825.,  0.,  0.,  0.]),
       array([0.5   , 0.625, 0.75 , 0.875, 1.    , 1.125, 1.25 , 1.375, 1.5   ]),
       <BarContainer object of 8 artists>)
```



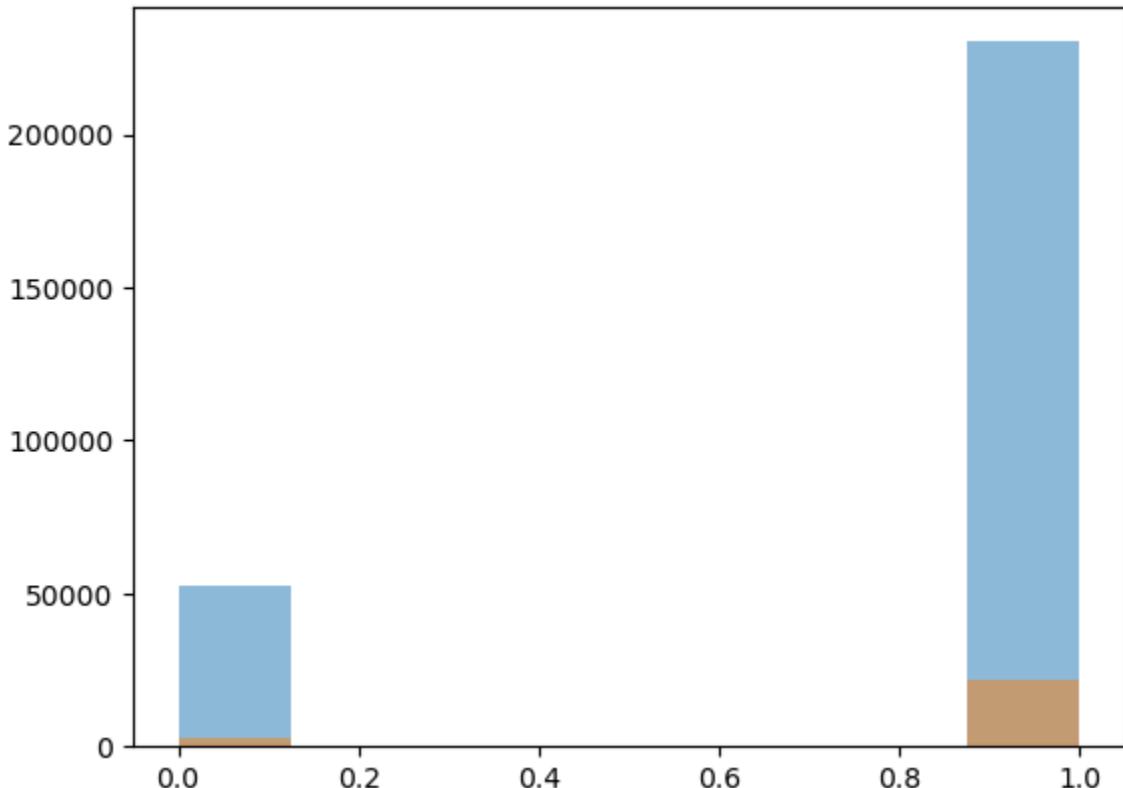
```
Out[86]: <Axes: xlabel='FLAG_MOBIL', ylabel='Density'>
```

```
Out[86]: <Axes: xlabel='FLAG_MOBIL', ylabel='Density'>
```



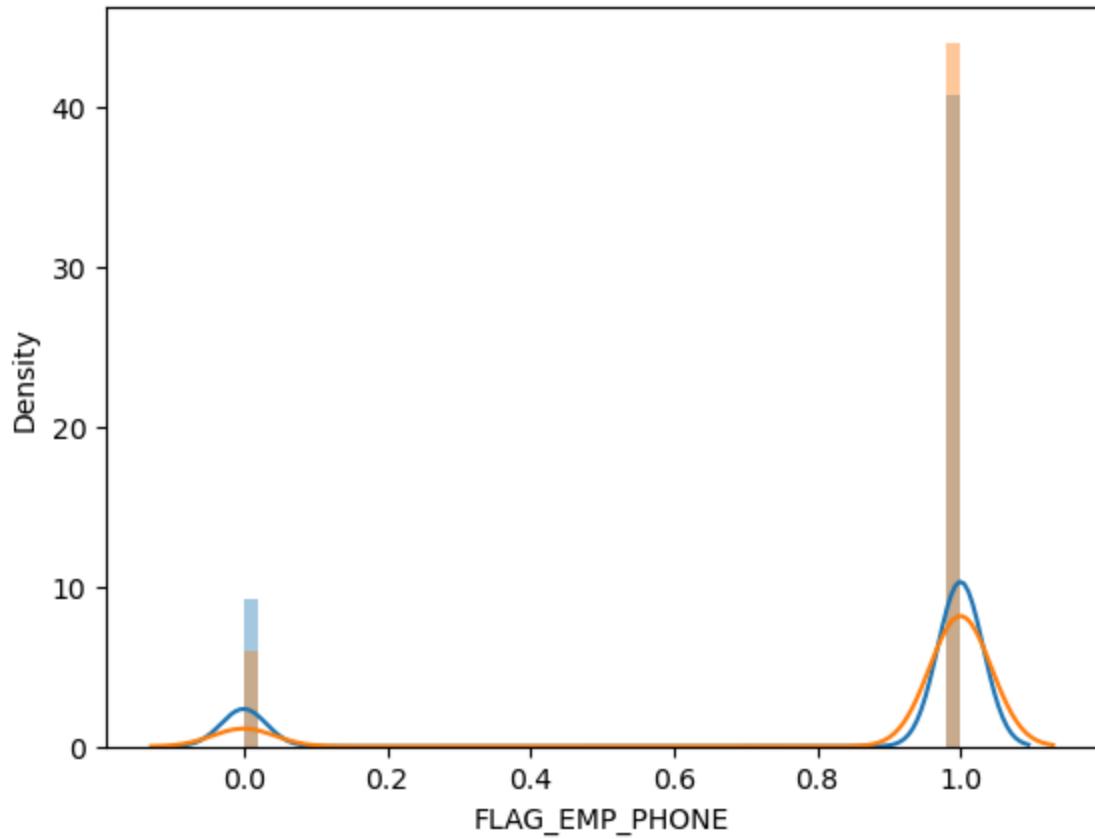
Plot of FLAG_EMP_PHONE

```
Out[86]: (array([ 52395.,      0.,      0.,      0.,      0.,      0.,
   230291.]),
 array([0.    , 0.125, 0.25 , 0.375, 0.5  , 0.625, 0.75 , 0.875, 1.    ]),
 <BarContainer object of 8 artists>
Out[86]: (array([ 2991.,      0.,      0.,      0.,      0.,      0.,      0.,
  21834.]),
 array([0.    , 0.125, 0.25 , 0.375, 0.5  , 0.625, 0.75 , 0.875, 1.    ]),
 <BarContainer object of 8 artists>)
```



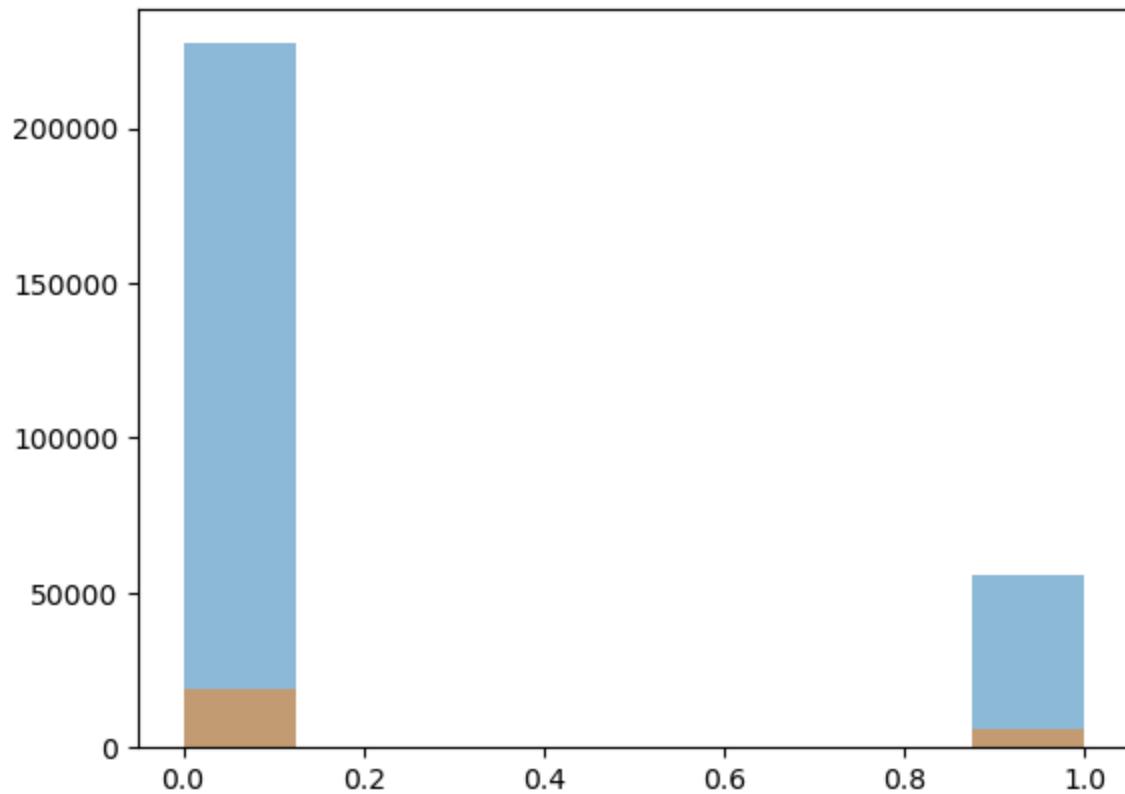
```
Out[86]: <Axes: xlabel='FLAG_EMP_PHONE', ylabel='Density'>
```

```
Out[86]: <Axes: xlabel='FLAG_EMP_PHONE', ylabel='Density'>
```



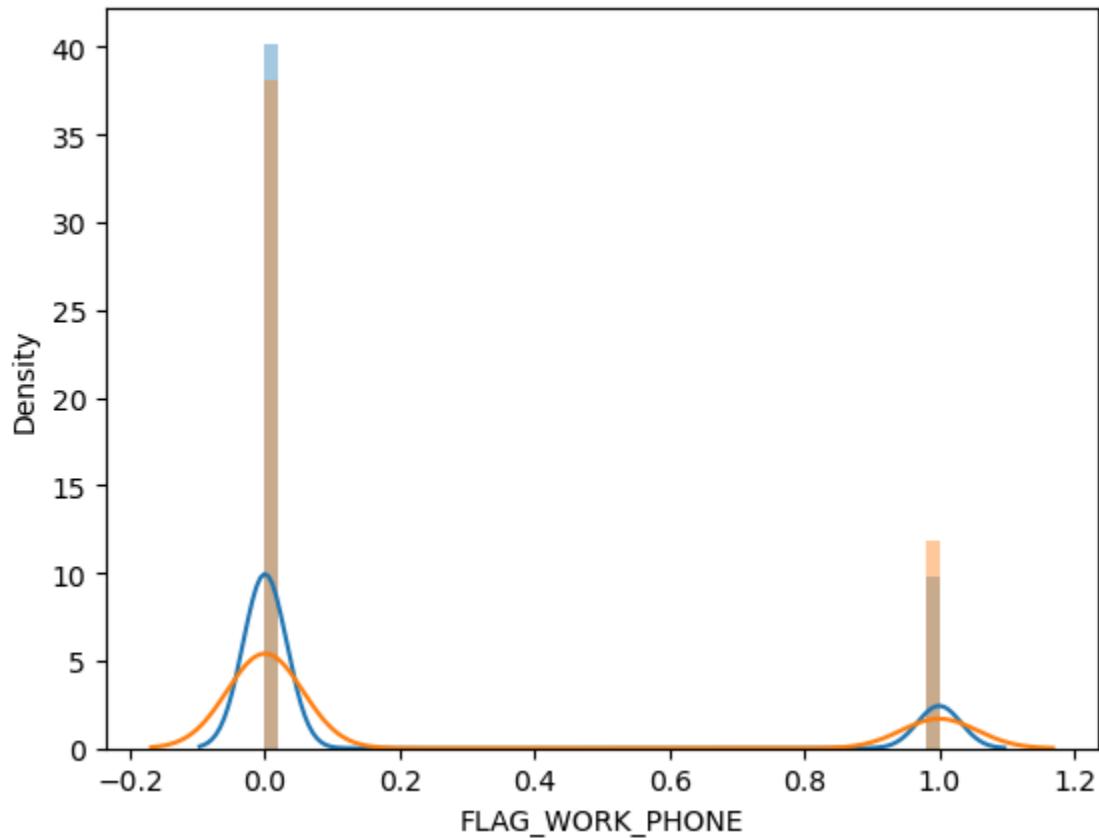
Plot of `FLAG_WORK_PHONE`

```
Out[86]: (array([227282.,      0.,      0.,      0.,      0.,      0.,      0.,  
      55404.]),  
 array([0.     , 0.125, 0.25 , 0.375, 0.5   , 0.625, 0.75 , 0.875, 1.     ]),  
 <BarContainer object of 8 artists>)  
Out[86]: (array([18921.,      0.,      0.,      0.,      0.,      0.,      0.,      5904.]),  
 array([0.     , 0.125, 0.25 , 0.375, 0.5   , 0.625, 0.75 , 0.875, 1.     ]),  
 <BarContainer object of 8 artists>)
```



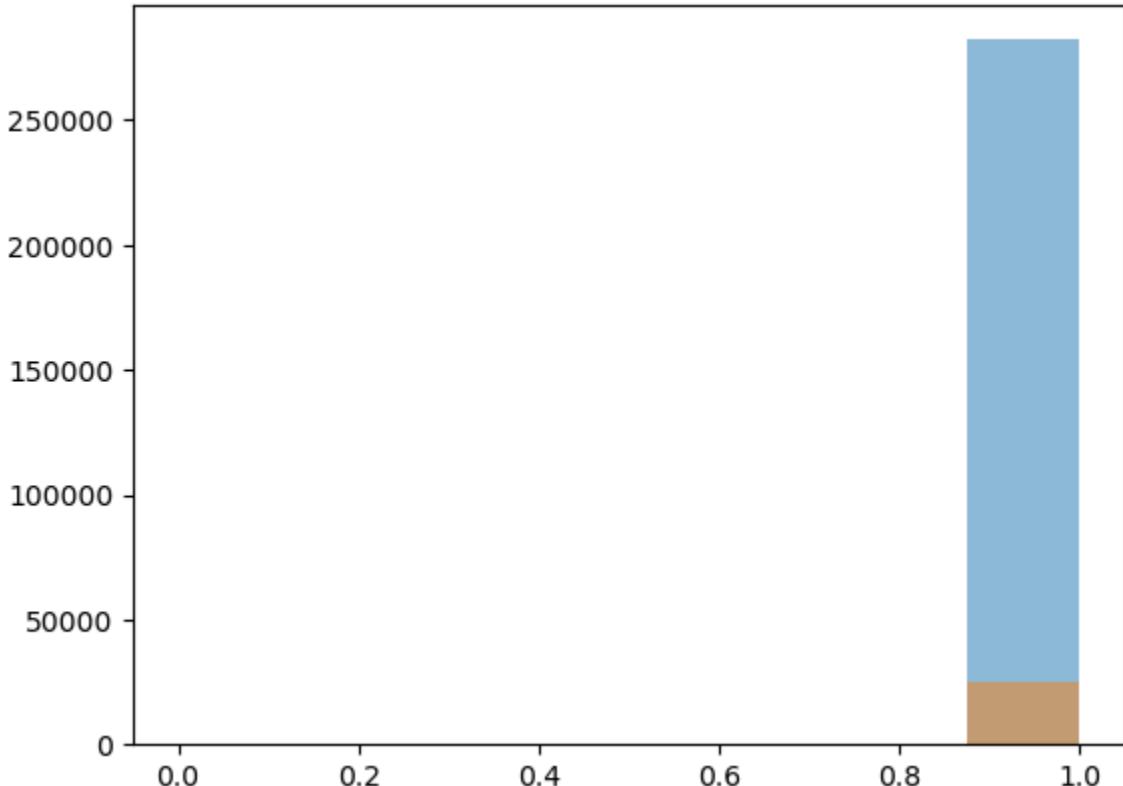
```
Out[86]: <Axes: xlabel='FLAG_WORK_PHONE', ylabel='Density'>
```

```
Out[86]: <Axes: xlabel='FLAG_WORK_PHONE', ylabel='Density'>
```



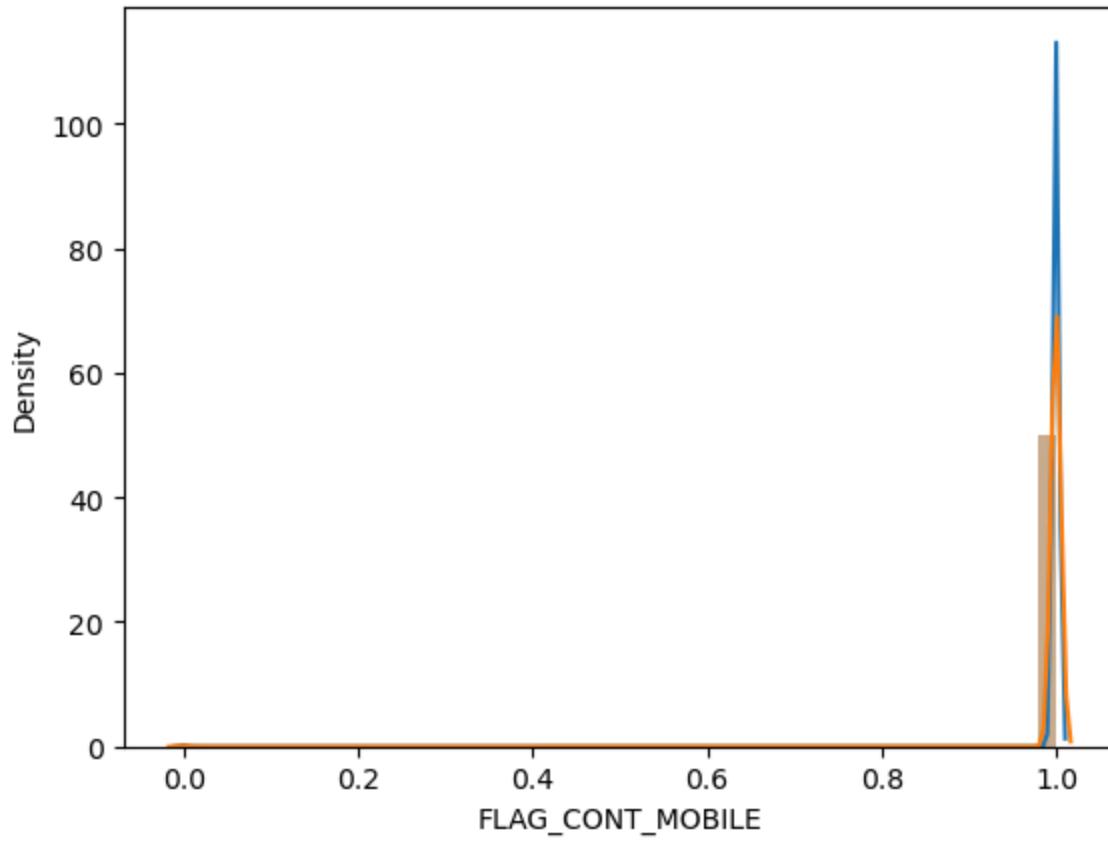
Plot of FLAG_CONT_MOBILE

```
Out[86]: (array([ 529.,      0.,      0.,      0.,      0.,      0.,
   282157.]),
 array([0.    , 0.125, 0.25 , 0.375, 0.5  , 0.625, 0.75 , 0.875, 1.    ]),
 <BarContainer object of 8 artists>
Out[86]: (array([ 45.,      0.,      0.,      0.,      0.,      0.,      0., 24780.]),
 array([0.    , 0.125, 0.25 , 0.375, 0.5  , 0.625, 0.75 , 0.875, 1.    ]),
 <BarContainer object of 8 artists>)
```



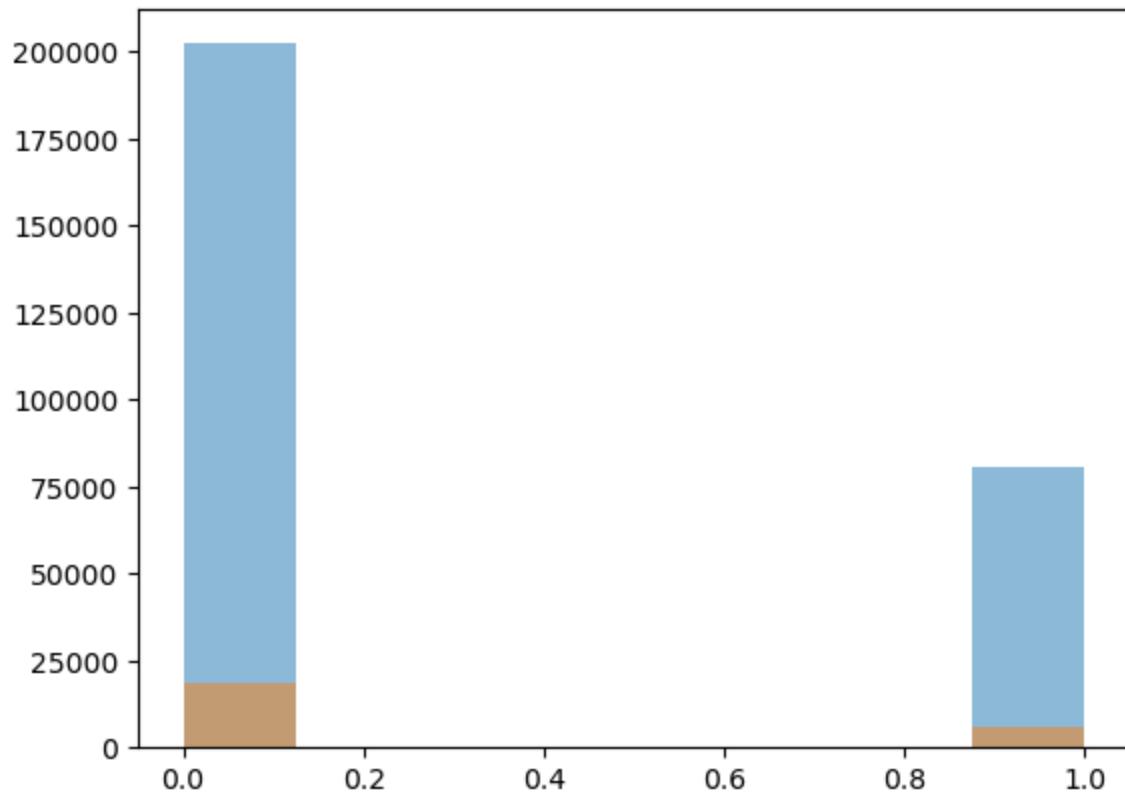
```
Out[86]: <Axes: xlabel='FLAG_CONT_MOBILE', ylabel='Density'>
```

```
Out[86]: <Axes: xlabel='FLAG_CONT_MOBILE', ylabel='Density'>
```



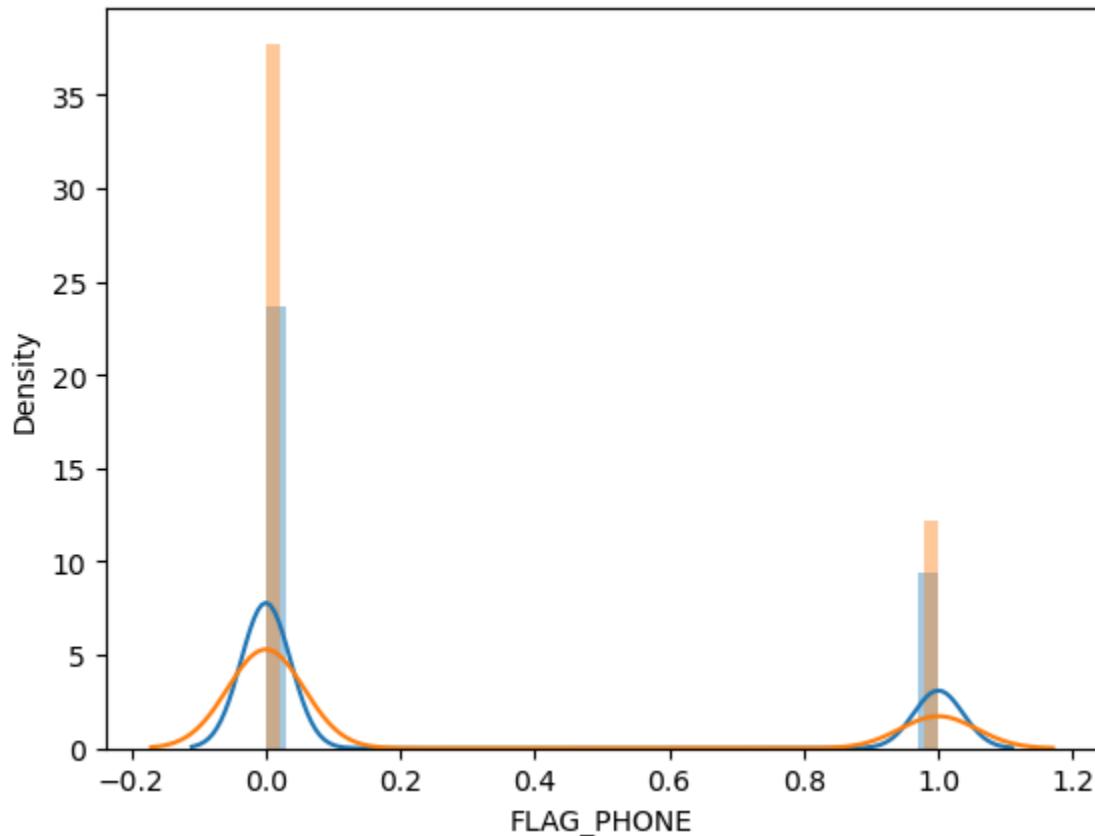
Plot of `FLAG_PHONE`

```
Out[86]: (array([202336.,      0.,      0.,      0.,      0.,      0.,      0.,      0.,  
     80350.]),  
 array([0.    , 0.125, 0.25 , 0.375, 0.5   , 0.625, 0.75 , 0.875, 1.    ]),  
 <BarContainer object of 8 artists>)  
Out[86]: (array([18744.,      0.,      0.,      0.,      0.,      0.,      0.,      0.,  
    6081.]),  
 array([0.    , 0.125, 0.25 , 0.375, 0.5   , 0.625, 0.75 , 0.875, 1.    ]),  
 <BarContainer object of 8 artists>)
```



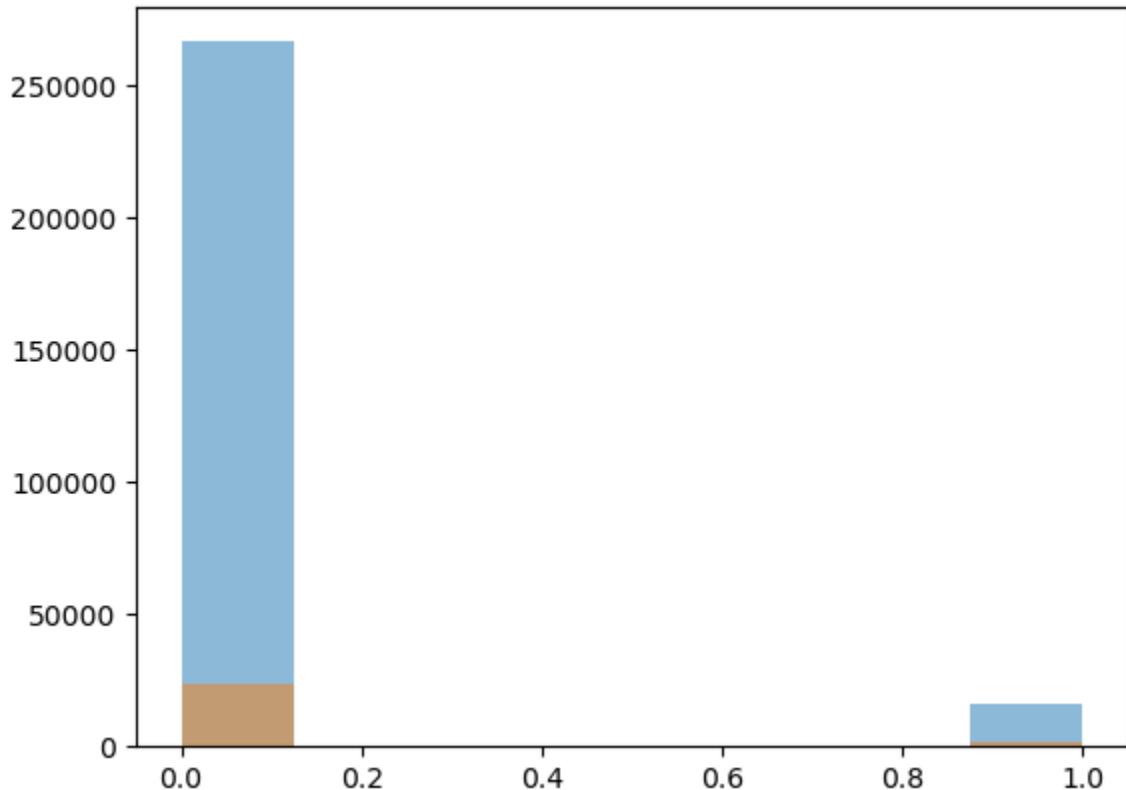
```
Out[86]: <Axes: xlabel='FLAG_PHONE', ylabel='Density'>
```

```
Out[86]: <Axes: xlabel='FLAG_PHONE', ylabel='Density'>
```



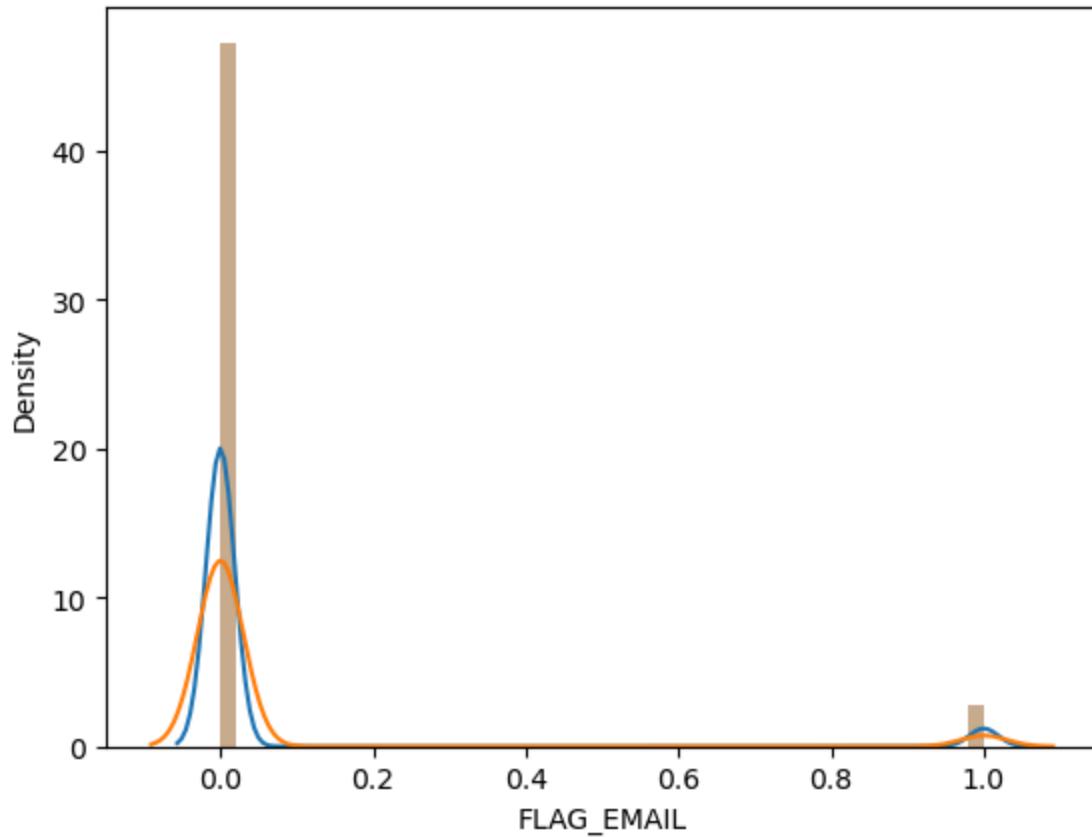
Plot of FLAG_EMAIL

```
Out[86]: (array([266618.,      0.,      0.,      0.,      0.,      0.,      0.,
   16068.]),
 array([0.    , 0.125, 0.25 , 0.375, 0.5  , 0.625, 0.75 , 0.875, 1.    ]),
 <BarContainer object of 8 artists>
Out[86]: (array([23451.,      0.,      0.,      0.,      0.,      0.,      0.,      1374.]),
 array([0.    , 0.125, 0.25 , 0.375, 0.5  , 0.625, 0.75 , 0.875, 1.    ]),
 <BarContainer object of 8 artists>)
```



```
Out[86]: <Axes: xlabel='FLAG_EMAIL', ylabel='Density'>
```

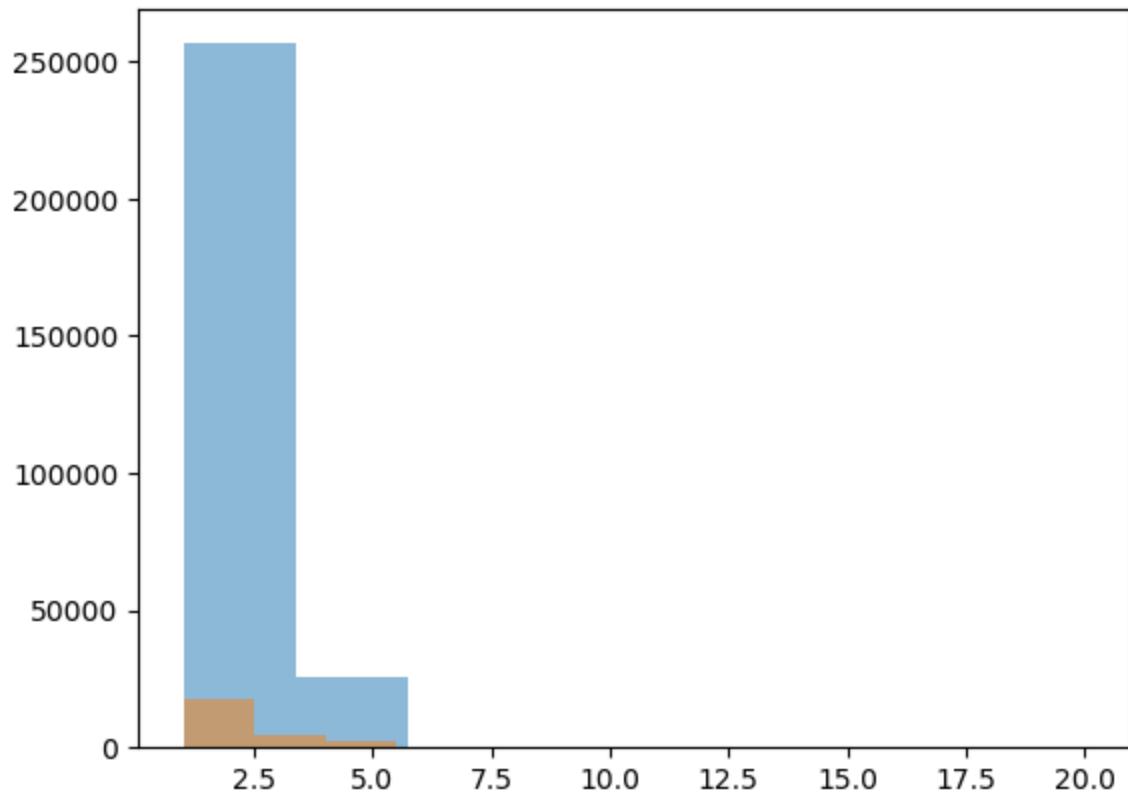
```
Out[86]: <Axes: xlabel='FLAG_EMAIL', ylabel='Density'>
```



Plot of CNT_FAM_MEMBERS

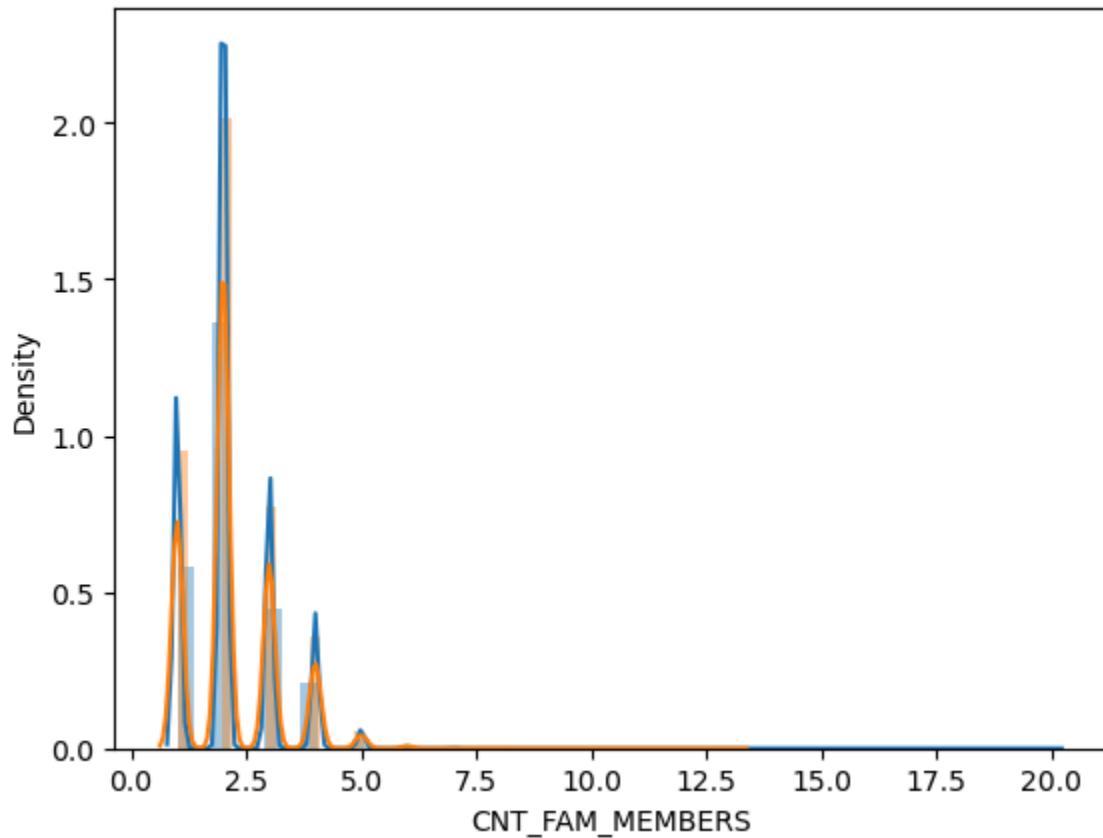
```
Out[86]: (array([2.56513e+05, 2.57120e+04, 4.42000e+02, 8.00000e+00, 2.00000e+00,
       3.00000e+00, 2.00000e+00, 2.00000e+00]),
       array([ 1. ,  3.375,  5.75 ,  8.125, 10.5 , 12.875, 15.25 , 17.625,
              20. ]),
       <BarContainer object of 8 artists>)

Out[86]: (array([1.7684e+04, 4.6080e+03, 2.4630e+03, 5.5000e+01, 1.2000e+01,
       0.0000e+00, 2.0000e+00, 1.0000e+00]),
       array([ 1. ,  2.5,  4. ,  5.5,  7. ,  8.5, 10. , 11.5, 13. ]),
       <BarContainer object of 8 artists>)
```



```
Out[86]: <Axes: xlabel='CNT_FAM_MEMBERS', ylabel='Density'>
```

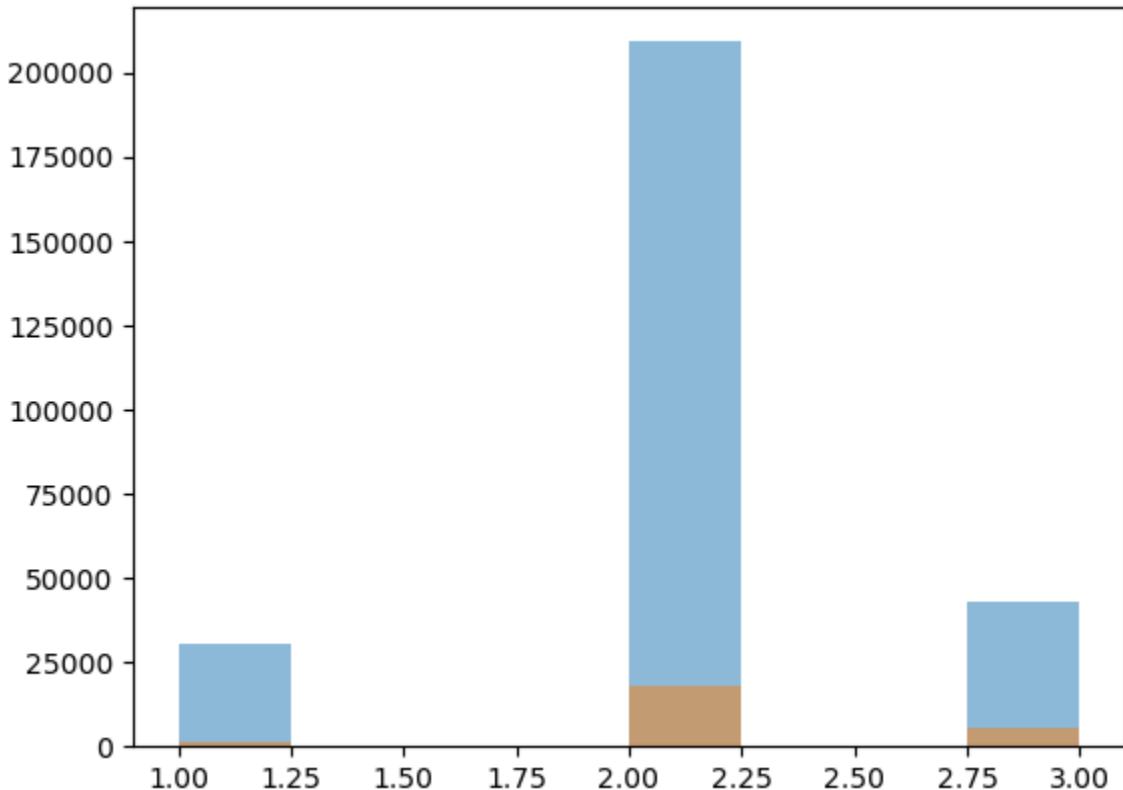
```
Out[86]: <Axes: xlabel='CNT_FAM_MEMBERS', ylabel='Density'>
```



Plot of REGION_RATING_CLIENT

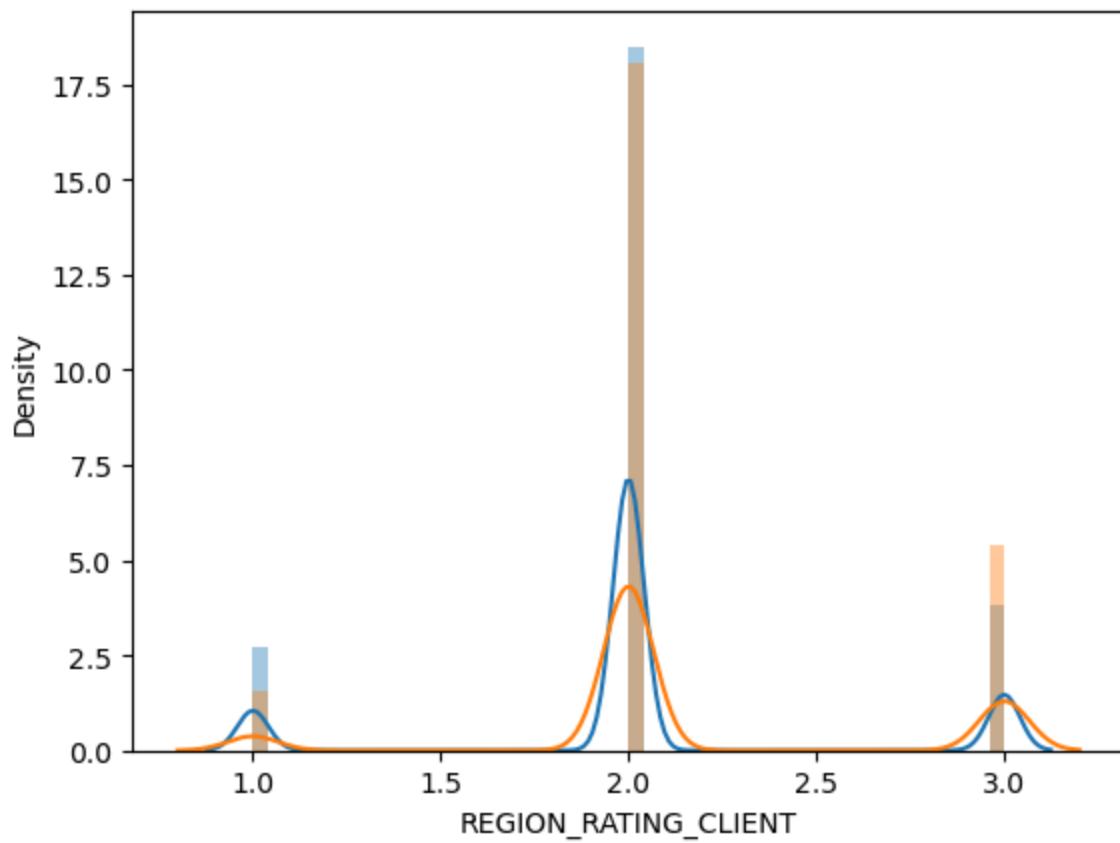
```
Out[86]: (array([ 30645.,      0.,      0.,      0., 209077.,      0.,      0.,
   42964.]),
 array([1. , 1.25, 1.5 , 1.75, 2. , 2.25, 2.5 , 2.75, 3. ]),
 <BarContainer object of 8 artists>)

Out[86]: (array([ 1552.,      0.,      0.,      0., 17907.,      0.,      0.,   5366.]),
 array([1. , 1.25, 1.5 , 1.75, 2. , 2.25, 2.5 , 2.75, 3. ]),
 <BarContainer object of 8 artists>)
```



```
Out[86]: <Axes: xlabel='REGION_RATING_CLIENT', ylabel='Density'>
```

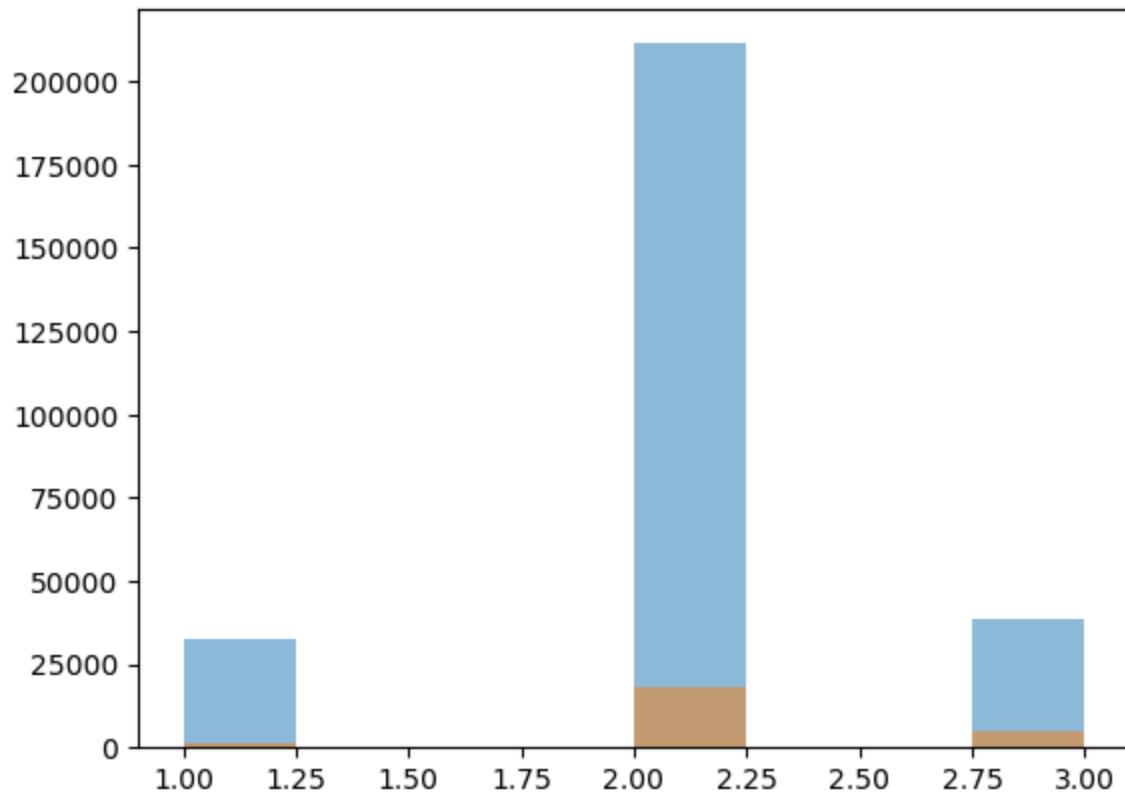
```
Out[86]: <Axes: xlabel='REGION_RATING_CLIENT', ylabel='Density'>
```



Plot of REGION_RATING_CLIENT_W_CITY

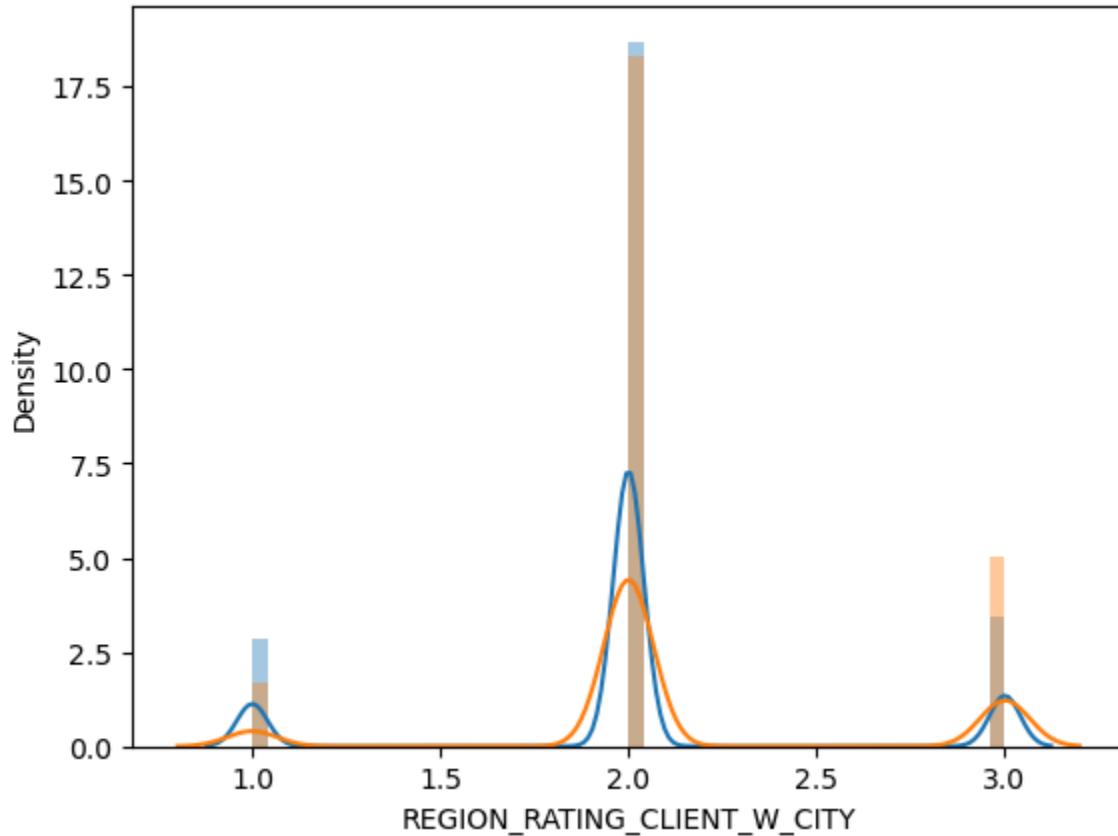
```
Out[86]: (array([ 32513.,      0.,      0.,      0., 211314.,      0.,      0.,
   38859.]),
 array([1. , 1.25, 1.5 , 1.75, 2. , 2.25, 2.5 , 2.75, 3. ]),
 <BarContainer object of 8 artists>)

Out[86]: (array([ 1654.,      0.,      0.,      0., 18170.,      0.,      0., 5001.]),
 array([1. , 1.25, 1.5 , 1.75, 2. , 2.25, 2.5 , 2.75, 3. ]),
 <BarContainer object of 8 artists>)
```



```
Out[86]: <Axes: xlabel='REGION_RATING_CLIENT_W_CITY', ylabel='Density'>
```

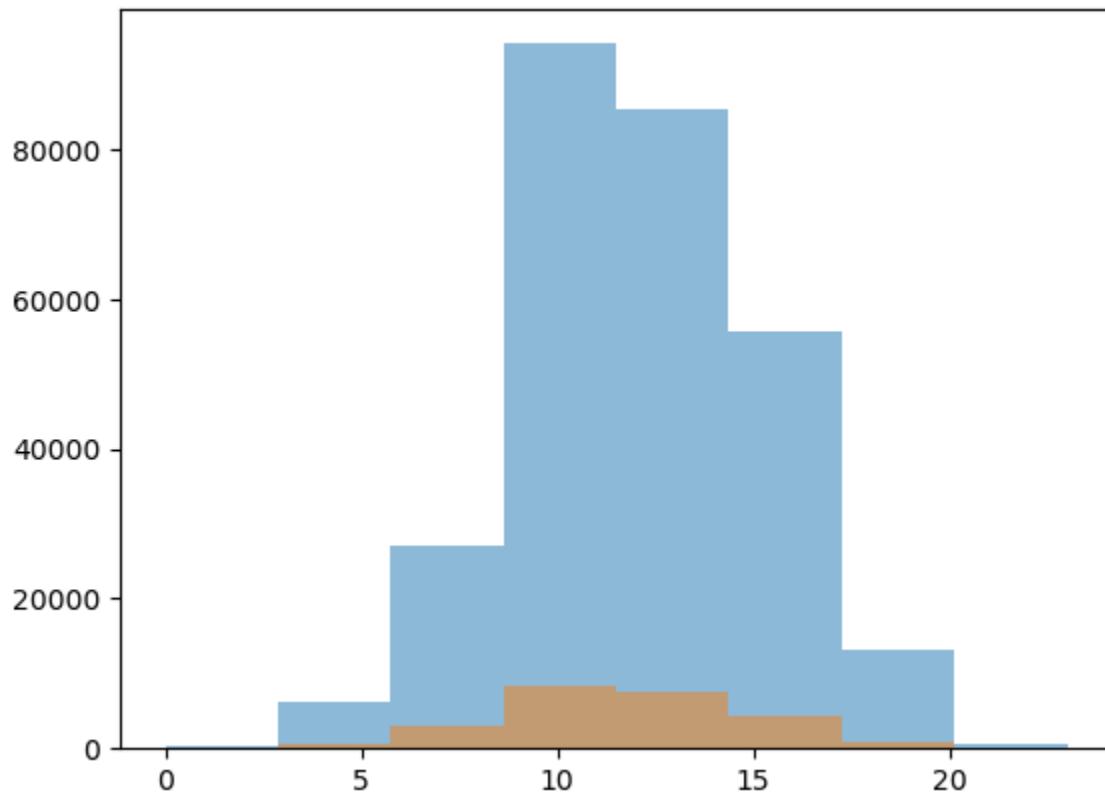
```
Out[86]: <Axes: xlabel='REGION_RATING_CLIENT_W_CITY', ylabel='Density'>
```



Plot of HOUR_APPR_PROCESS_START

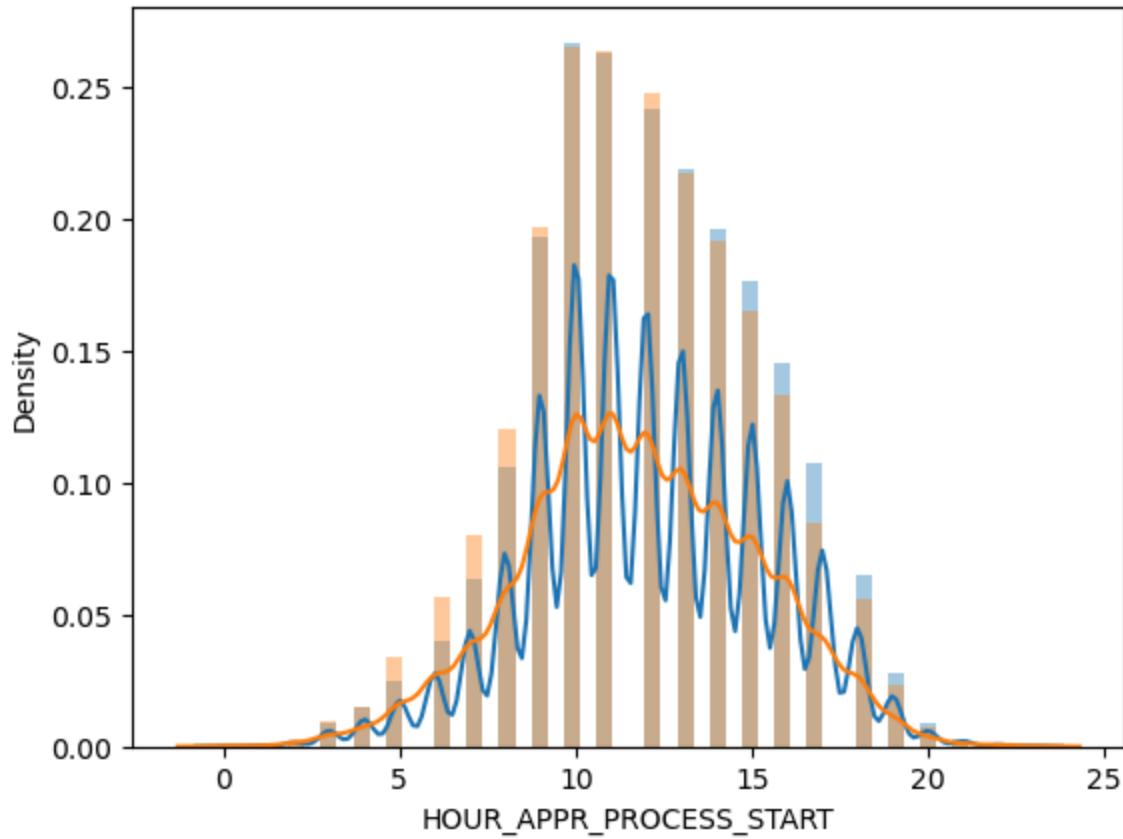
```
Out[86]: (array([ 388., 6293., 27165., 94056., 85373., 55750., 13110., 551.]),  
          array([ 0. , 2.875, 5.75 , 8.625, 11.5 , 14.375, 17.25 , 20.125,  
                 23. ]),  
          <BarContainer object of 8 artists>)
```

```
Out[86]: (array([ 43., 665., 2935., 8279., 7501., 4374., 983., 45.]),  
          array([ 0. , 2.875, 5.75 , 8.625, 11.5 , 14.375, 17.25 , 20.125,  
                 23. ]),  
          <BarContainer object of 8 artists>)
```



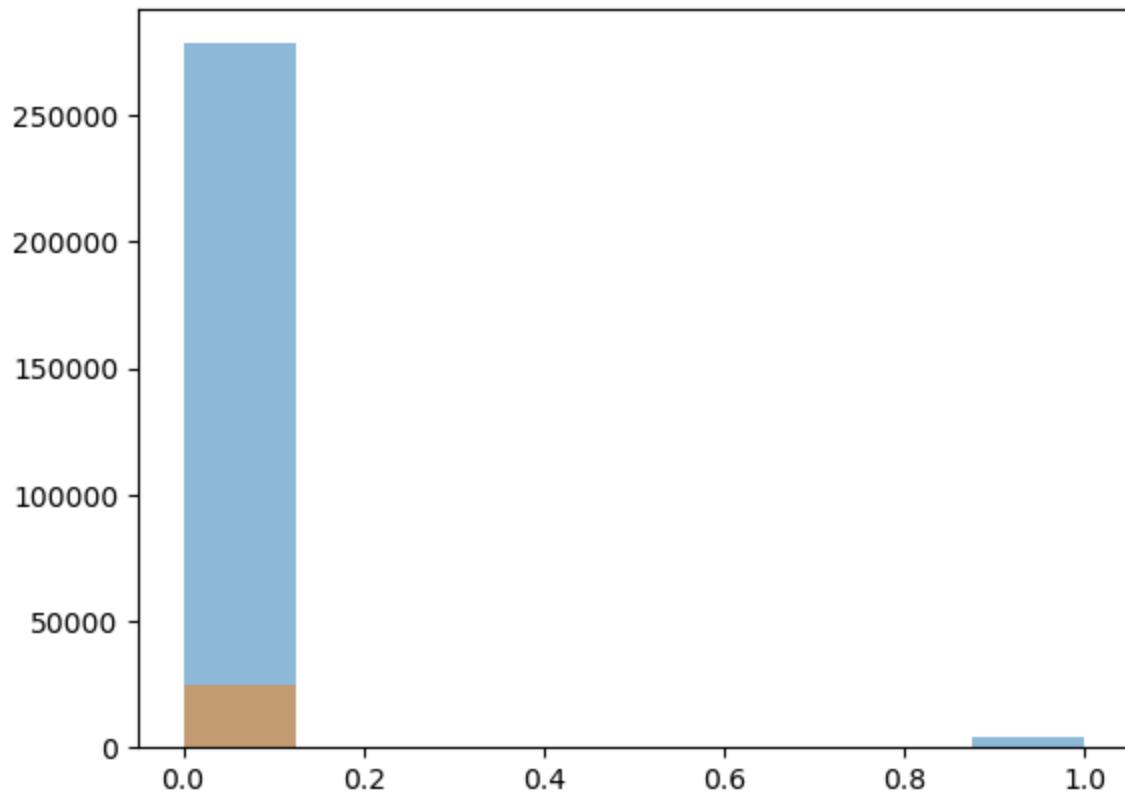
```
Out[86]: <Axes: xlabel='HOUR_APPR_PROCESS_START', ylabel='Density'>
```

```
Out[86]: <Axes: xlabel='HOUR_APPR_PROCESS_START', ylabel='Density'>
```



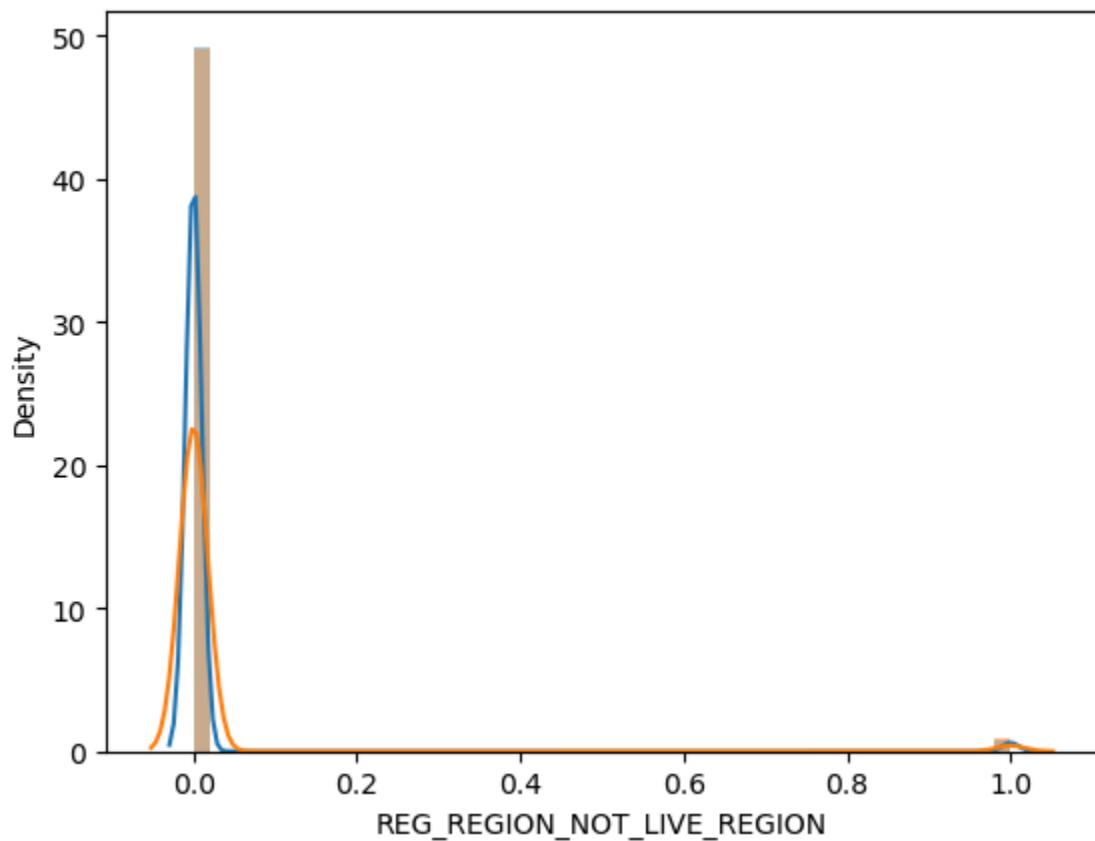
Plot of REG_REGION_NOT_LIVE_REGION

```
Out[86]: (array([278462.,      0.,      0.,      0.,      0.,      0.,      0.,      0.,      4224.]),  
          array([0.      , 0.125, 0.25 , 0.375, 0.5   , 0.625, 0.75 , 0.875, 1.     ]),  
          <BarContainer object of 8 artists>)  
  
Out[86]: (array([24392.,      0.,      0.,      0.,      0.,      0.,      0.,      0.,      433.]),  
          array([0.      , 0.125, 0.25 , 0.375, 0.5   , 0.625, 0.75 , 0.875, 1.     ]),  
          <BarContainer object of 8 artists>)
```



```
Out[86]: <Axes: xlabel='REG_REGION_NOT_LIVE_REGION', ylabel='Density'>
```

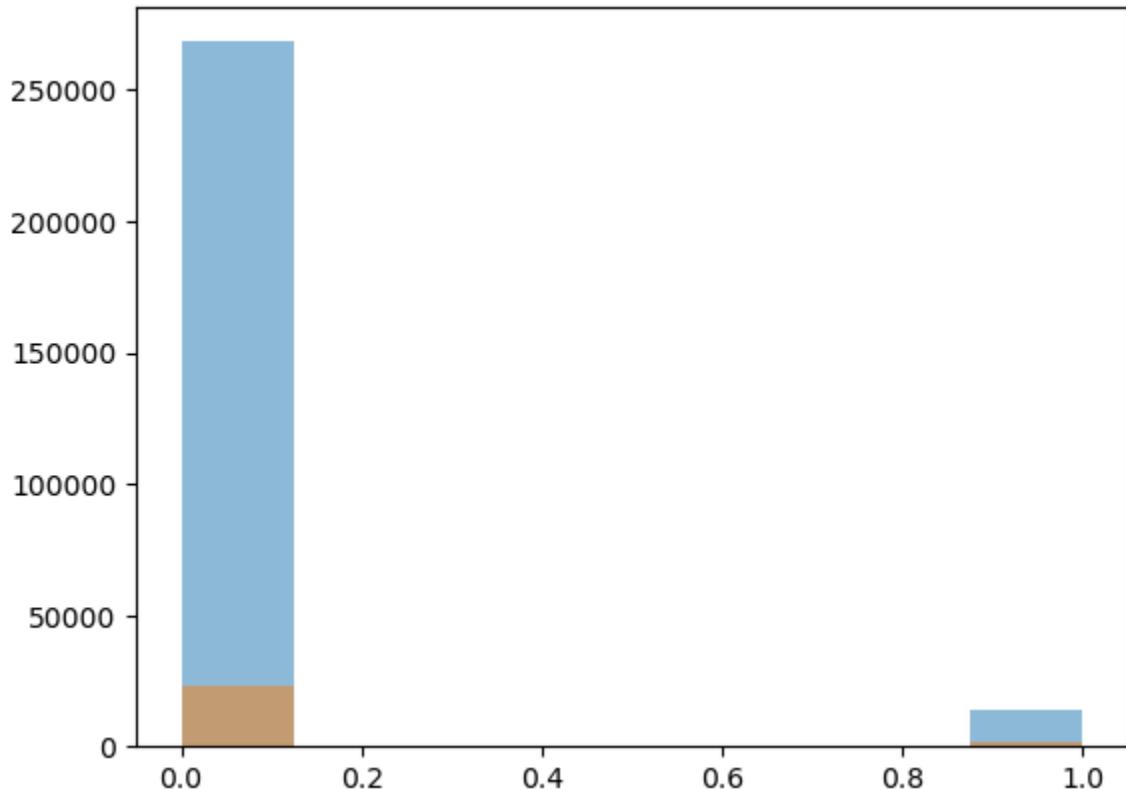
```
Out[86]: <Axes: xlabel='REG_REGION_NOT_LIVE_REGION', ylabel='Density'>
```



Plot of REG_REGION_NOT_WORK_REGION

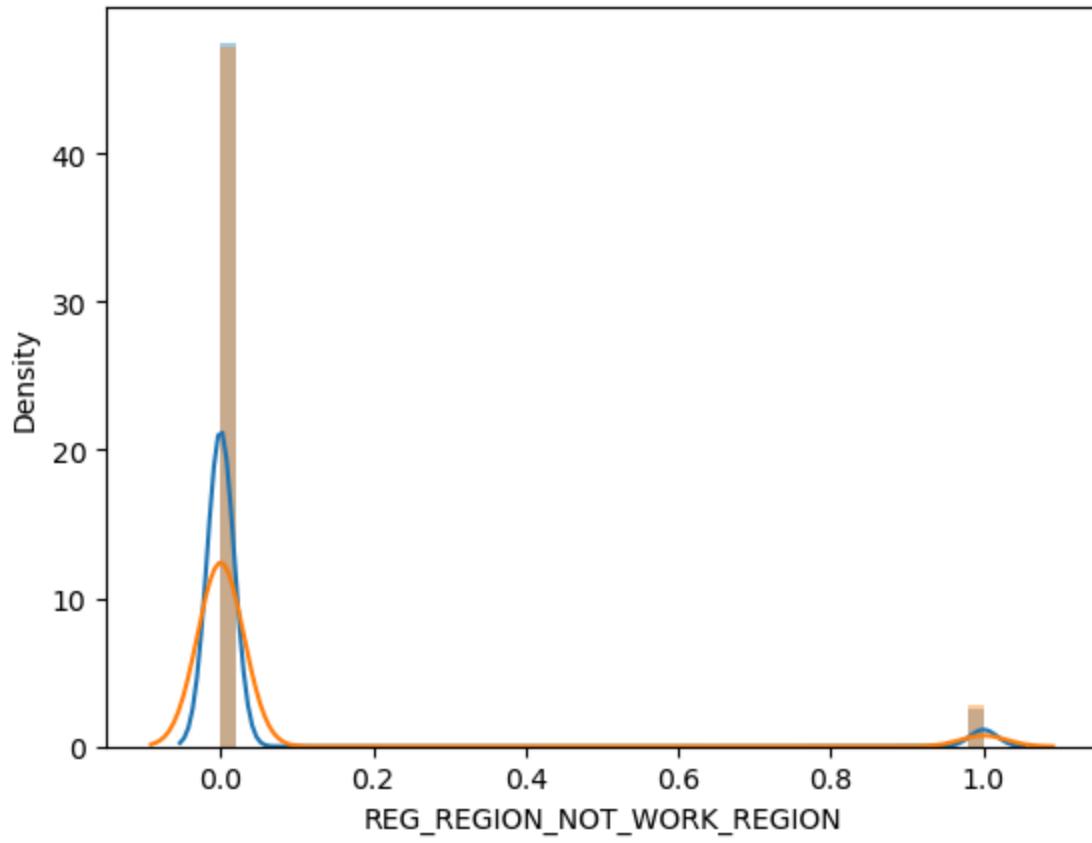
```
Out[86]: (array([268462.,      0.,      0.,      0.,      0.,      0.,      0.,
   14224.]),
 array([0.     , 0.125, 0.25 , 0.375, 0.5   , 0.625, 0.75 , 0.875, 1.     ]),
 <BarContainer object of 8 artists>)

Out[86]: (array([23437.,      0.,      0.,      0.,      0.,      0.,      0.,      1388.]),
 array([0.     , 0.125, 0.25 , 0.375, 0.5   , 0.625, 0.75 , 0.875, 1.     ]),
 <BarContainer object of 8 artists>)
```



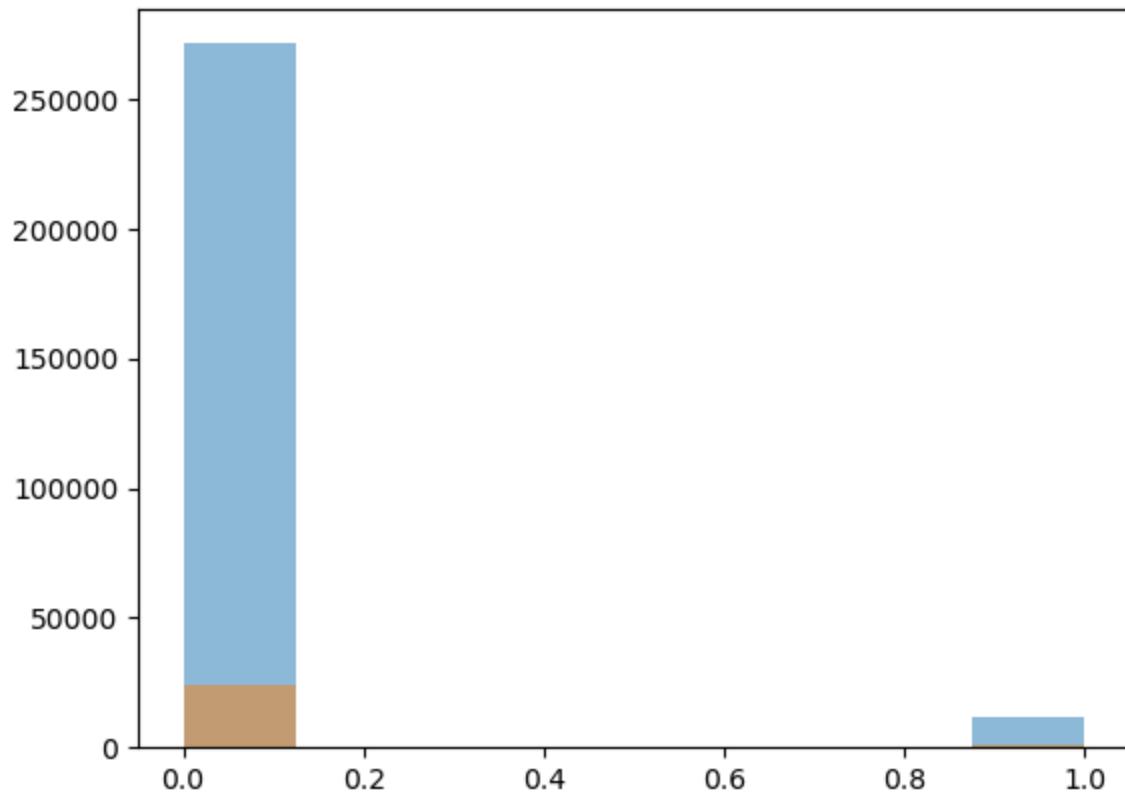
```
Out[86]: <Axes: xlabel='REG_REGION_NOT_WORK_REGION', ylabel='Density'>
```

```
Out[86]: <Axes: xlabel='REG_REGION_NOT_WORK_REGION', ylabel='Density'>
```



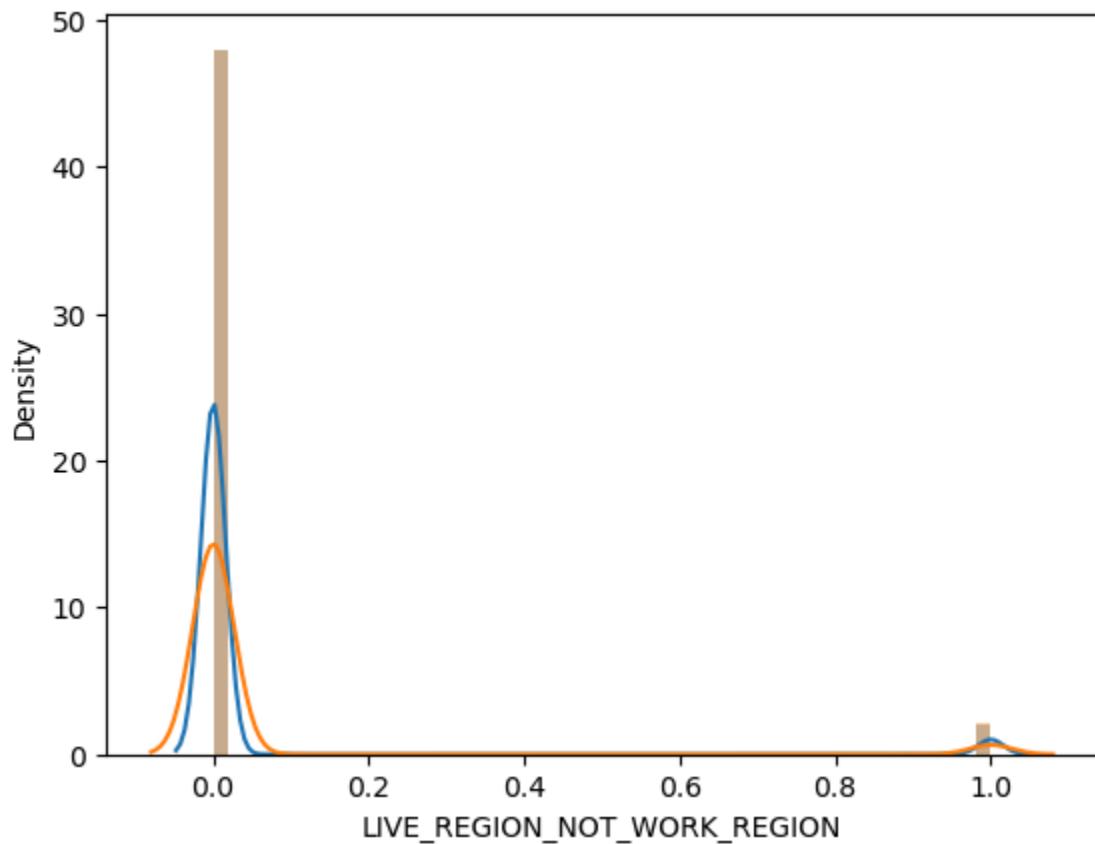
Plot of `LIVE_REGION_NOT_WORK_REGION`

```
Out[86]: (array([271239.,      0.,      0.,      0.,      0.,      0.,      0.,  
    11447.]),  
 array([0.      , 0.125, 0.25 , 0.375, 0.5   , 0.625, 0.75 , 0.875, 1.     ]),  
 <BarContainer object of 8 artists>)  
Out[86]: (array([23769.,      0.,      0.,      0.,      0.,      0.,      0., 1056.]),  
 array([0.      , 0.125, 0.25 , 0.375, 0.5   , 0.625, 0.75 , 0.875, 1.     ]),  
 <BarContainer object of 8 artists>)
```



```
Out[86]: <Axes: xlabel='LIVE_REGION_NOT_WORK_REGION', ylabel='Density'>
```

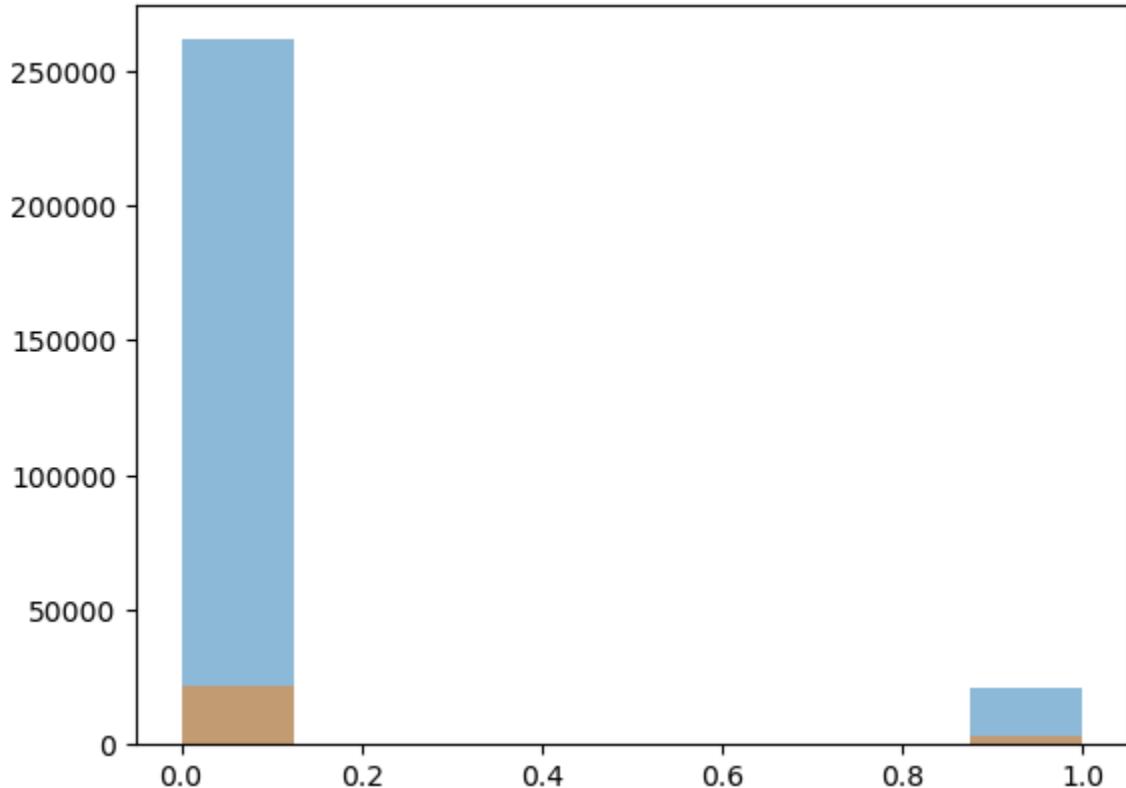
```
Out[86]: <Axes: xlabel='LIVE_REGION_NOT_WORK_REGION', ylabel='Density'>
```



Plot of REG_CITY_NOT_LIVE_CITY

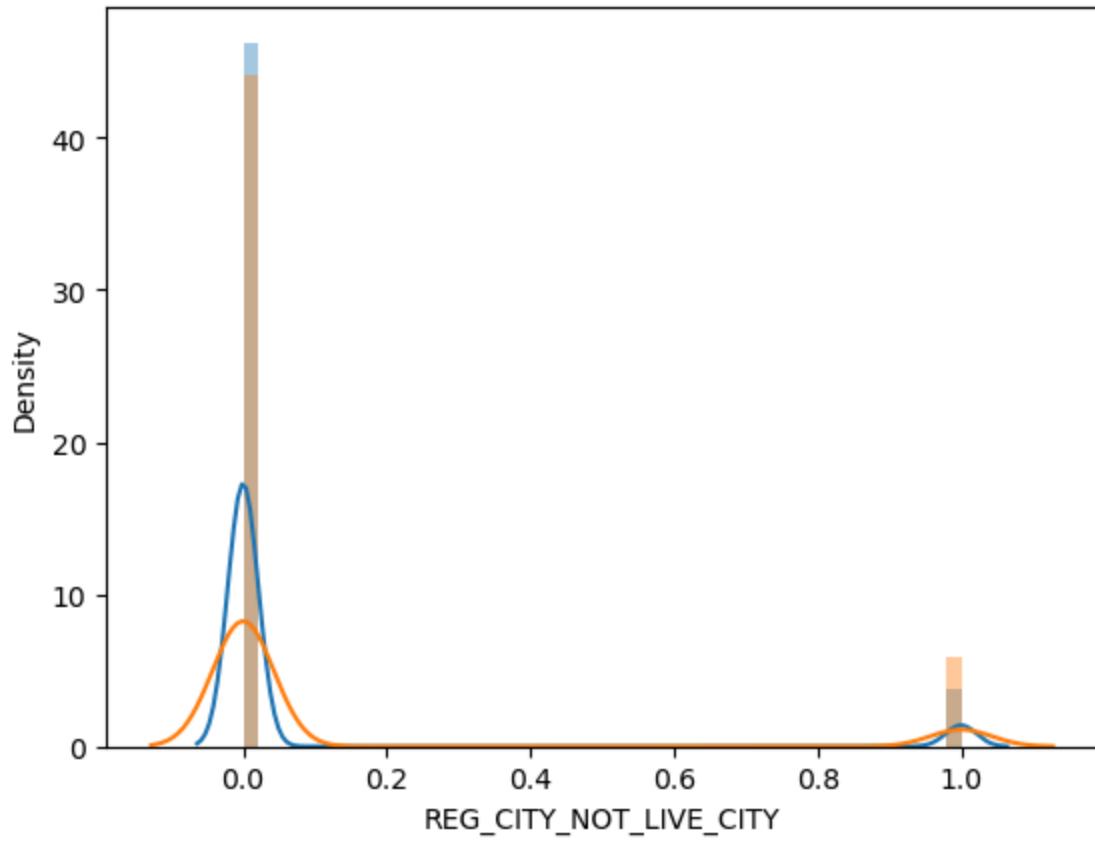
```
Out[86]: (array([261586.,      0.,      0.,      0.,      0.,      0.,      0.,
   21100.]),
 array([0.    , 0.125, 0.25 , 0.375, 0.5   , 0.625, 0.75 , 0.875, 1.    ]),
 <BarContainer object of 8 artists>)

Out[86]: (array([21886.,      0.,      0.,      0.,      0.,      0.,      0.,      2939.]),
 array([0.    , 0.125, 0.25 , 0.375, 0.5   , 0.625, 0.75 , 0.875, 1.    ]),
 <BarContainer object of 8 artists>)
```



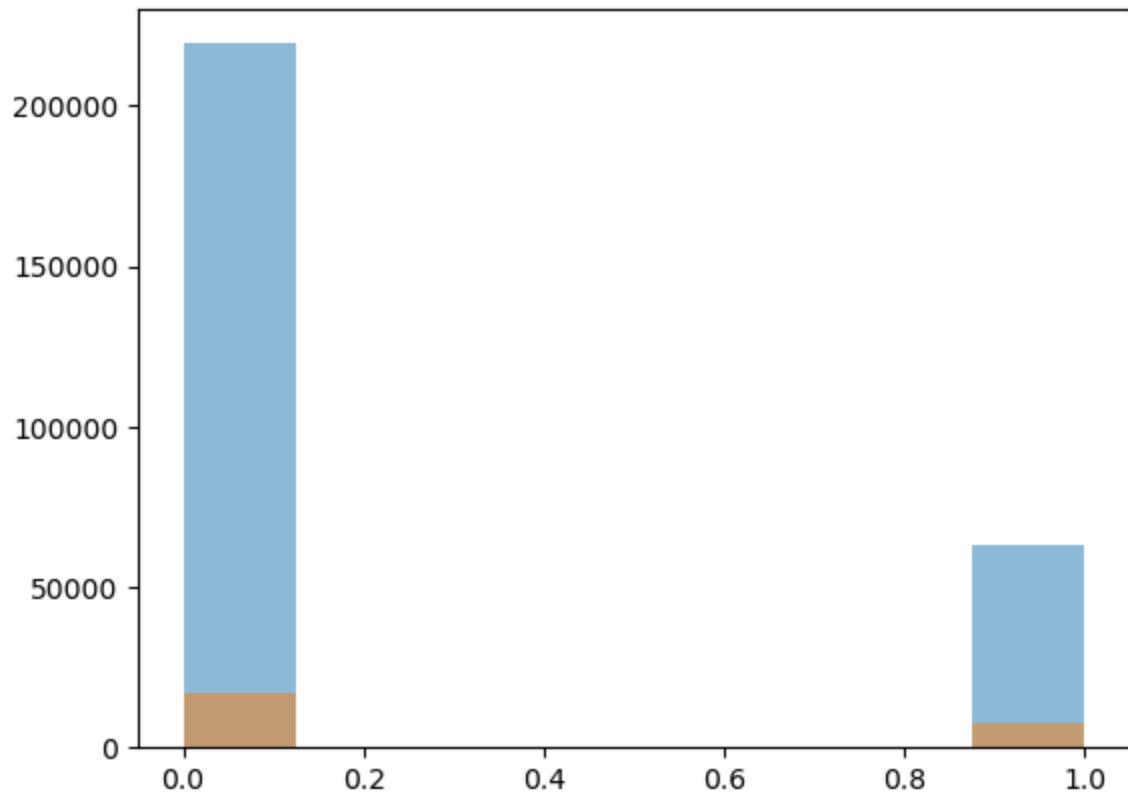
```
Out[86]: <Axes: xlabel='REG_CITY_NOT_LIVE_CITY', ylabel='Density'>
```

```
Out[86]: <Axes: xlabel='REG_CITY_NOT_LIVE_CITY', ylabel='Density'>
```



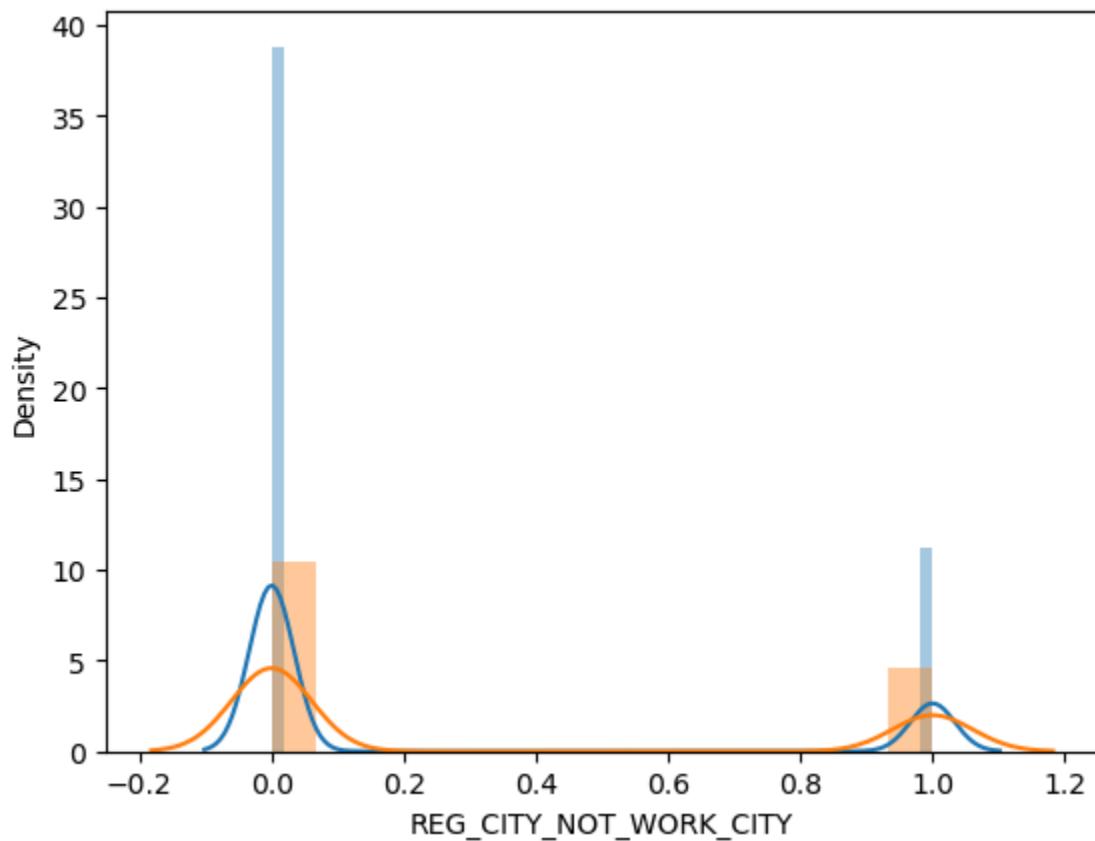
Plot of `REG_CITY_NOT_WORK_CITY`

```
Out[86]: (array([219339.,      0.,      0.,      0.,      0.,      0.,      0.,  
       63347.]),  
 array([0.      , 0.125, 0.25 , 0.375, 0.5   , 0.625, 0.75 , 0.875, 1.     ]),  
 <BarContainer object of 8 artists>)  
Out[86]: (array([17305.,      0.,      0.,      0.,      0.,      0.,      0.,    7520.]),  
 array([0.      , 0.125, 0.25 , 0.375, 0.5   , 0.625, 0.75 , 0.875, 1.     ]),  
 <BarContainer object of 8 artists>)
```



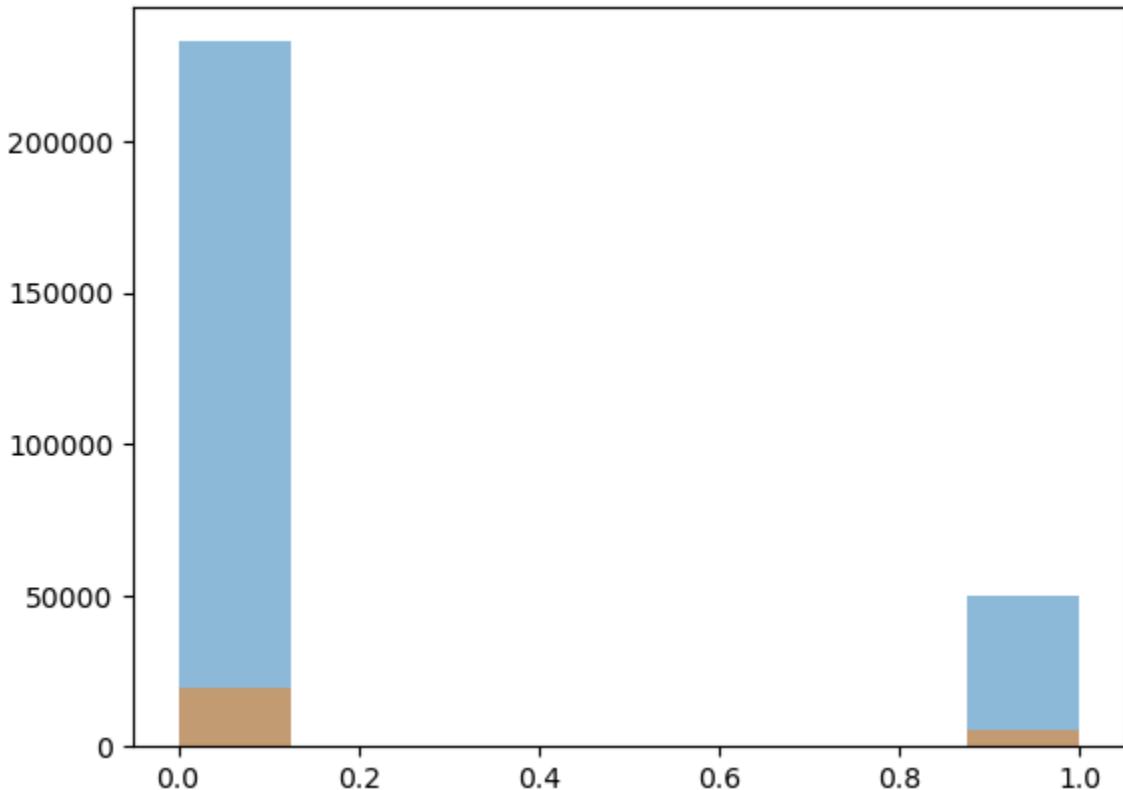
```
Out[86]: <Axes: xlabel='REG_CITY_NOT_WORK_CITY', ylabel='Density'>
```

```
Out[86]: <Axes: xlabel='REG_CITY_NOT_WORK_CITY', ylabel='Density'>
```



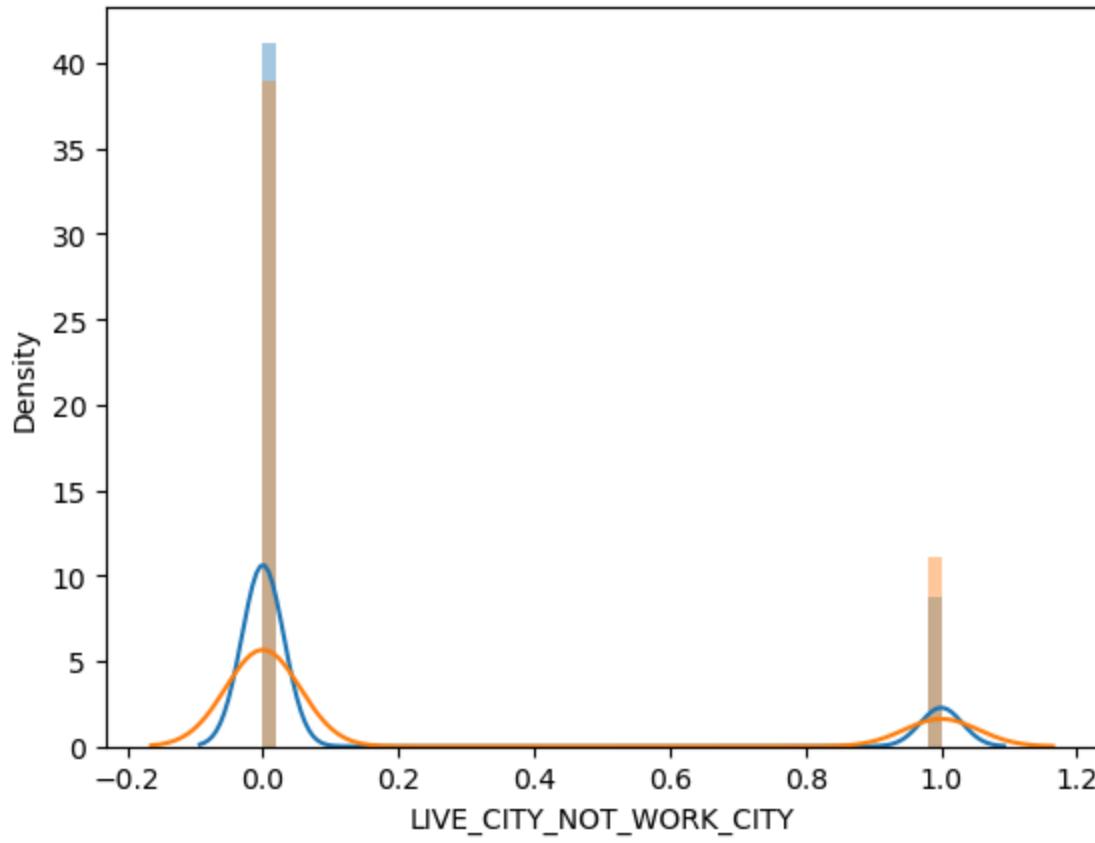
Plot of LIVE_CITY_NOT_WORK_CITY

```
Out[86]: (array([232974.,      0.,      0.,      0.,      0.,      0.,
   49712.]),  
 array([0.    , 0.125, 0.25 , 0.375, 0.5  , 0.625, 0.75 , 0.875, 1.    ]),  
 <BarContainer object of 8 artists>)  
Out[86]: (array([19322.,      0.,      0.,      0.,      0.,      0.,      0.,  5503.]),  
 array([0.    , 0.125, 0.25 , 0.375, 0.5  , 0.625, 0.75 , 0.875, 1.    ]),  
 <BarContainer object of 8 artists>)
```



```
Out[86]: <Axes: xlabel='LIVE_CITY_NOT_WORK_CITY', ylabel='Density'>
```

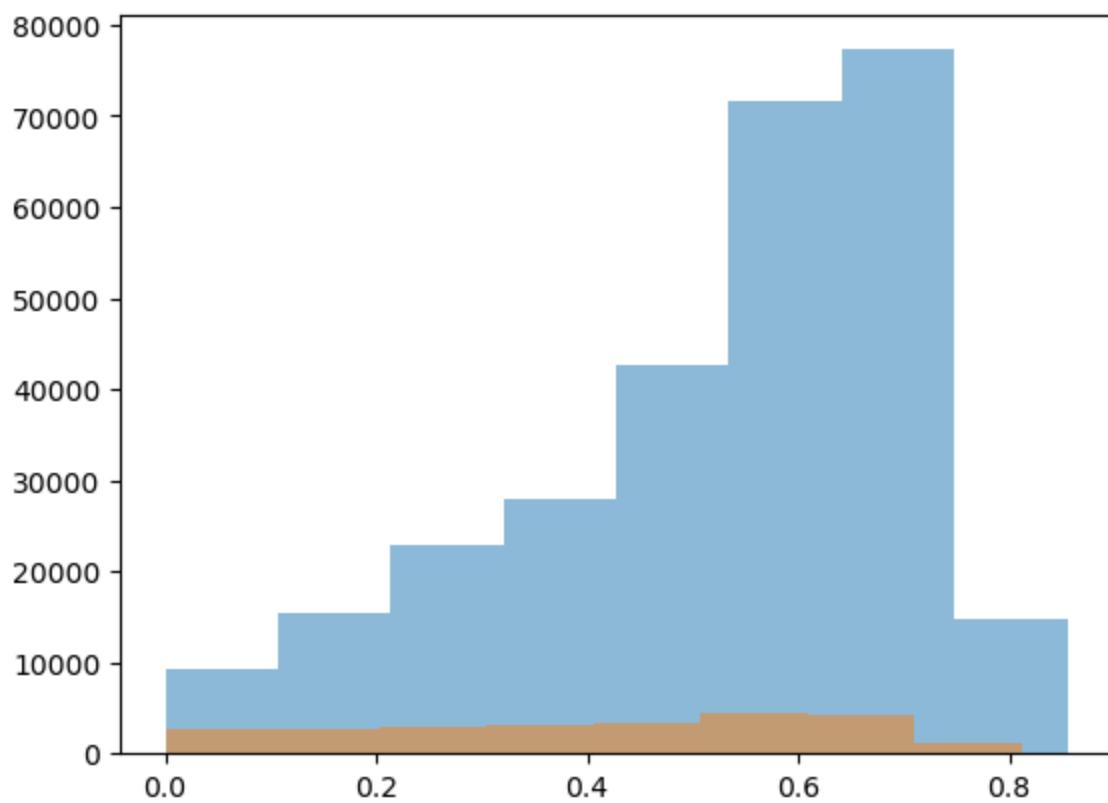
```
Out[86]: <Axes: xlabel='LIVE_CITY_NOT_WORK_CITY', ylabel='Density'>
```



Plot of EXT_SOURCE_2

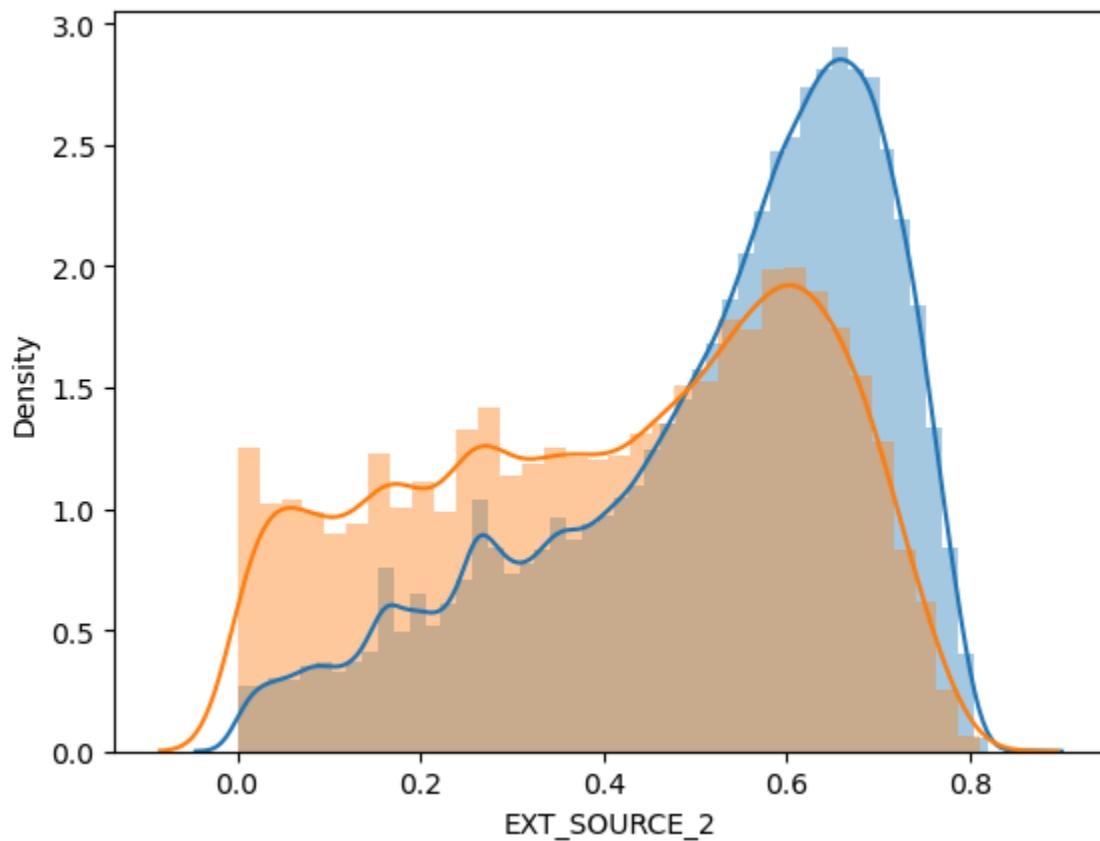
```
Out[86]: (array([ 9350., 15435., 22909., 28051., 42747., 71565., 77285., 14736.]),
          array([8.17361652e-08, 1.06875030e-01, 2.13749978e-01, 3.20624926e-01,
                 4.27499874e-01, 5.34374822e-01, 6.41249770e-01, 7.48124718e-01,
                 8.54999666e-01]),
          <BarContainer object of 8 artists>)

Out[86]: (array([2674., 2668., 2968., 3052., 3429., 4506., 4258., 1218.]),
          array([5.00210876e-06, 1.01488094e-01, 2.02971186e-01, 3.04454278e-01,
                 4.05937370e-01, 5.07420462e-01, 6.08903554e-01, 7.10386646e-01,
                 8.11869738e-01]),
          <BarContainer object of 8 artists>)
```



```
Out[86]: <Axes: xlabel='EXT_SOURCE_2', ylabel='Density'>
```

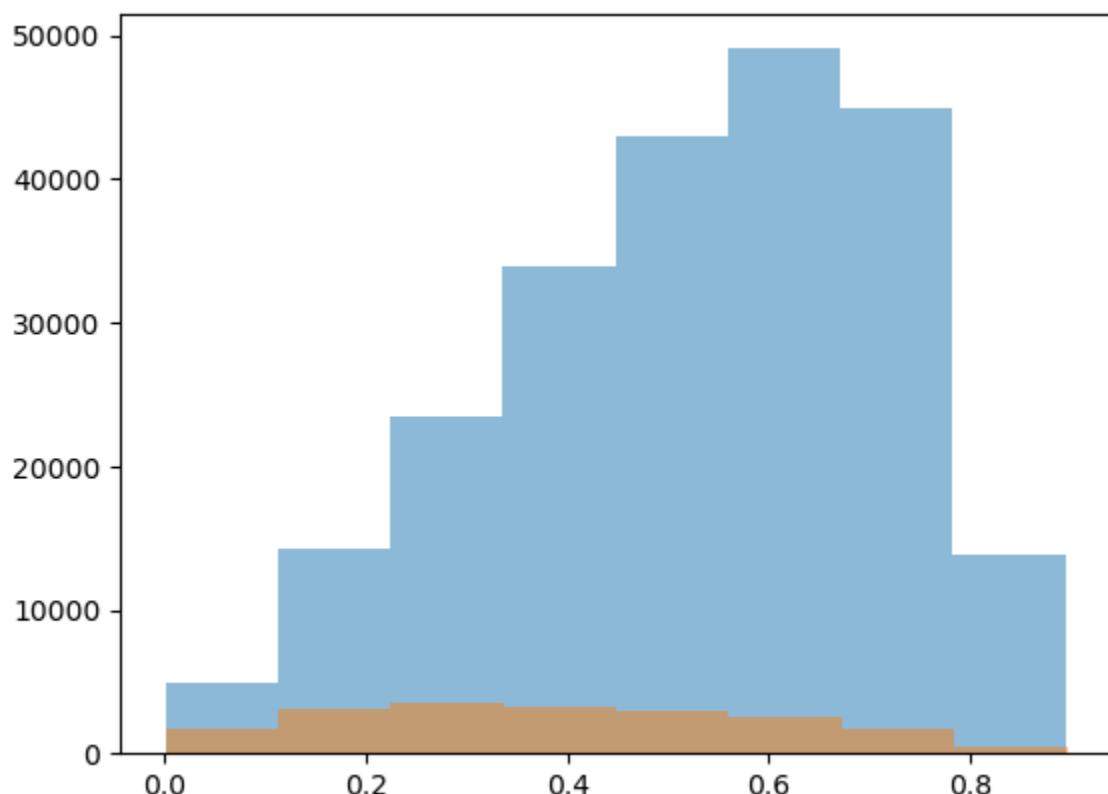
```
Out[86]: <Axes: xlabel='EXT_SOURCE_2', ylabel='Density'>
```



```
-----  
Plot of EXT_SOURCE_3
```

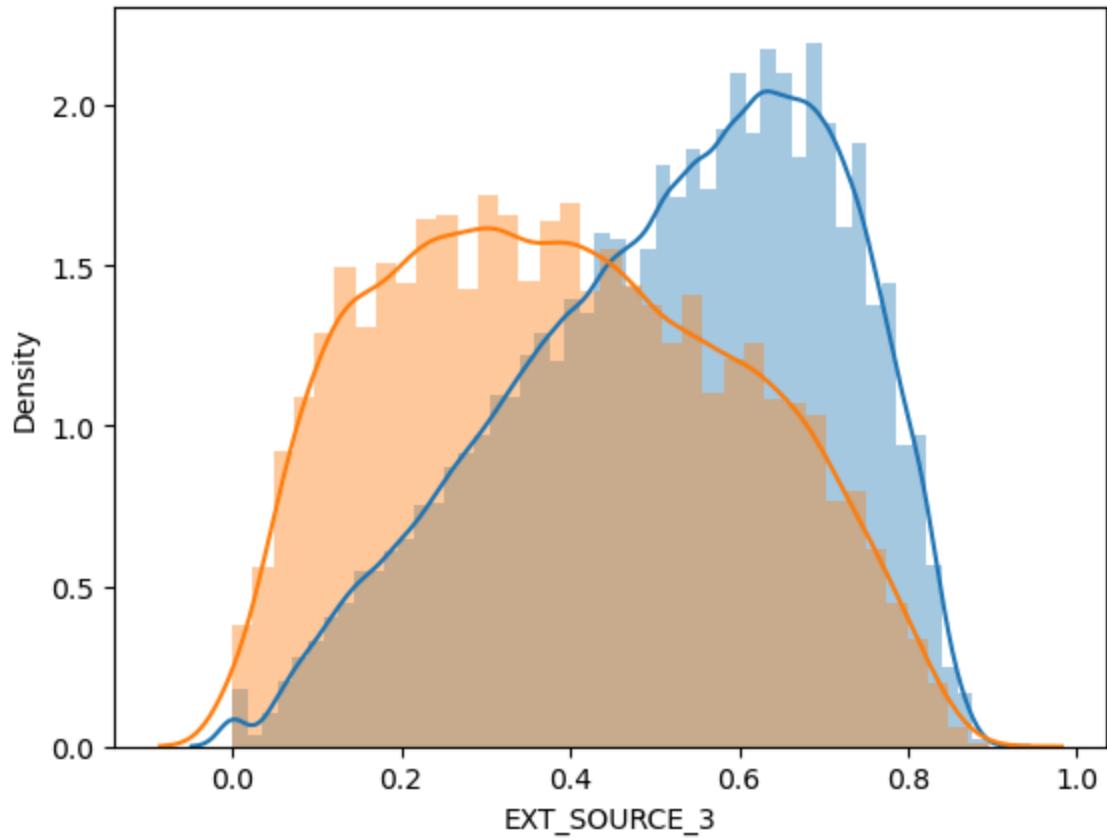
```
Out[86]: (array([ 4938., 14288., 23476., 33861., 42934., 49059., 44982., 13860.]),  
          array([5.27265239e-04, 1.12208366e-01, 2.23889468e-01, 3.35570569e-01,  
                 4.47251670e-01, 5.58932771e-01, 6.70613872e-01, 7.82294973e-01,  
                 8.93976075e-01]),  
         <BarContainer object of 8 artists>)
```

```
Out[86]: (array([1707., 3130., 3473., 3292., 2959., 2490., 1687., 410.]),  
          array([5.27265239e-04, 1.12462551e-01, 2.24397836e-01, 3.36333122e-01,  
                 4.48268407e-01, 5.60203693e-01, 6.72138978e-01, 7.84074264e-01,  
                 8.96009549e-01]),  
         <BarContainer object of 8 artists>)
```



```
Out[86]: <Axes: xlabel='EXT_SOURCE_3', ylabel='Density'>
```

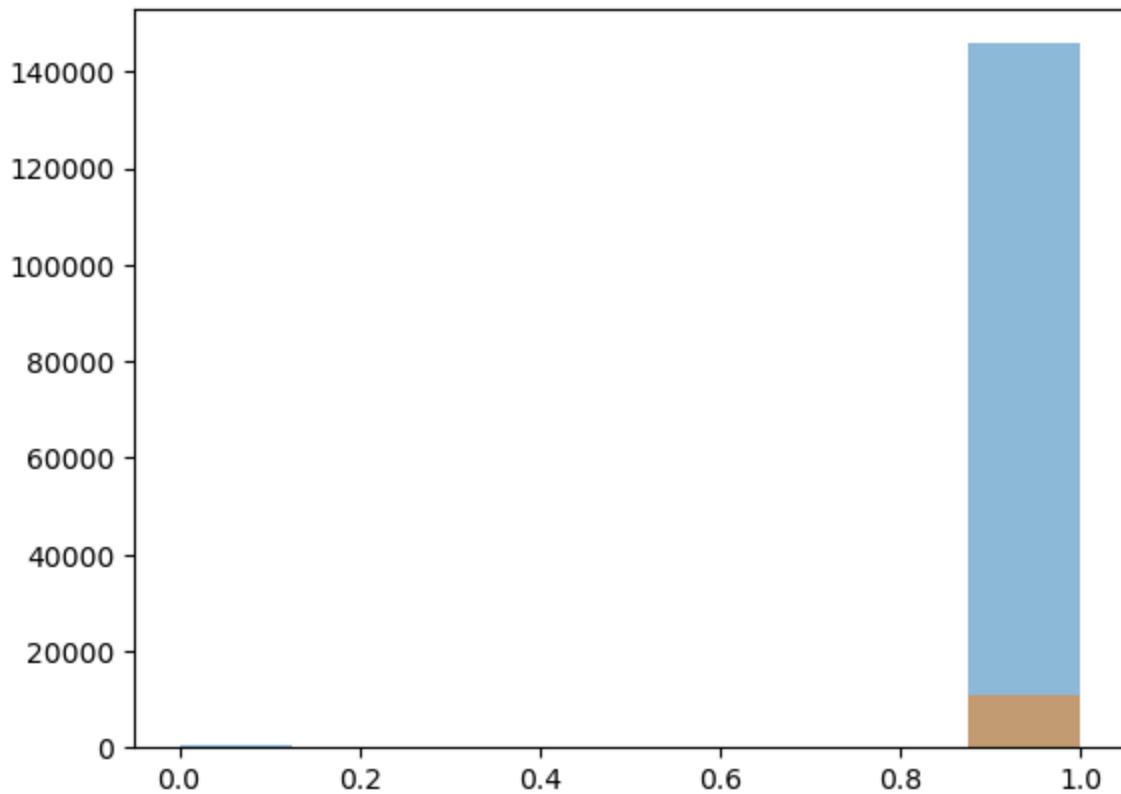
```
Out[86]: <Axes: xlabel='EXT_SOURCE_3', ylabel='Density'>
```



Plot of YEARS_BEGINEXPLUATATION_AVG

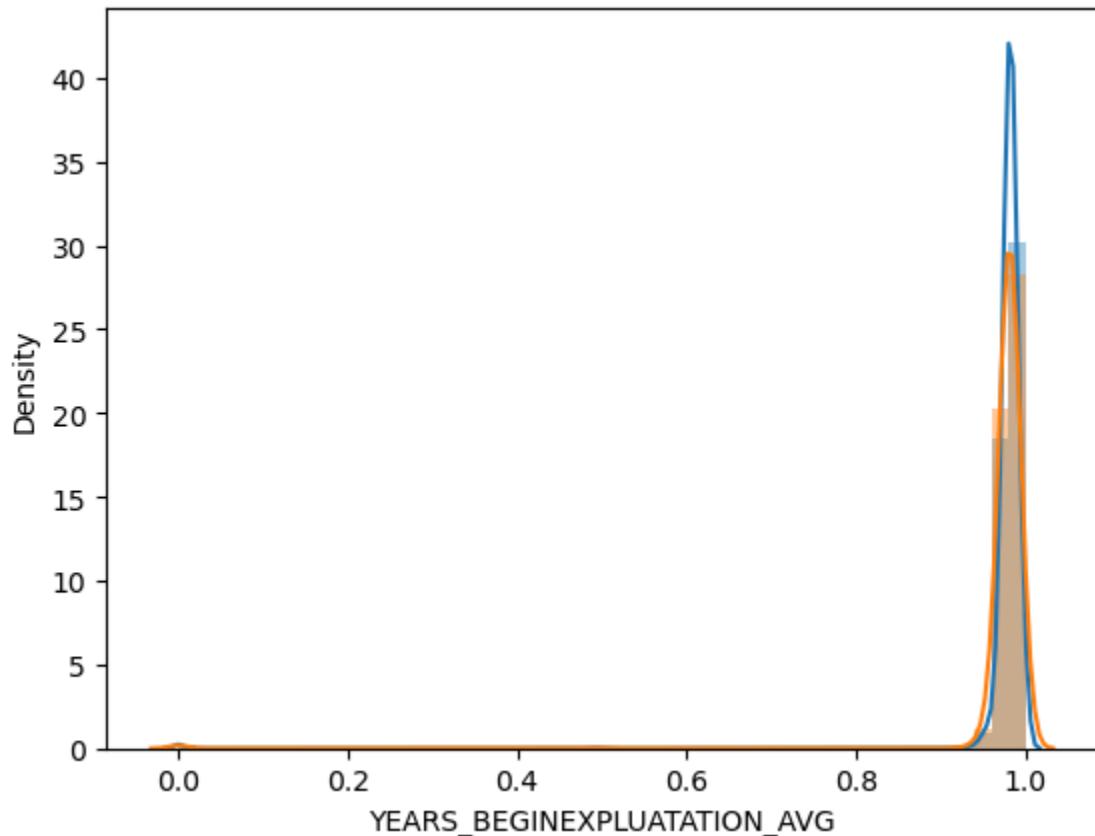
```
Out[86]: (array([4.68000e+02, 1.00000e+01, 2.10000e+01, 8.60000e+01, 7.00000e+00,
       3.60000e+01, 1.40000e+01, 1.45845e+05]),
       array([0.    , 0.125, 0.25 , 0.375, 0.5   , 0.625, 0.75 , 0.875, 1.    ]),
       <BarContainer object of 8 artists>)

Out[86]: (array([5.1000e+01, 0.0000e+00, 3.0000e+00, 7.0000e+00, 0.0000e+00,
       1.0000e+00, 0.0000e+00, 1.0955e+04]),
       array([0.    , 0.125, 0.25 , 0.375, 0.5   , 0.625, 0.75 , 0.875, 1.    ]),
       <BarContainer object of 8 artists>)
```



```
Out[86]: <Axes: xlabel='YEARS_BEGINEXPLUATATION_AVG', ylabel='Density'>
```

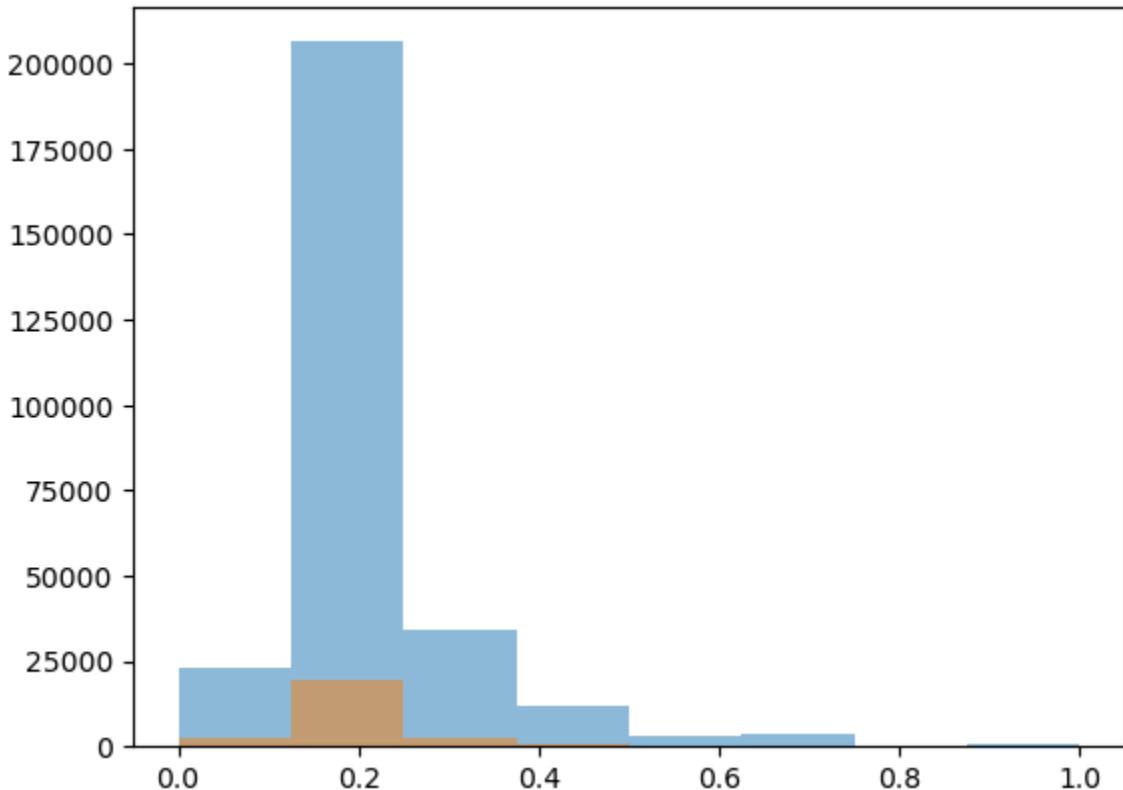
```
Out[86]: <Axes: xlabel='YEARS_BEGINEXPLUATATION_AVG', ylabel='Density'>
```



Plot of FLOORSMAX_AVG

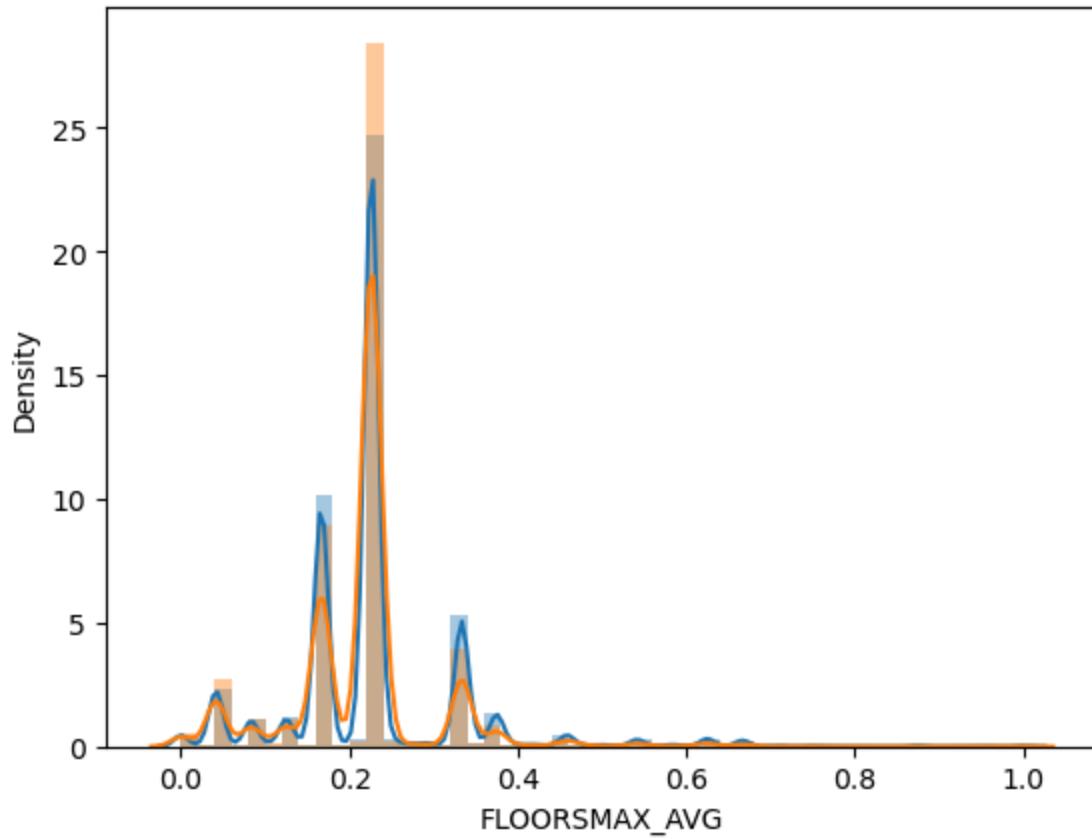
```
Out[86]: (array([ 22989., 206413., 34235., 11784., 2844., 3623., 312.,
   486.]),
 array([0. , 0.125, 0.25 , 0.375, 0.5 , 0.625, 0.75 , 0.875, 1. ]),
 <BarContainer object of 8 artists>)

Out[86]: (array([2.2640e+03, 1.9325e+04, 2.2280e+03, 6.6800e+02, 1.3900e+02,
 1.6600e+02, 1.4000e+01, 2.1000e+01]),
 array([0. , 0.125, 0.25 , 0.375, 0.5 , 0.625, 0.75 , 0.875, 1. ]),
 <BarContainer object of 8 artists>)
```



```
Out[86]: <Axes: xlabel='FLOORSMAX_AVG', ylabel='Density'>
```

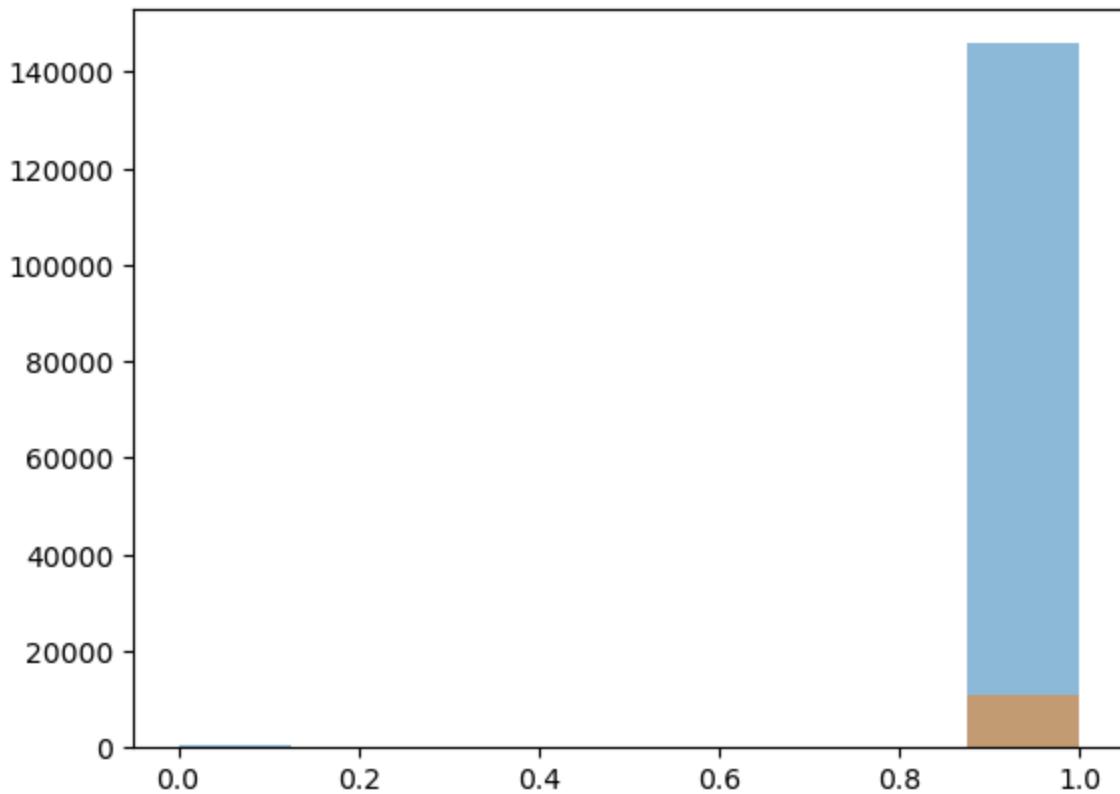
```
Out[86]: <Axes: xlabel='FLOORSMAX_AVG', ylabel='Density'>
```



Plot of YEARS_BEGINEXPLUATATION_MODE

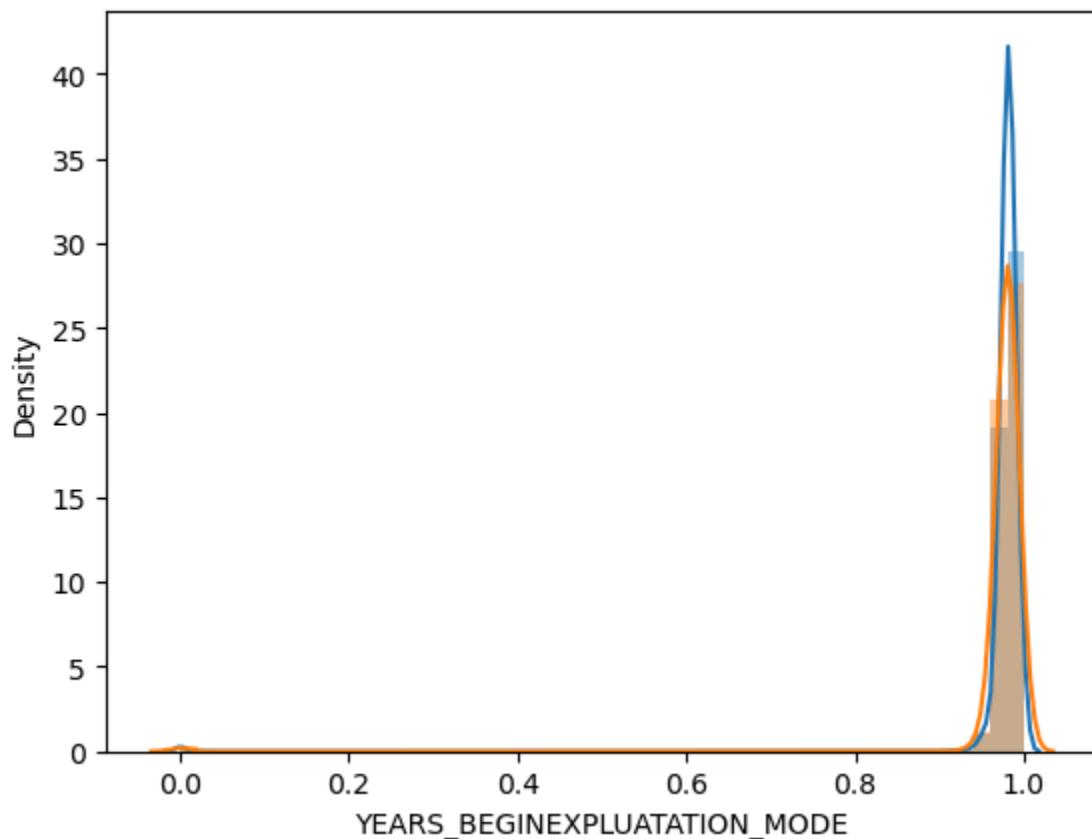
```
Out[86]: (array([6.1000e+02, 0.0000e+00, 0.0000e+00, 2.0000e+00, 5.0000e+00,
       1.0000e+00, 9.0000e+00, 1.4586e+05]),
       array([0.    , 0.125, 0.25 , 0.375, 0.5   , 0.625, 0.75 , 0.875, 1.    ]),
       <BarContainer object of 8 artists>)

Out[86]: (array([6.2000e+01, 0.0000e+00, 0.0000e+00, 0.0000e+00, 0.0000e+00,
       0.0000e+00, 1.0000e+00, 1.0954e+04]),
       array([0.    , 0.125, 0.25 , 0.375, 0.5   , 0.625, 0.75 , 0.875, 1.    ]),
       <BarContainer object of 8 artists>)
```



```
Out[86]: <Axes: xlabel='YEARS_BEGINEXPLUATATION_MODE', ylabel='Density'>
```

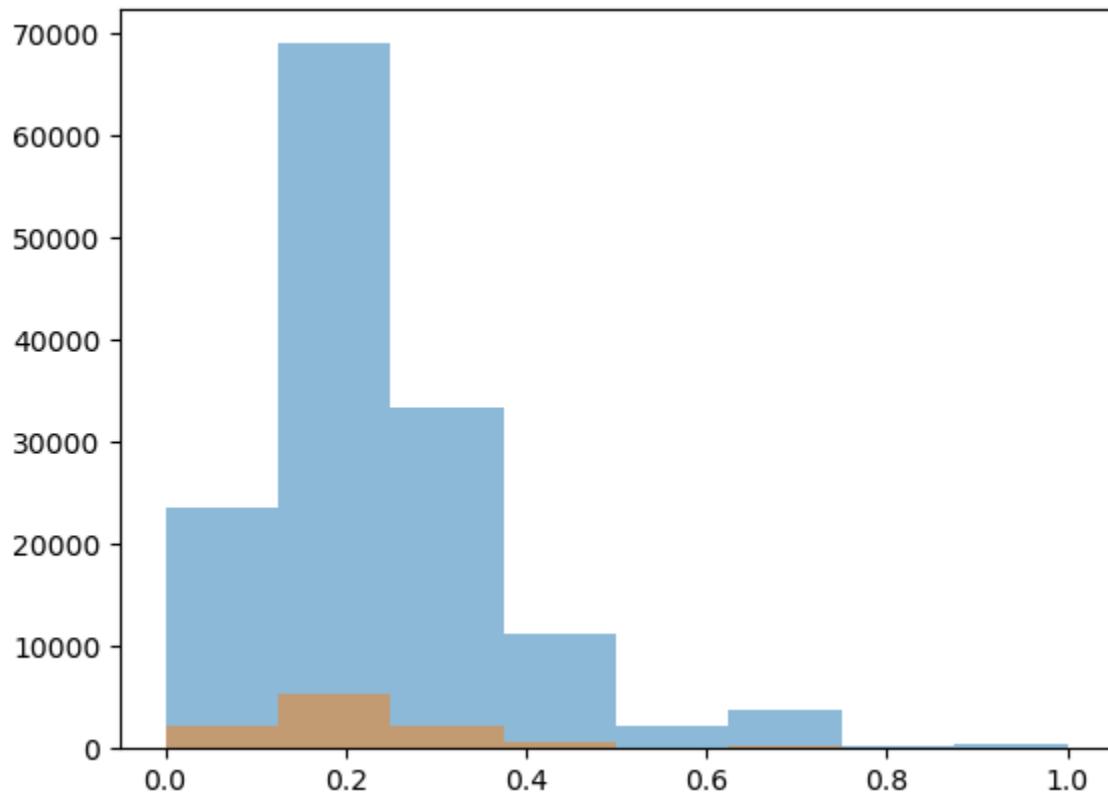
```
Out[86]: <Axes: xlabel='YEARS_BEGINEXPLUATATION_MODE', ylabel='Density'>
```



Plot of FLOORSMAX_MODE

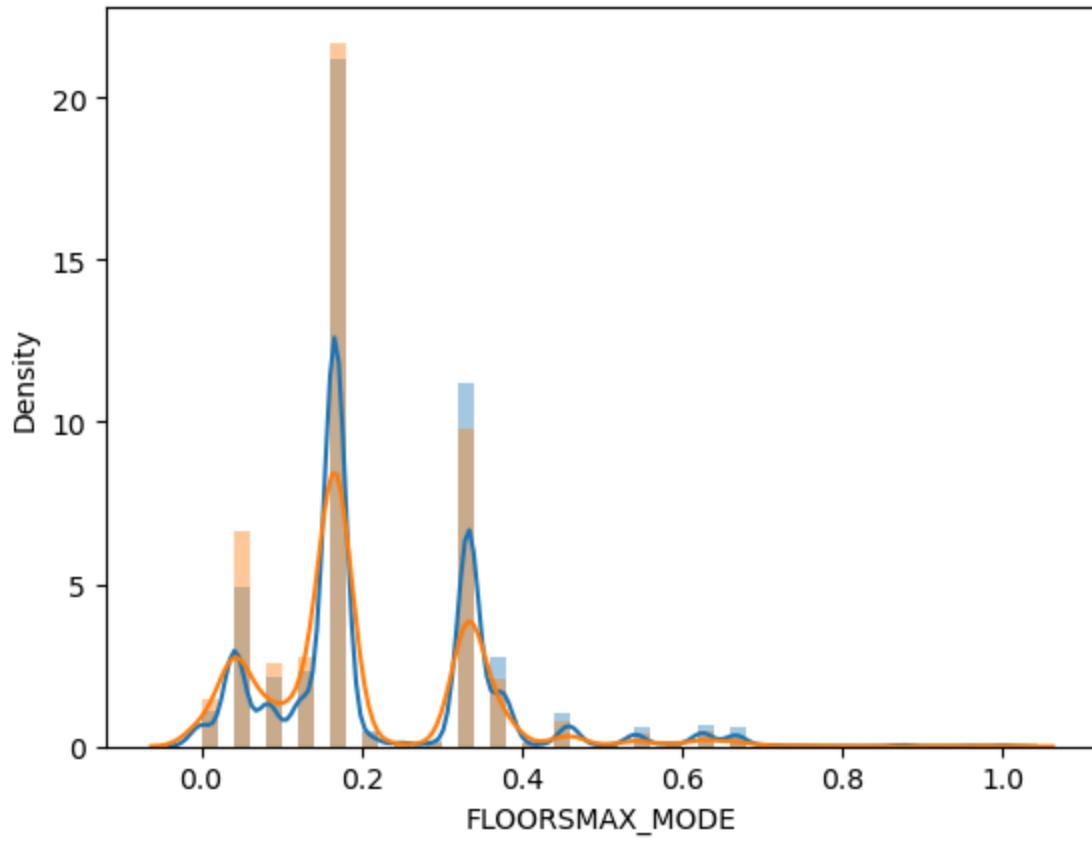
```
Out[86]: (array([23519., 68969., 33297., 11140., 2296., 3812., 193., 504.]),  
 array([0. , 0.125, 0.25 , 0.375, 0.5 , 0.625, 0.75 , 0.875, 1. ]),  
 <BarContainer object of 8 artists>)
```

```
Out[86]: (array([2297., 5353., 2159., 633., 115., 172., 11., 21.]),  
 array([0. , 0.125, 0.25 , 0.375, 0.5 , 0.625, 0.75 , 0.875, 1. ]),  
 <BarContainer object of 8 artists>)
```



```
Out[86]: <Axes: xlabel='FLOORSMAX_MODE', ylabel='Density'>
```

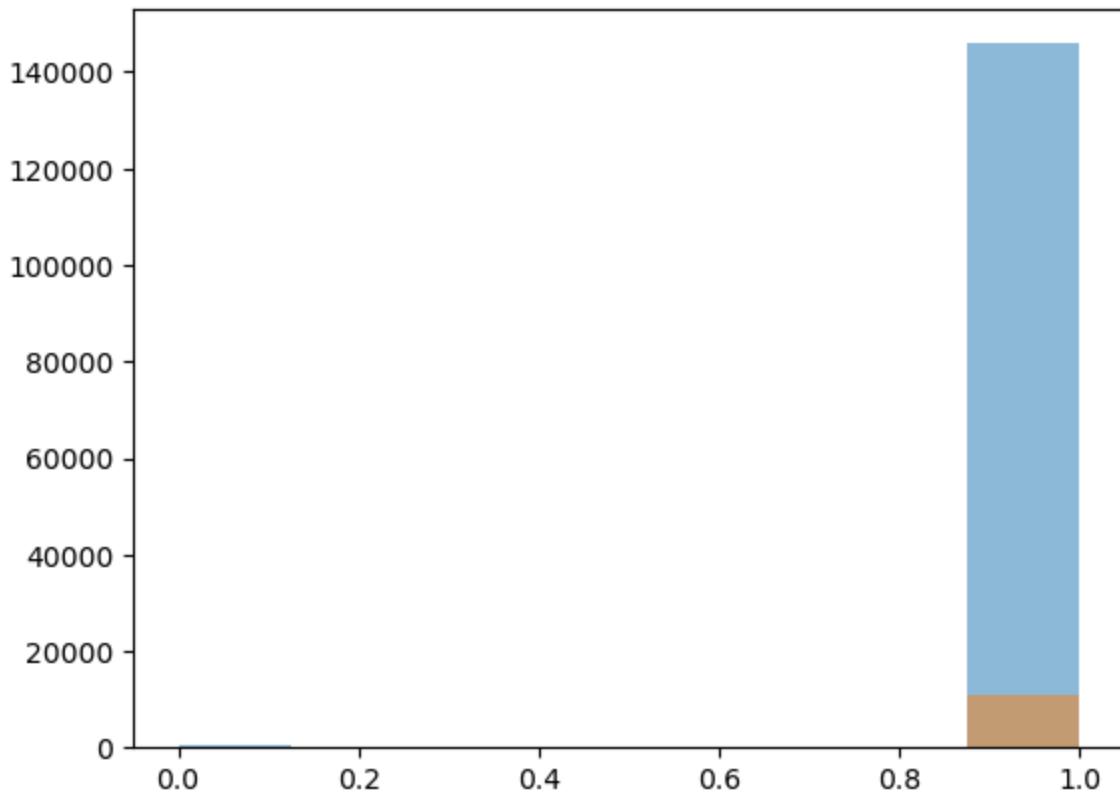
```
Out[86]: <Axes: xlabel='FLOORSMAX_MODE', ylabel='Density'>
```



Plot of `YEARS_BEGINEXPLUATATION_MEDI`

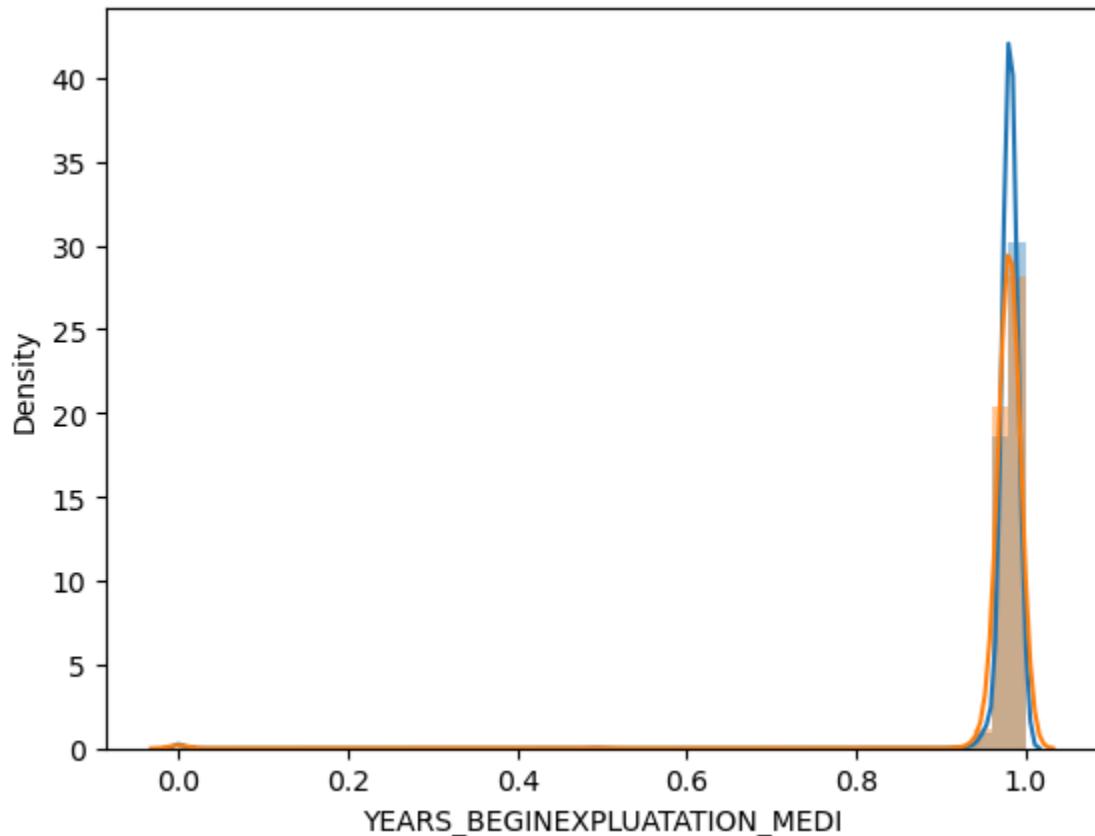
```
Out[86]: (array([4.99000e+02, 0.00000e+00, 0.00000e+00, 8.60000e+01, 5.00000e+00,
       1.00000e+00, 7.00000e+00, 1.45889e+05]),
       array([0.    , 0.125, 0.25 , 0.375, 0.5   , 0.625, 0.75 , 0.875, 1.    ]),
       <BarContainer object of 8 artists>)
```

```
Out[86]: (array([5.4000e+01, 0.0000e+00, 0.0000e+00, 7.0000e+00, 0.0000e+00,
       0.0000e+00, 0.0000e+00, 1.0956e+04]),
       array([0.    , 0.125, 0.25 , 0.375, 0.5   , 0.625, 0.75 , 0.875, 1.    ]),
       <BarContainer object of 8 artists>)
```



```
Out[86]: <Axes: xlabel='YEARS_BEGINEXPLUATATION_MEDI', ylabel='Density'>
```

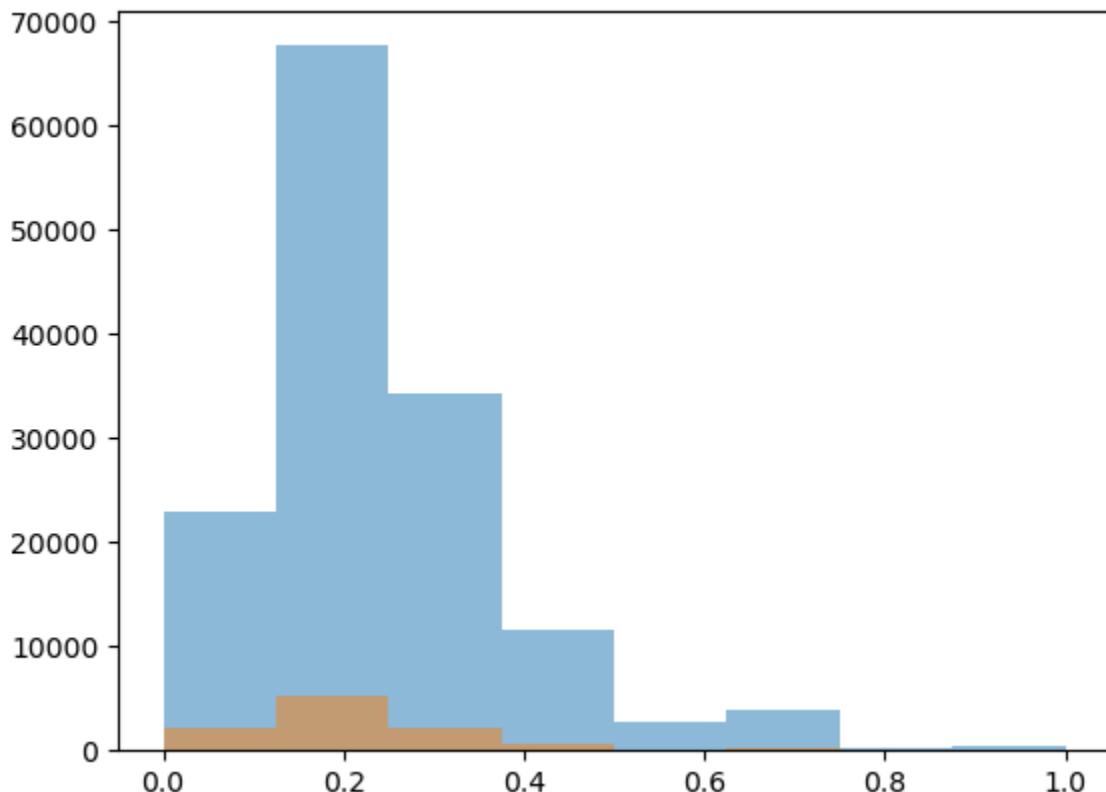
```
Out[86]: <Axes: xlabel='YEARS_BEGINEXPLUATATION_MEDI', ylabel='Density'>
```



Plot of FLOORSMAX_MEDI

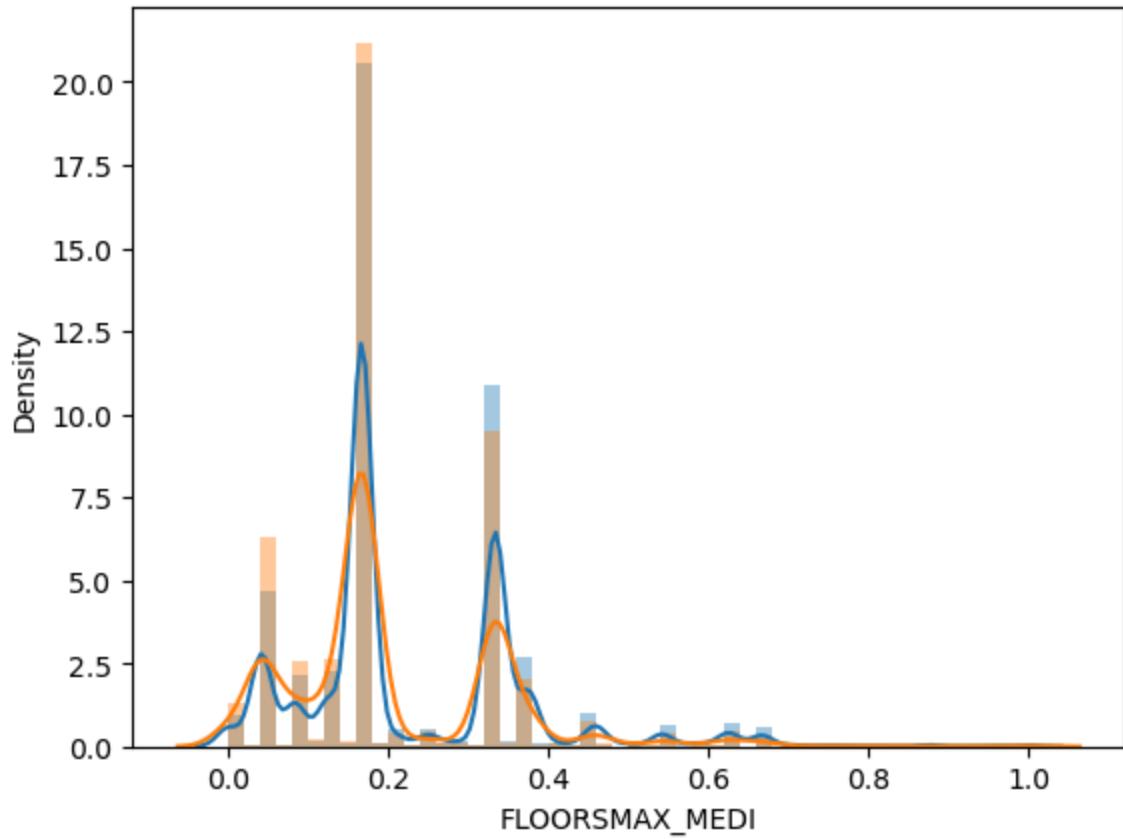
```
Out[86]: (array([22944., 67645., 34194., 11662., 2690., 3810., 274., 511.]),  
 array([0. , 0.125, 0.25 , 0.375, 0.5 , 0.625, 0.75 , 0.875, 1. ]),  
 <BarContainer object of 8 artists>)
```

```
Out[86]: (array([2258., 5279., 2217., 667., 129., 175., 15., 21.]),  
 array([0. , 0.125, 0.25 , 0.375, 0.5 , 0.625, 0.75 , 0.875, 1. ]),  
 <BarContainer object of 8 artists>)
```



```
Out[86]: <Axes: xlabel='FLOORSMAX_MEDI', ylabel='Density'>
```

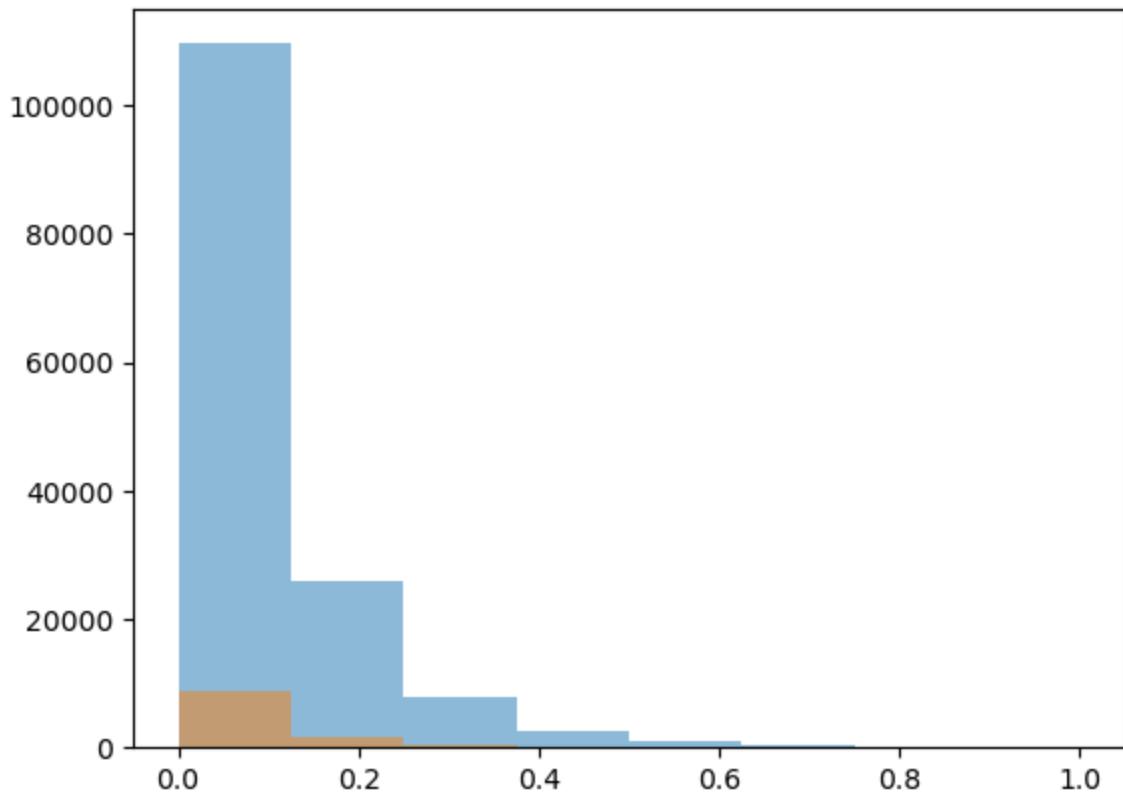
```
Out[86]: <Axes: xlabel='FLOORSMAX_MEDI', ylabel='Density'>
```



Plot of TOTALAREA_MODE

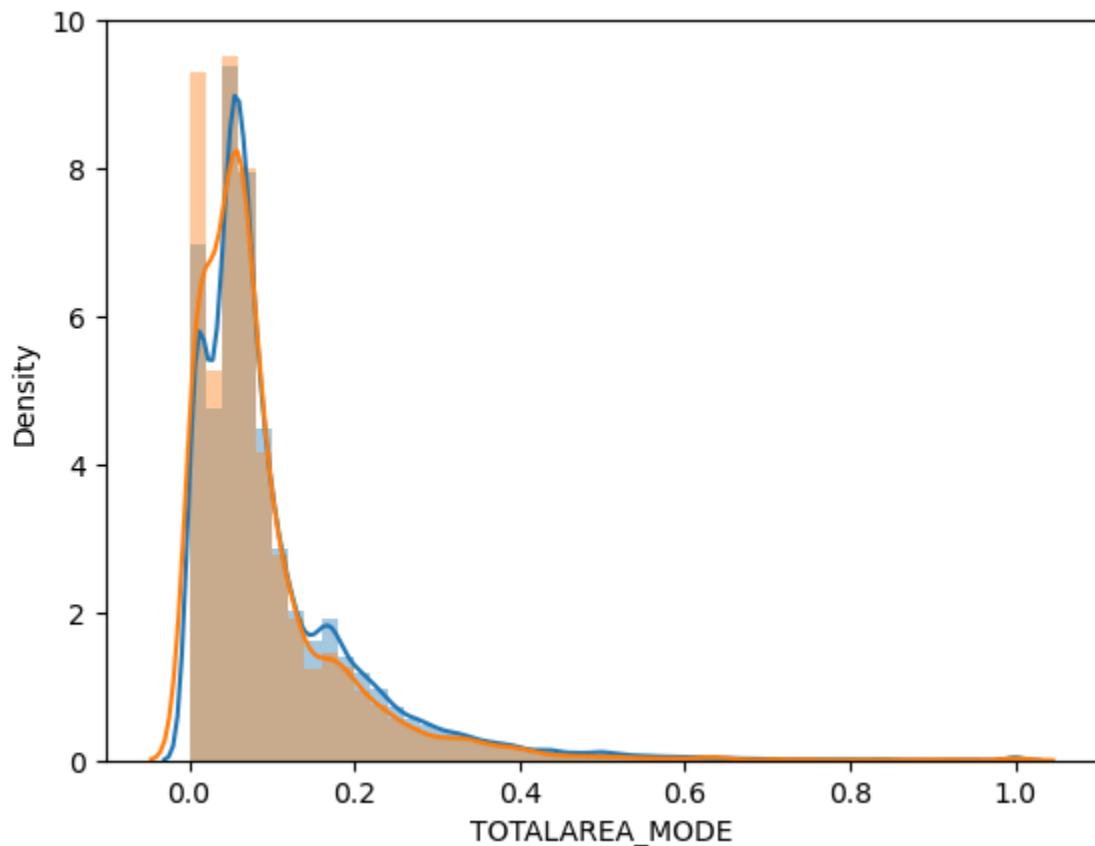
```
Out[86]: (array([109651.,  26086.,   7770.,   2587.,   1116.,    382.,    147.,
       222.]),
       array([0.     , 0.125, 0.25 , 0.375, 0.5   , 0.625, 0.75 , 0.875, 1.     ]),
       <BarContainer object of 8 artists>)

Out[86]: (array([8.826e+03, 1.602e+03, 4.570e+02, 1.380e+02, 4.800e+01, 2.500e+01,
       6.000e+00, 1.700e+01]),
       array([0.     , 0.125, 0.25 , 0.375, 0.5   , 0.625, 0.75 , 0.875, 1.     ]),
       <BarContainer object of 8 artists>)
```



```
Out[86]: <Axes: xlabel='TOTALAREA_MODE', ylabel='Density'>
```

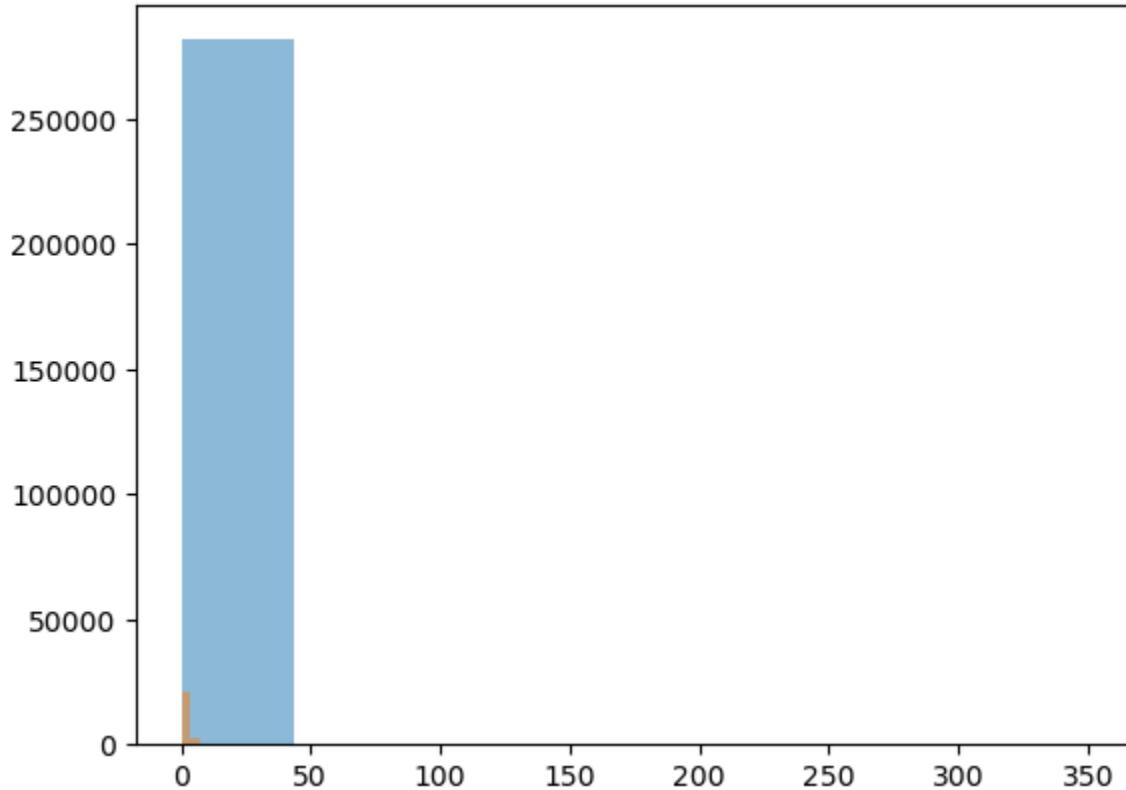
```
Out[86]: <Axes: xlabel='TOTALAREA_MODE', ylabel='Density'>
```



Plot of OBS_30_CNT_SOCIAL_CIRCLE

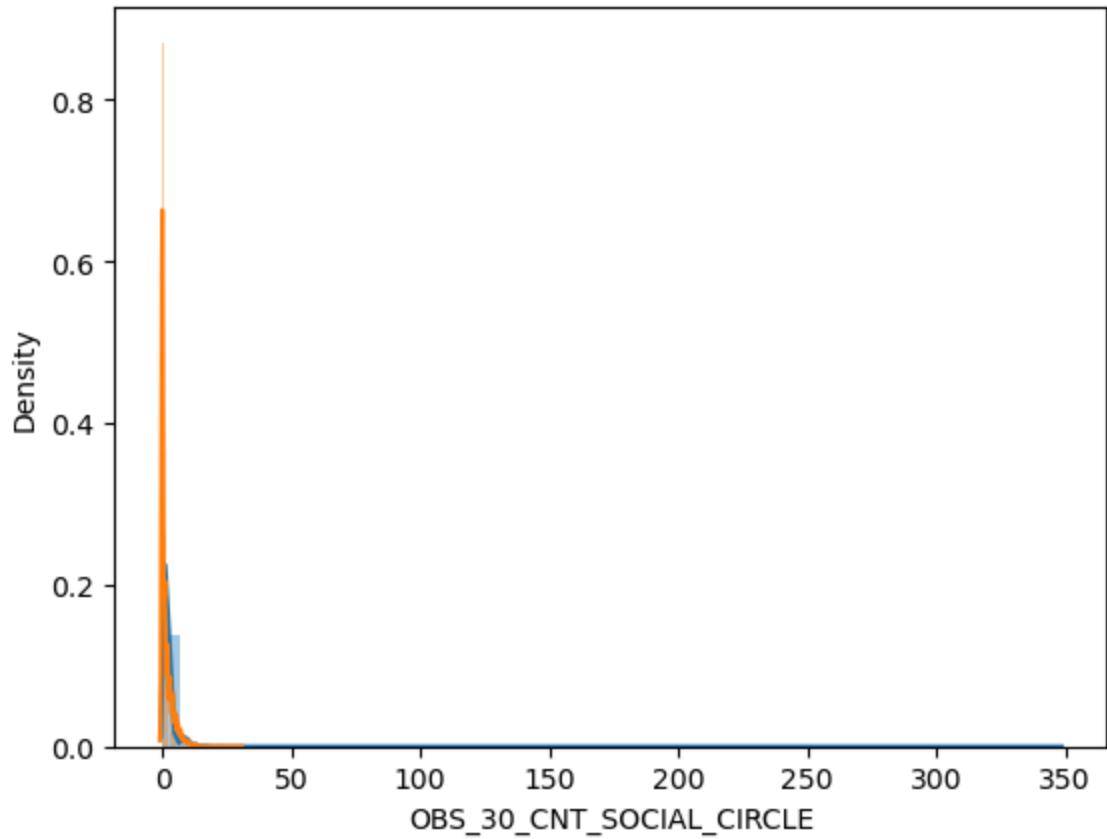
```
Out[86]: (array([2.81699e+05, 1.00000e+00, 0.00000e+00, 0.00000e+00, 0.00000e+00,
       0.00000e+00, 0.00000e+00, 1.00000e+00]),
       array([ 0. , 43.5, 87. , 130.5, 174. , 217.5, 261. , 304.5, 348. ]),
       <BarContainer object of 8 artists>)

Out[86]: (array([2.0978e+04, 3.0210e+03, 6.1800e+02, 1.2900e+02, 2.8000e+01,
       1.1000e+01, 3.0000e+00, 1.0000e+00]),
       array([ 0. , 3.75, 7.5 , 11.25, 15. , 18.75, 22.5 , 26.25, 30. ]),
       <BarContainer object of 8 artists>)
```



```
Out[86]: <Axes: xlabel='OBS_30_CNT_SOCIAL_CIRCLE', ylabel='Density'>
```

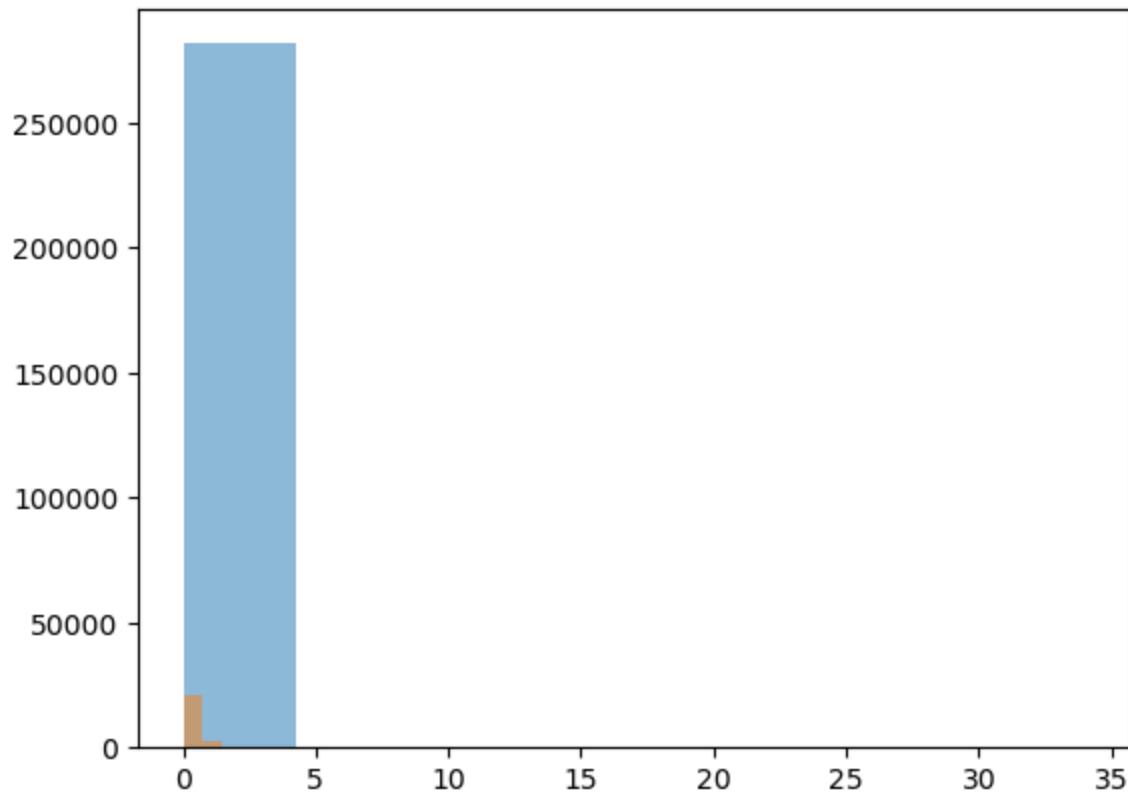
```
Out[86]: <Axes: xlabel='OBS_30_CNT_SOCIAL_CIRCLE', ylabel='Density'>
```



Plot of `DEF_30_CNT_SOCIAL_CIRCLE`

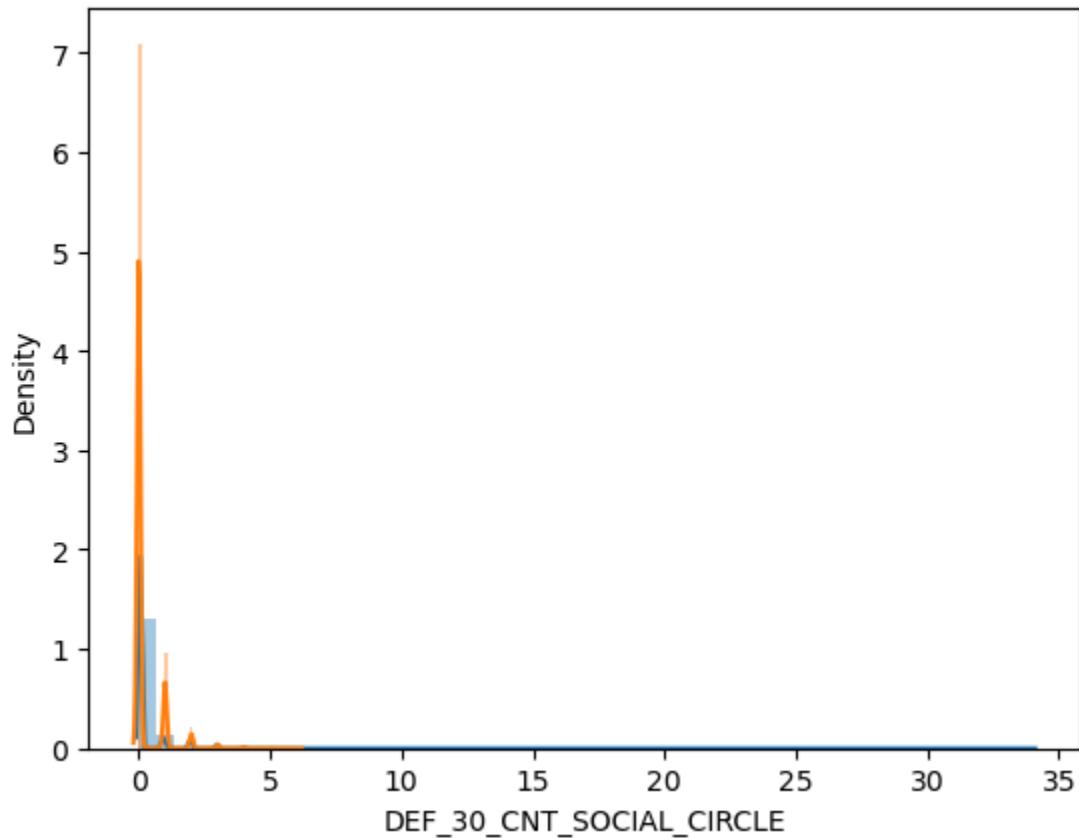
```
Out[86]: (array([2.8164e+05, 6.0000e+01, 0.0000e+00, 0.0000e+00, 0.0000e+00,
   0.0000e+00, 0.0000e+00, 1.0000e+00]),
 array([ 0. ,  4.25,  8.5 , 12.75, 17. , 21.25, 25.5 , 29.75, 34. ]),
 <BarContainer object of 8 artists>)

Out[86]: (array([2.1098e+04, 2.8740e+03, 6.2000e+02, 0.0000e+00, 1.5300e+02,
   3.5000e+01, 8.0000e+00, 1.0000e+00]),
 array([0. , 0.75, 1.5 , 2.25, 3. , 3.75, 4.5 , 5.25, 6. ]),
 <BarContainer object of 8 artists>)
```



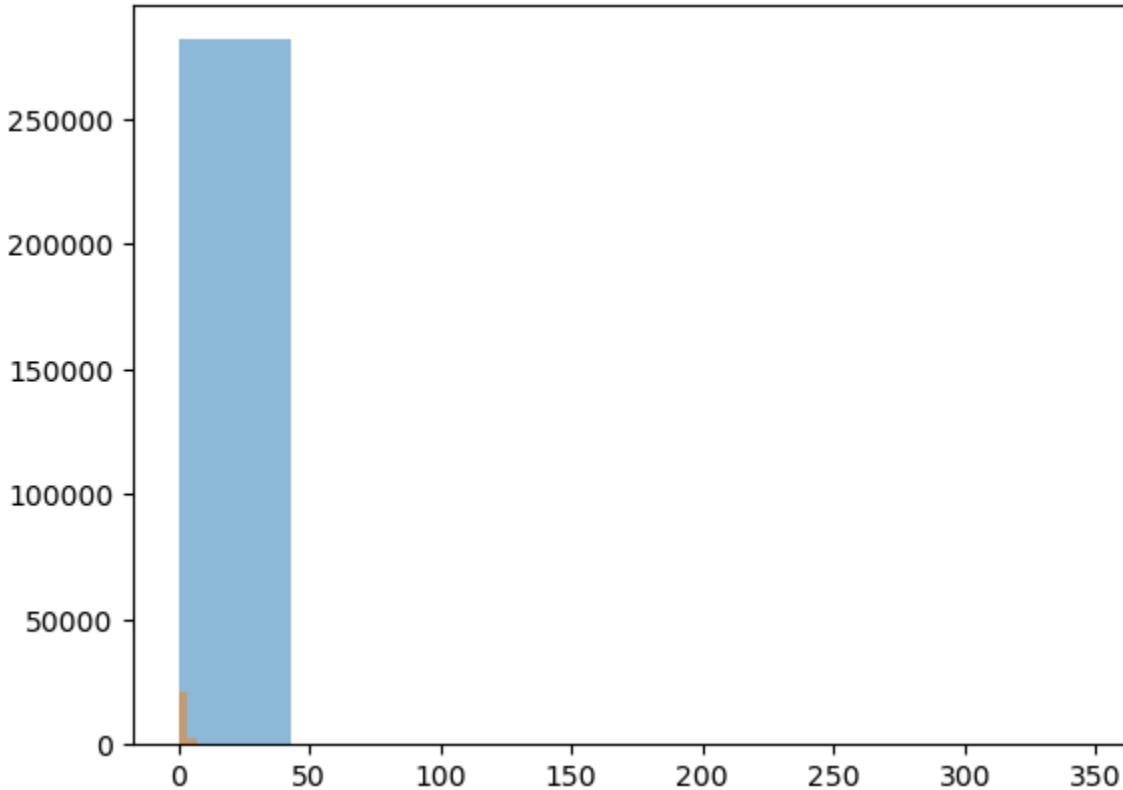
```
Out[86]: <Axes: xlabel='DEF_30_CNT_SOCIAL_CIRCLE', ylabel='Density'>
```

```
Out[86]: <Axes: xlabel='DEF_30_CNT_SOCIAL_CIRCLE', ylabel='Density'>
```



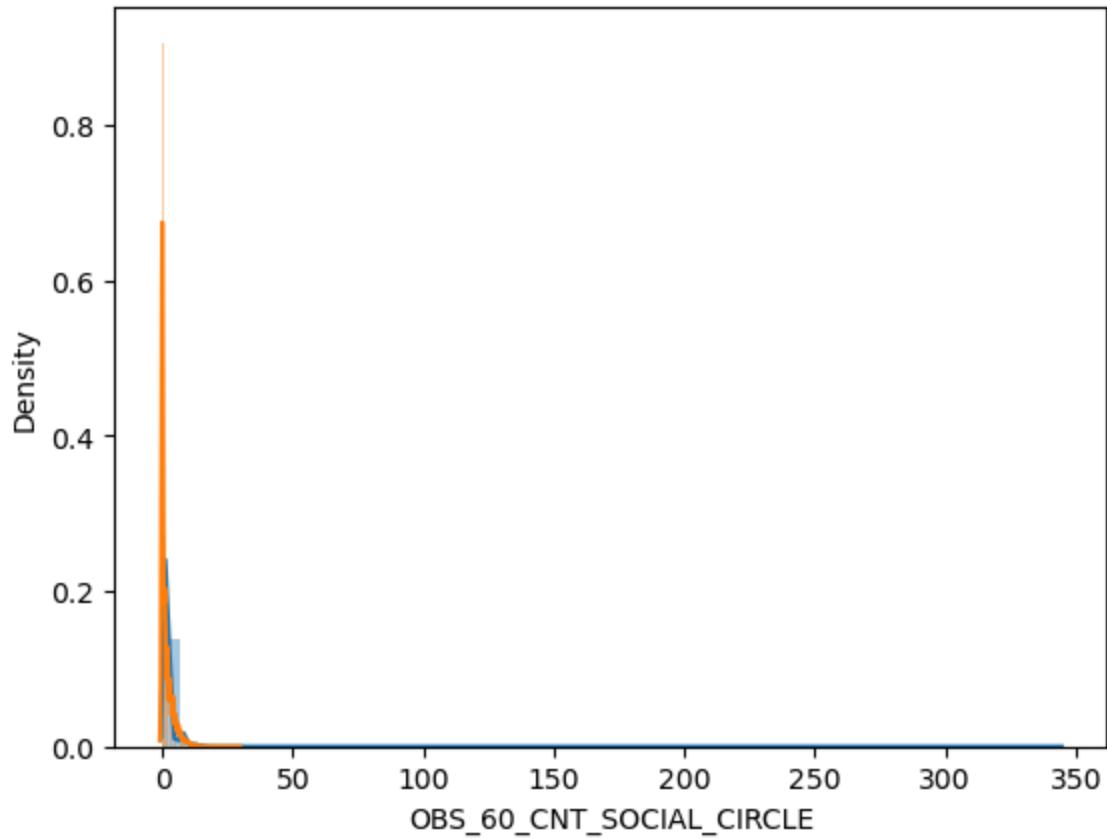
Plot of OBS_60_CNT_SOCIAL_CIRCLE

```
Out[86]: (array([2.81699e+05, 1.00000e+00, 0.00000e+00, 0.00000e+00, 0.00000e+00,
   0.00000e+00, 0.00000e+00, 1.00000e+00]),  
 array([ 0.,  43.,  86., 129., 172., 215., 258., 301., 344.]),  
 <BarContainer object of 8 artists>)  
  
Out[86]: (array([2.1033e+04, 2.9860e+03, 5.3000e+02, 1.9700e+02, 2.9000e+01,  
 8.0000e+00, 5.0000e+00, 1.0000e+00]),  
 array([ 0. ,  3.625,  7.25 , 10.875, 14.5 , 18.125, 21.75 , 25.375,  
 29. ]),  
 <BarContainer object of 8 artists>)
```



```
Out[86]: <Axes: xlabel='OBS_60_CNT_SOCIAL_CIRCLE', ylabel='Density'>
```

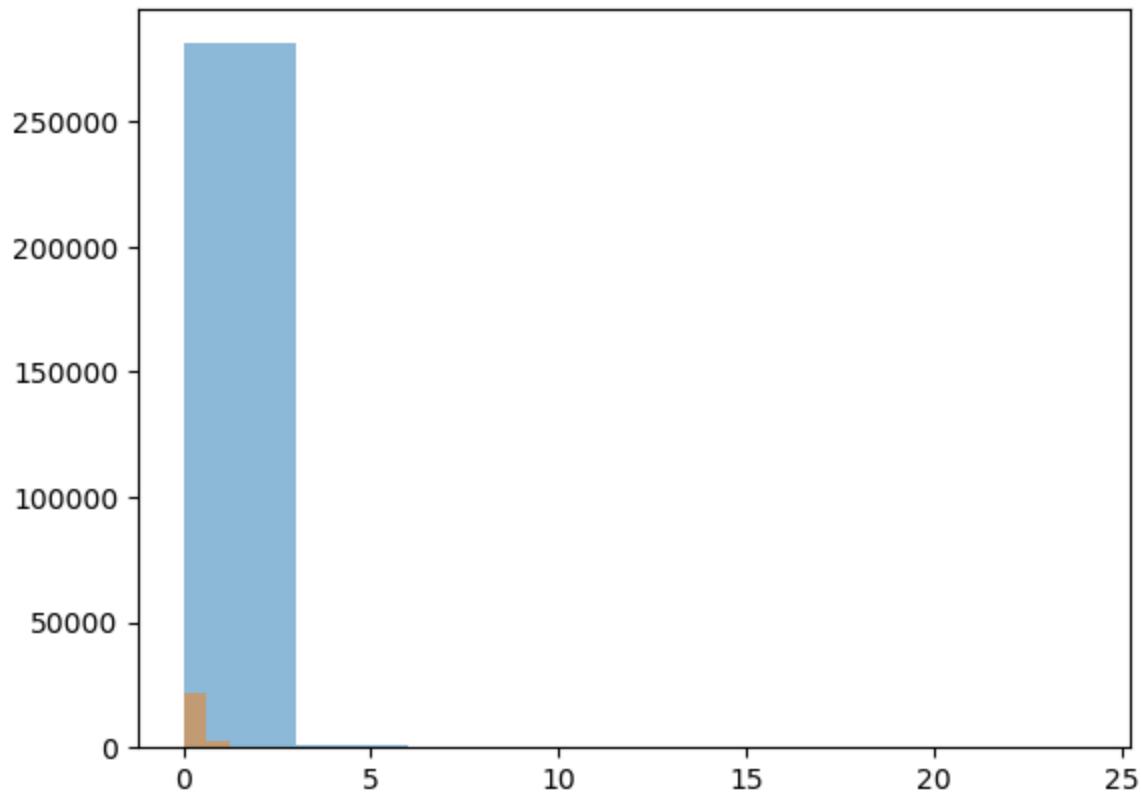
```
Out[86]: <Axes: xlabel='OBS_60_CNT_SOCIAL_CIRCLE', ylabel='Density'>
```



Plot of `DEF_60_CNT_SOCIAL_CIRCLE`

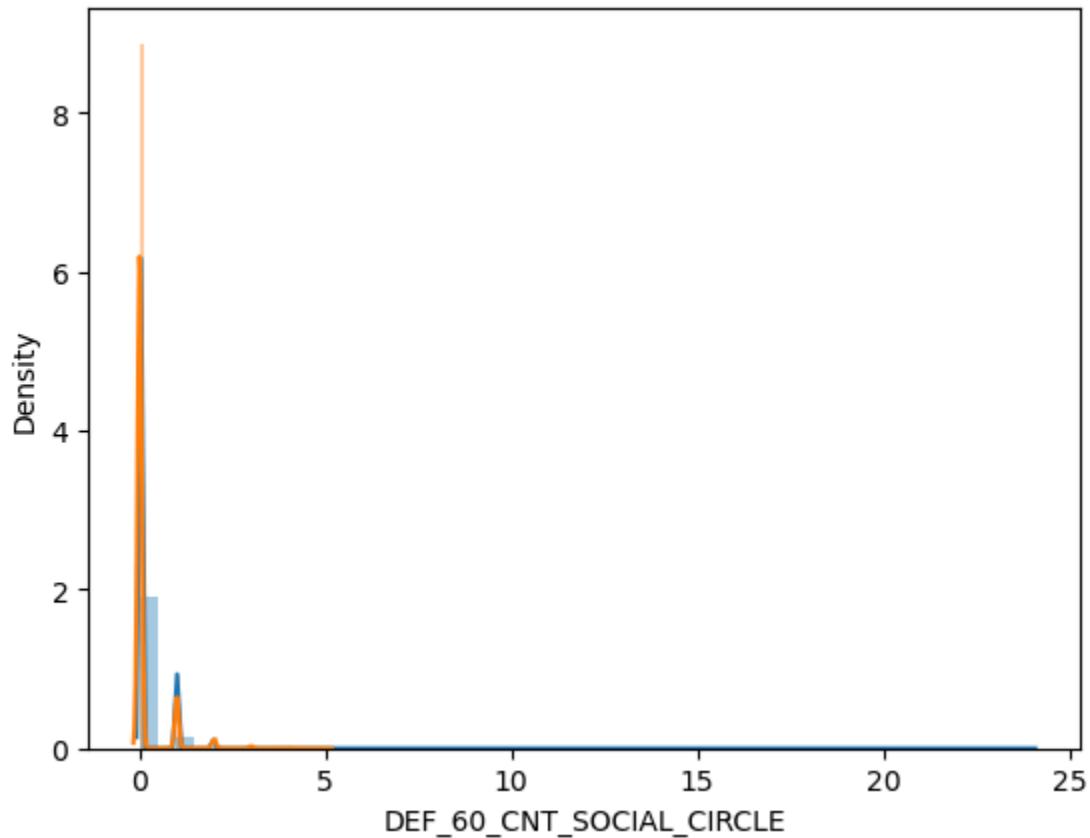
```
Out[86]: (array([2.81056e+05, 6.40000e+02, 4.00000e+00, 0.00000e+00, 0.00000e+00,
   0.00000e+00, 0.00000e+00, 1.00000e+00]),
 array([ 0.,  3.,  6.,  9., 12., 15., 18., 21., 24.]),
 <BarContainer object of 8 artists>)
```

```
Out[86]: (array([2.1994e+04, 2.2970e+03, 0.0000e+00, 3.8500e+02, 9.5000e+01,
   0.0000e+00, 1.5000e+01, 3.0000e+00]),
 array([0.    , 0.625, 1.25 , 1.875, 2.5  , 3.125, 3.75 , 4.375, 5.    ]),
 <BarContainer object of 8 artists>)
```



```
Out[86]: <Axes: xlabel='DEF_60_CNT_SOCIAL_CIRCLE', ylabel='Density'>
```

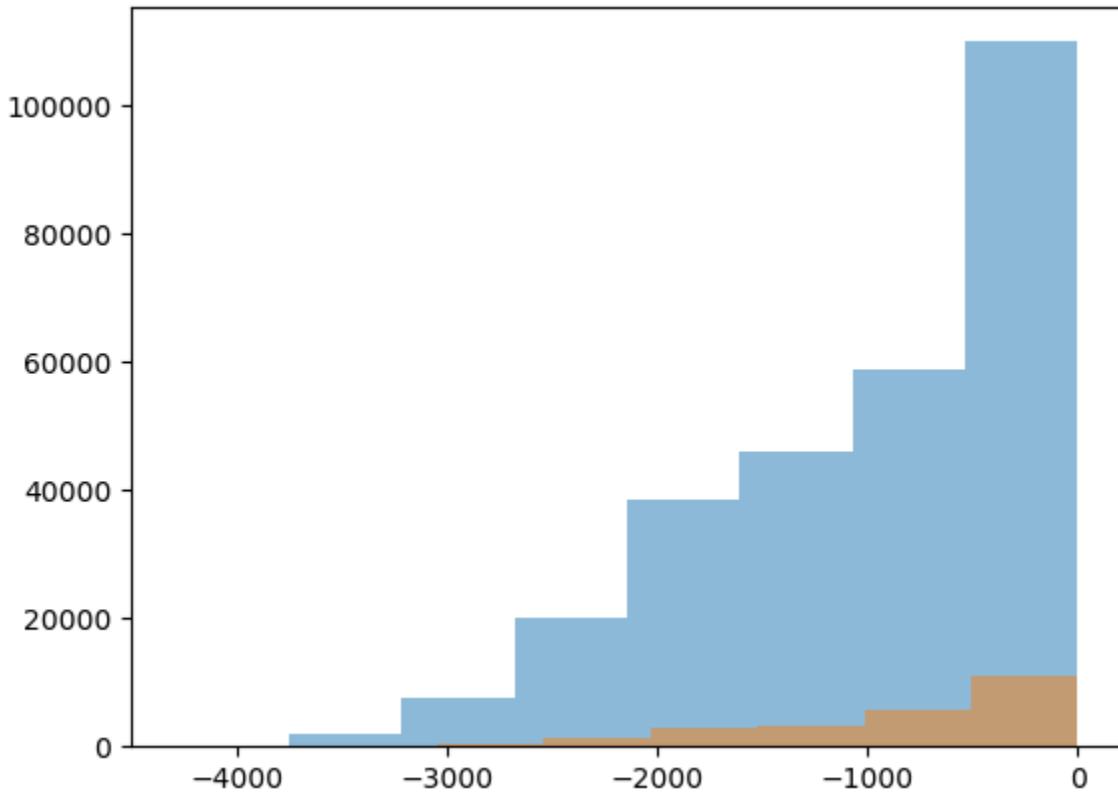
```
Out[86]: <Axes: xlabel='DEF_60_CNT_SOCIAL_CIRCLE', ylabel='Density'>
```



Plot of DAYS_LAST_PHONE_CHANGE

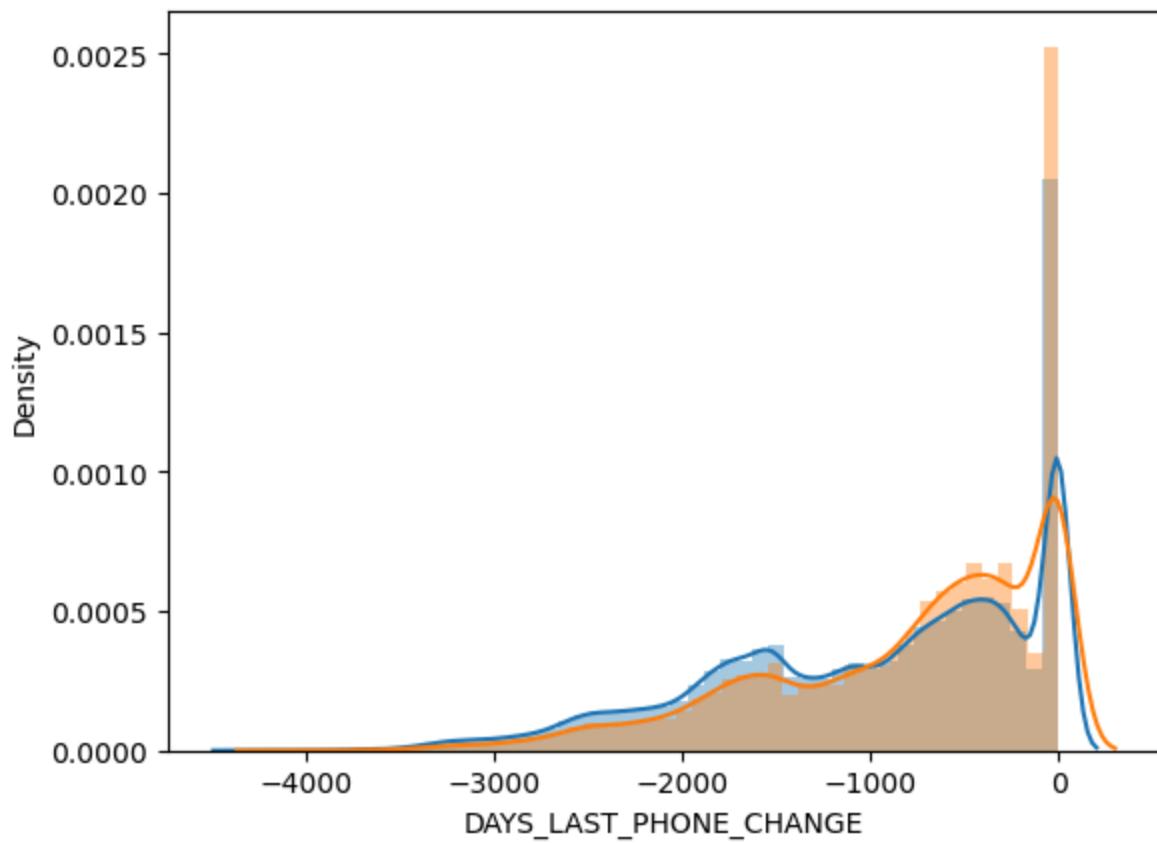
```
Out[86]: (array([ 111.,  2095.,  7428., 19991., 38558., 45868., 58816.,
   109818.]),
 array([-4292. , -3755.5, -3219. , -2682.5, -2146. , -1609.5, -1073. ,
  -536.5,      0. ]),
 <BarContainer object of 8 artists>)

Out[86]: (array([ 15., 173., 553., 1263., 2866., 3173., 5664., 11118.]),
 array([-4070. , -3561.25, -3052.5, -2543.75, -2035. , -1526.25,
  -1017.5, -508.75,      0. ]),
 <BarContainer object of 8 artists>)
```



```
Out[86]: <Axes: xlabel='DAYSLASTPHONECHANGE', ylabel='Density'>
```

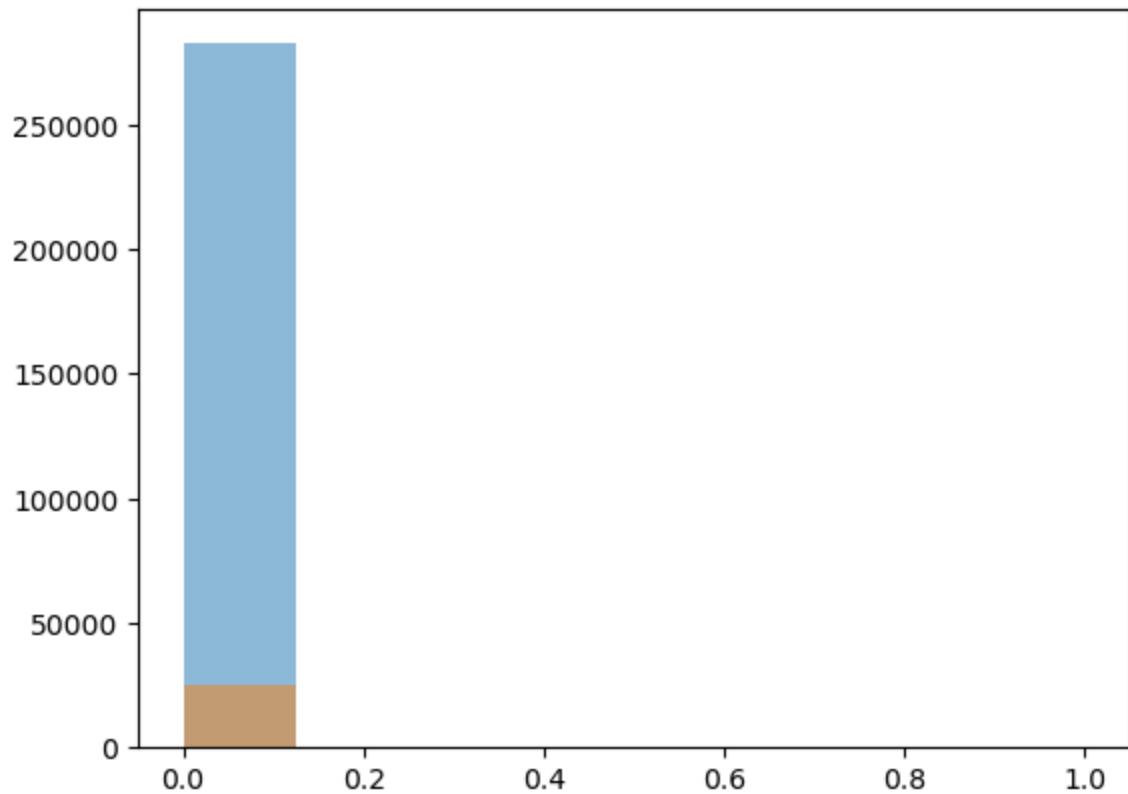
```
Out[86]: <Axes: xlabel='DAYSLASTPHONECHANGE', ylabel='Density'>
```



Plot of FLAG_DOCUMENT_2

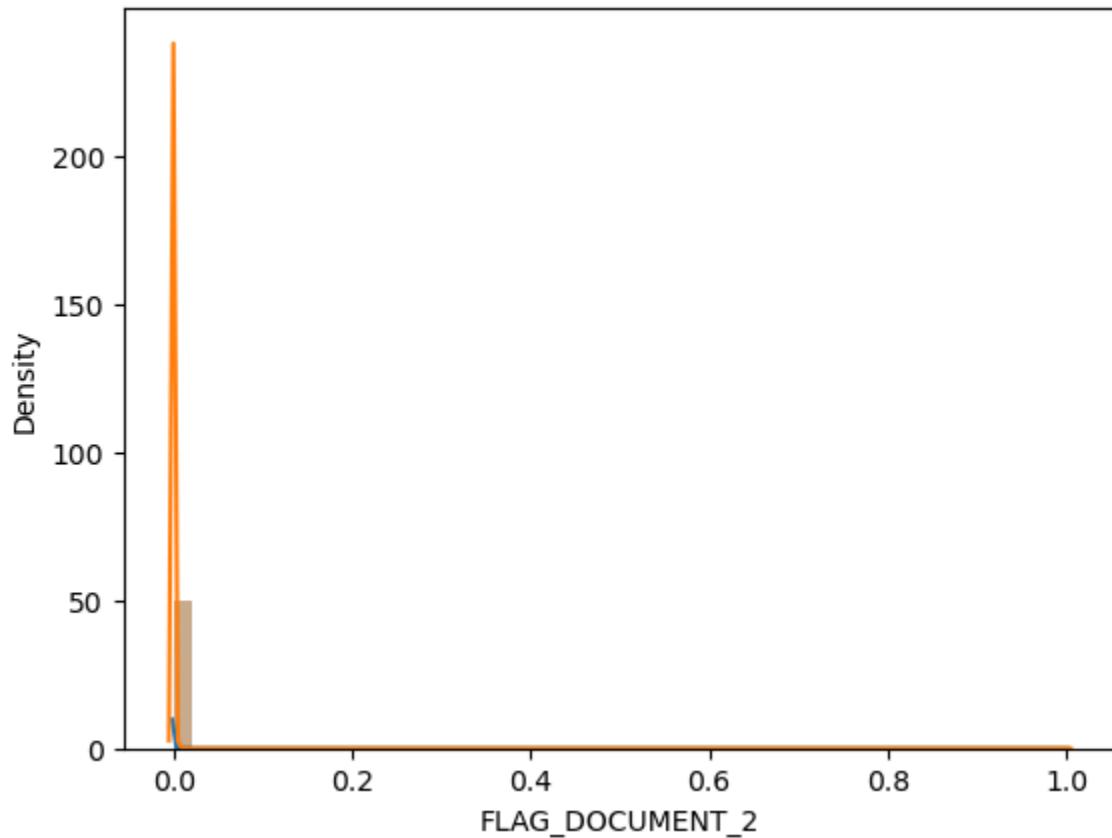
```
Out[86]: (array([2.82677e+05, 0.00000e+00, 0.00000e+00, 0.00000e+00, 0.00000e+00,
       0.00000e+00, 0.00000e+00, 9.00000e+00]),
       array([0.    , 0.125, 0.25 , 0.375, 0.5   , 0.625, 0.75 , 0.875, 1.    ]),
       <BarContainer object of 8 artists>)

Out[86]: (array([2.4821e+04, 0.0000e+00, 0.0000e+00, 0.0000e+00, 0.0000e+00,
       0.0000e+00, 0.0000e+00, 4.0000e+00]),
       array([0.    , 0.125, 0.25 , 0.375, 0.5   , 0.625, 0.75 , 0.875, 1.    ]),
       <BarContainer object of 8 artists>)
```



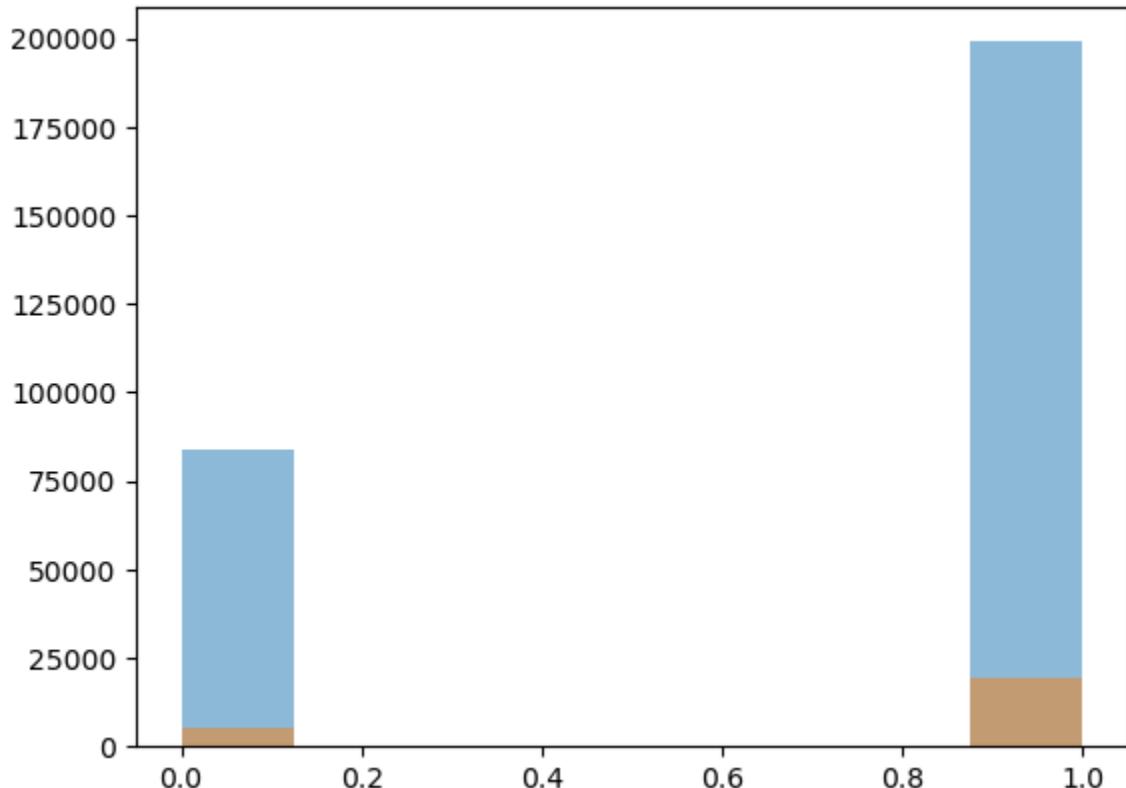
```
Out[86]: <Axes: xlabel='FLAG_DOCUMENT_2', ylabel='Density'>
```

```
Out[86]: <Axes: xlabel='FLAG_DOCUMENT_2', ylabel='Density'>
```



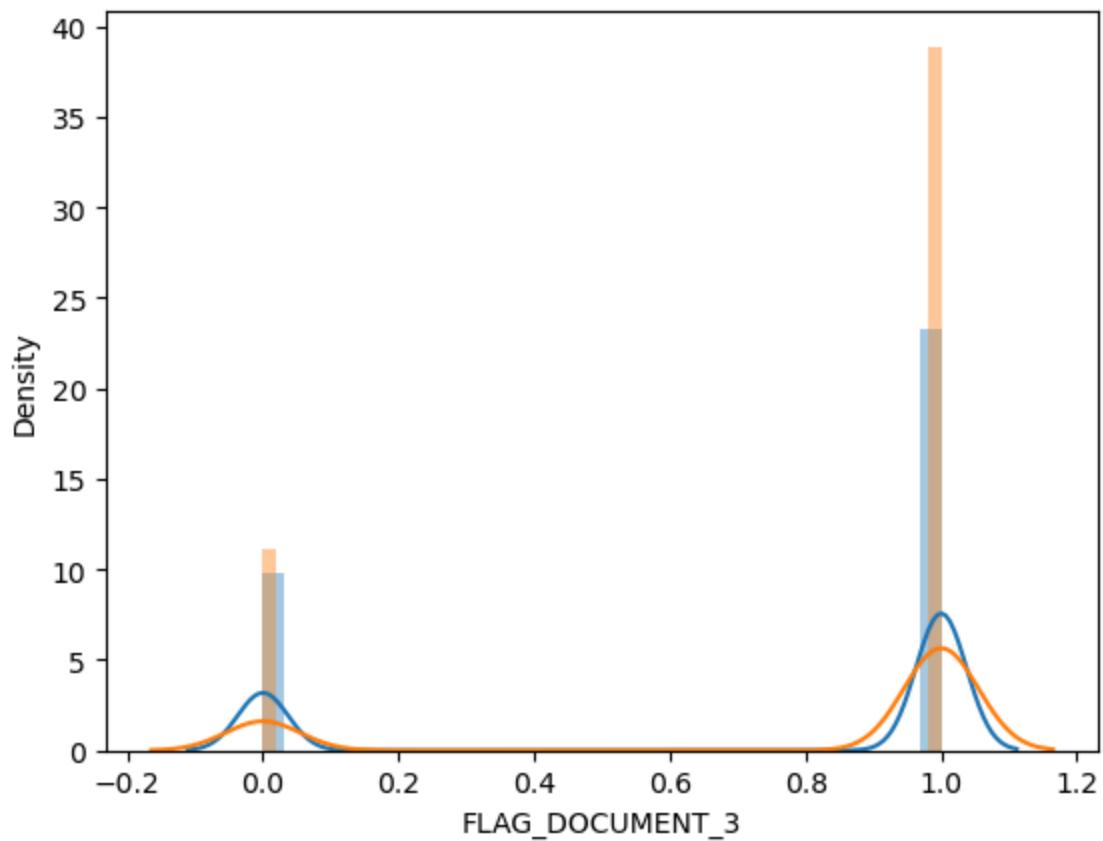
Plot of FLAG_DOCUMENT_3

```
Out[86]: (array([ 83658.,      0.,      0.,      0.,      0.,      0.,      0.,
   199028.]),
 array([0.    , 0.125, 0.25 , 0.375, 0.5   , 0.625, 0.75 , 0.875, 1.    ]),
 <BarContainer object of 8 artists>
Out[86]: (array([ 5513.,      0.,      0.,      0.,      0.,      0.,      0.,      0.,
  19312.]),
 array([0.    , 0.125, 0.25 , 0.375, 0.5   , 0.625, 0.75 , 0.875, 1.    ]),
 <BarContainer object of 8 artists>)
```



```
Out[86]: <Axes: xlabel='FLAG_DOCUMENT_3', ylabel='Density'>
```

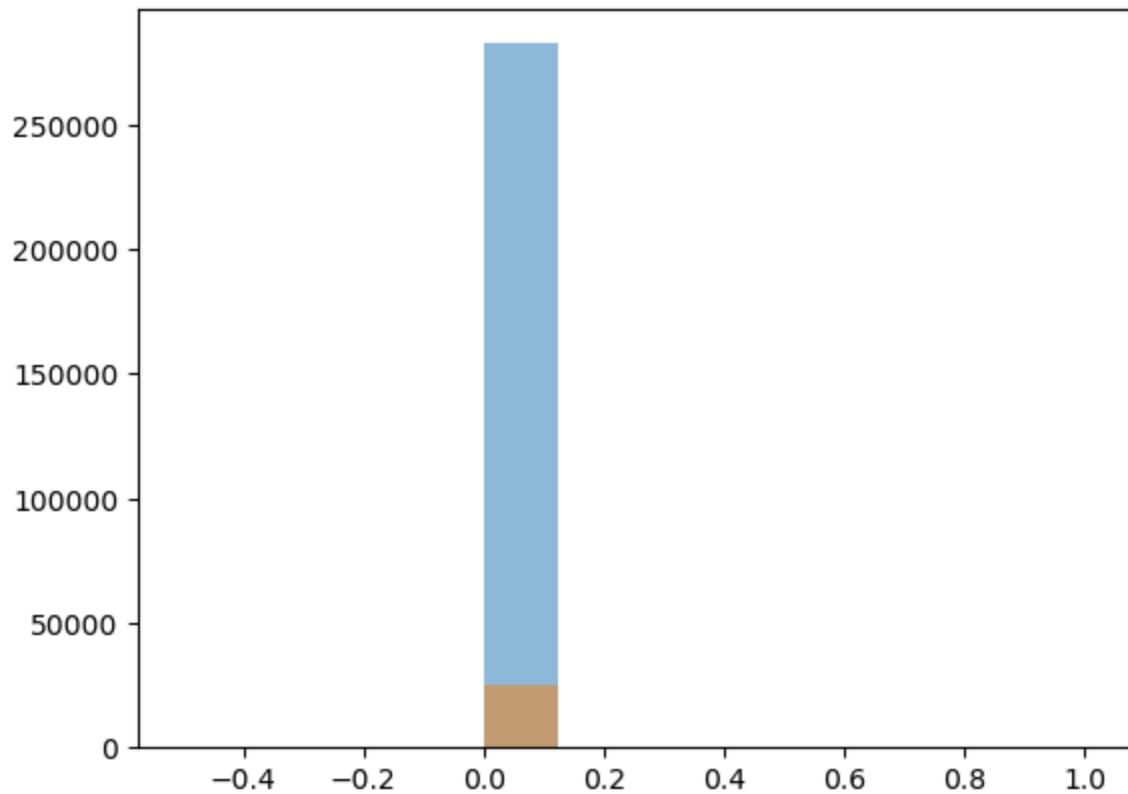
```
Out[86]: <Axes: xlabel='FLAG_DOCUMENT_3', ylabel='Density'>
```



Plot of `FLAG_DOCUMENT_4`

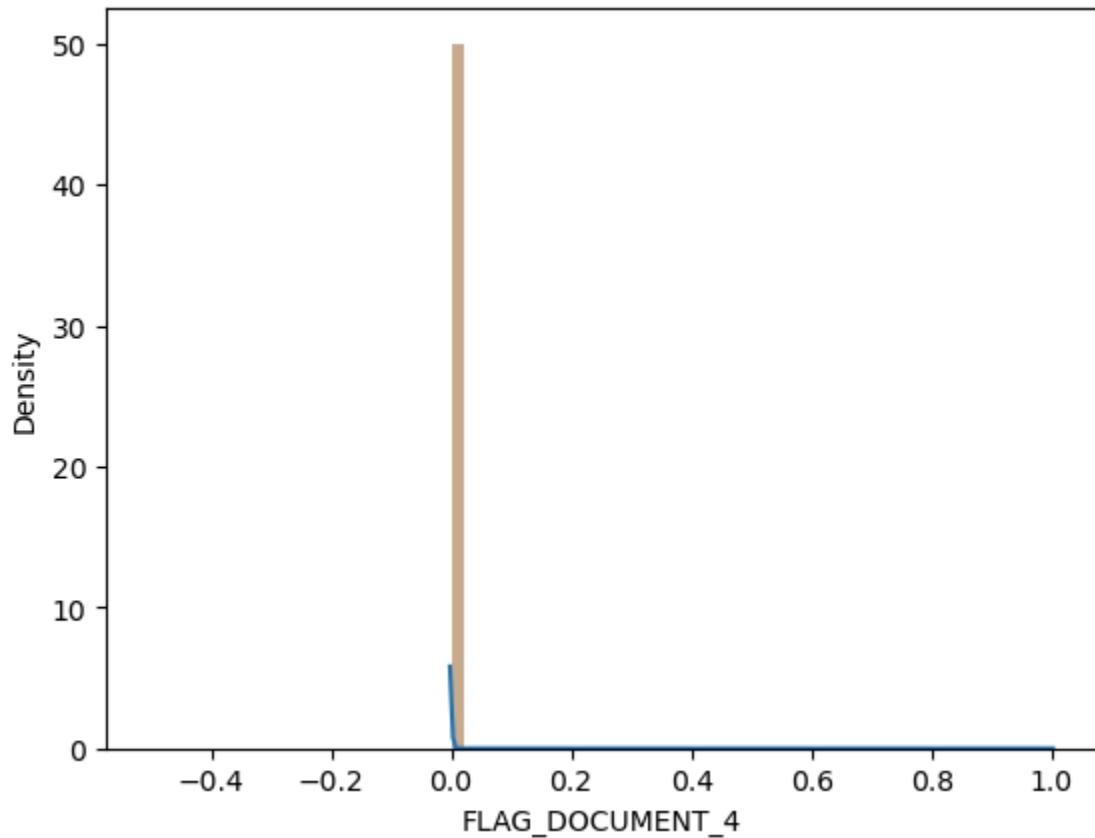
```
Out[86]: (array([2.82661e+05, 0.00000e+00, 0.00000e+00, 0.00000e+00, 0.00000e+00,
       0.00000e+00, 0.00000e+00, 2.50000e+01]),
       array([0.    , 0.125, 0.25 , 0.375, 0.5   , 0.625, 0.75 , 0.875, 1.    ]),
       <BarContainer object of 8 artists>)

Out[86]: (array([    0.,     0.,     0.,     0., 24825.,     0.,     0.,     0.]),
       array([-0.5  , -0.375, -0.25 , -0.125,  0.    ,  0.125,  0.25 ,  0.375,
              0.5  ]),
       <BarContainer object of 8 artists>)
```



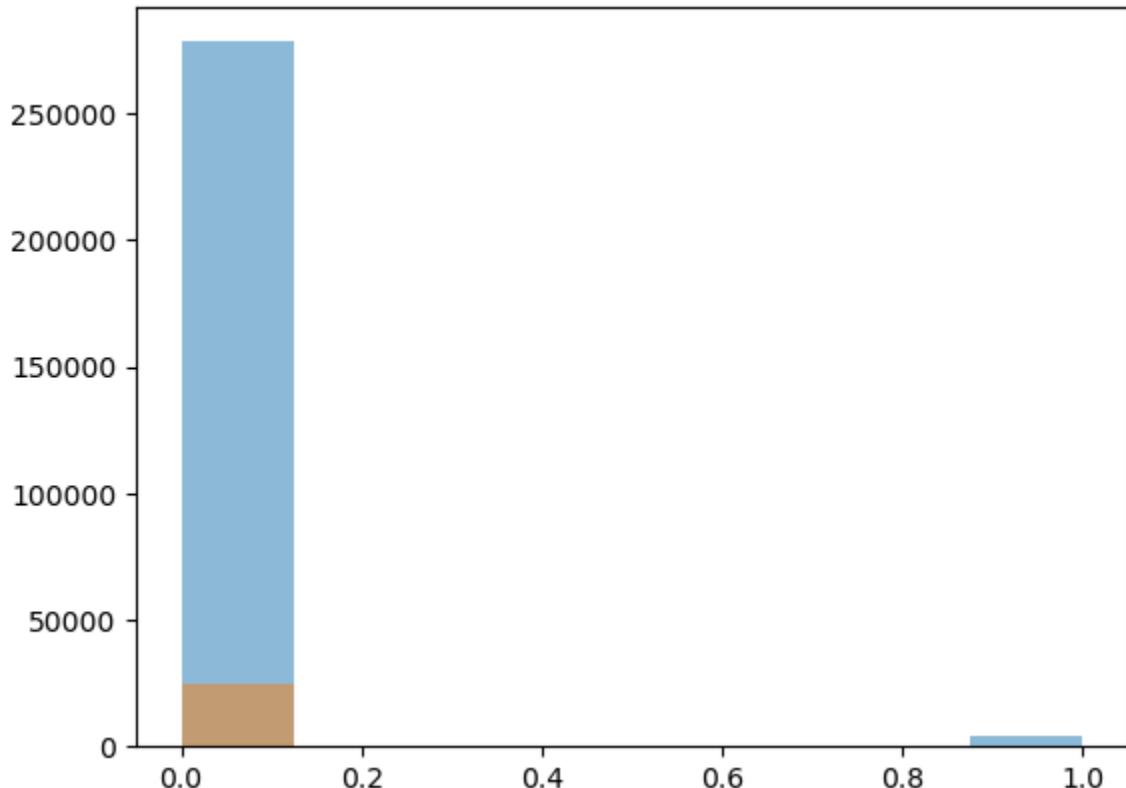
```
Out[86]: <Axes: xlabel='FLAG_DOCUMENT_4', ylabel='Density'>
```

```
Out[86]: <Axes: xlabel='FLAG_DOCUMENT_4', ylabel='Density'>
```



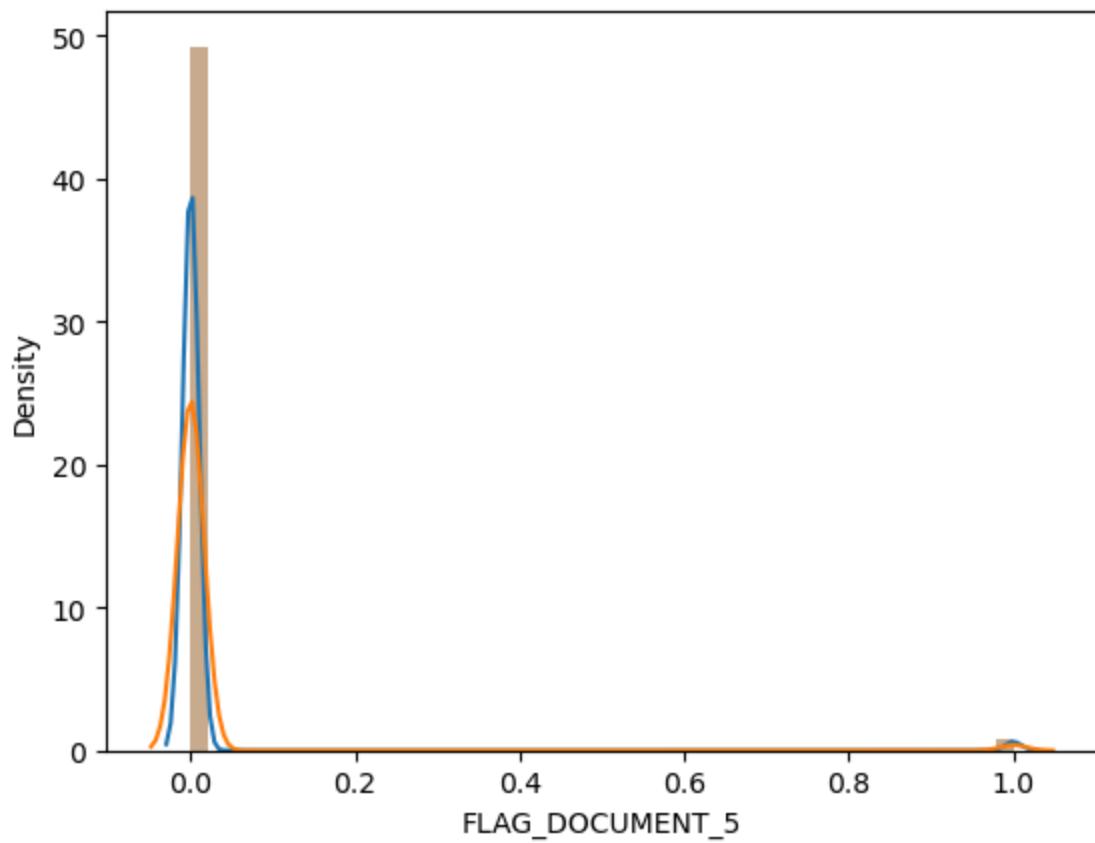
Plot of FLAG_DOCUMENT_5

```
Out[86]: (array([278410.,      0.,      0.,      0.,      0.,      0.,      0.,
   4276.]),
 array([0.    , 0.125, 0.25 , 0.375, 0.5  , 0.625, 0.75 , 0.875, 1.    ]),
 <BarContainer object of 8 artists>
Out[86]: (array([24453.,      0.,      0.,      0.,      0.,      0.,      0.,      372.]),
 array([0.    , 0.125, 0.25 , 0.375, 0.5  , 0.625, 0.75 , 0.875, 1.    ]),
 <BarContainer object of 8 artists>)
```



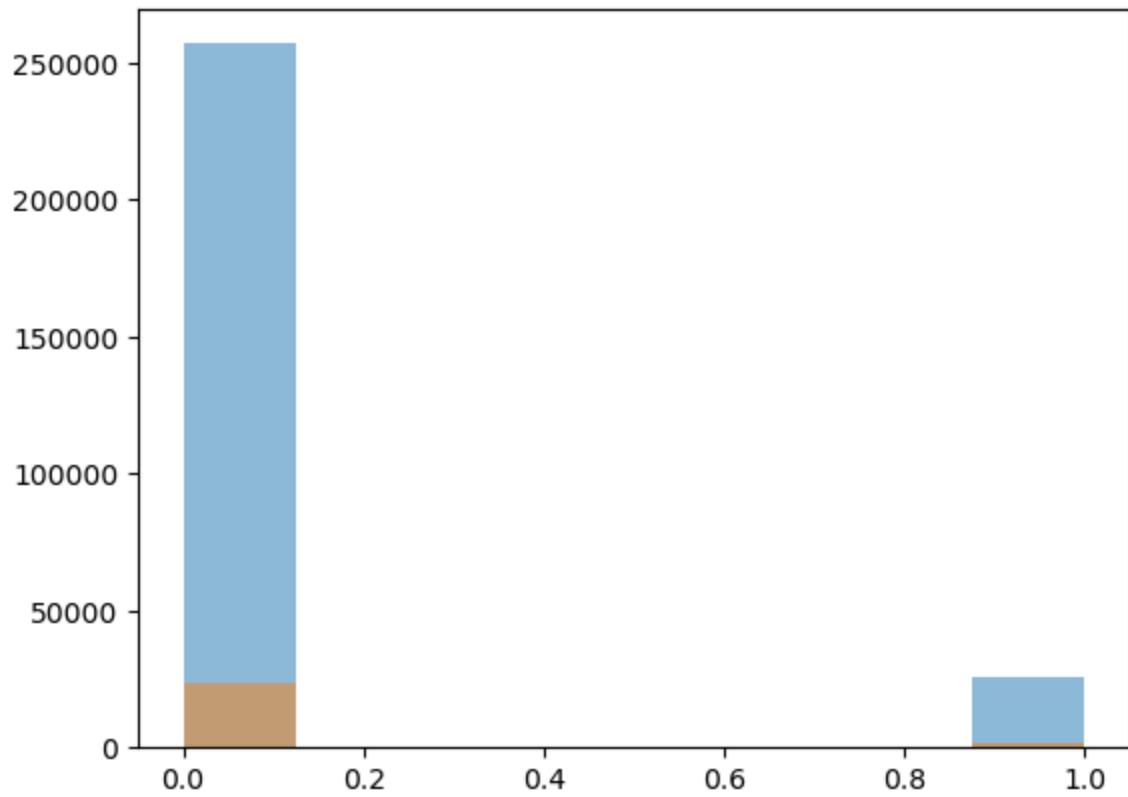
```
Out[86]: <Axes: xlabel='FLAG_DOCUMENT_5', ylabel='Density'>
```

```
Out[86]: <Axes: xlabel='FLAG_DOCUMENT_5', ylabel='Density'>
```



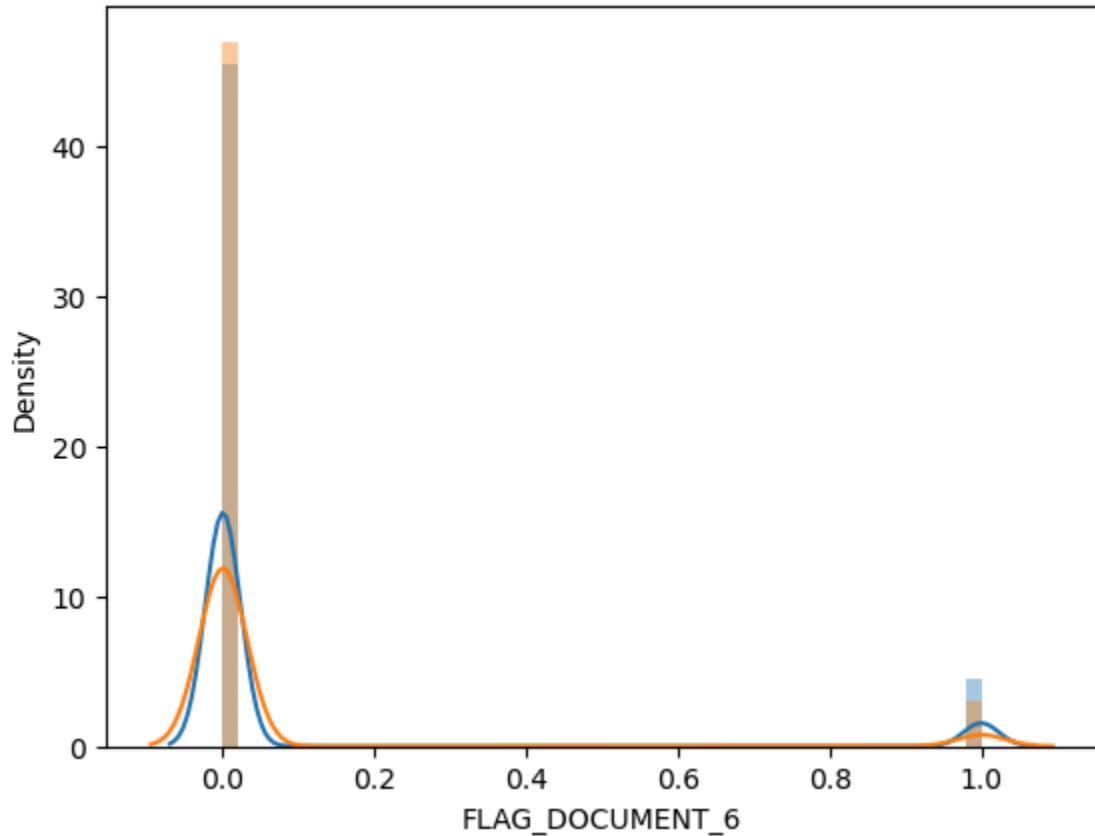
Plot of `FLAG_DOCUMENT_6`

```
Out[86]: (array([257115.,      0.,      0.,      0.,      0.,      0.,      0.,
   25571.]),
 array([0.     , 0.125, 0.25 , 0.375, 0.5   , 0.625, 0.75 , 0.875, 1.     ]),
 <BarContainer object of 8 artists>
Out[86]: (array([23318.,      0.,      0.,      0.,      0.,      0.,      0.,    1507.]),
 array([0.     , 0.125, 0.25 , 0.375, 0.5   , 0.625, 0.75 , 0.875, 1.     ]),
 <BarContainer object of 8 artists>)
```



```
Out[86]: <Axes: xlabel='FLAG_DOCUMENT_6', ylabel='Density'>
```

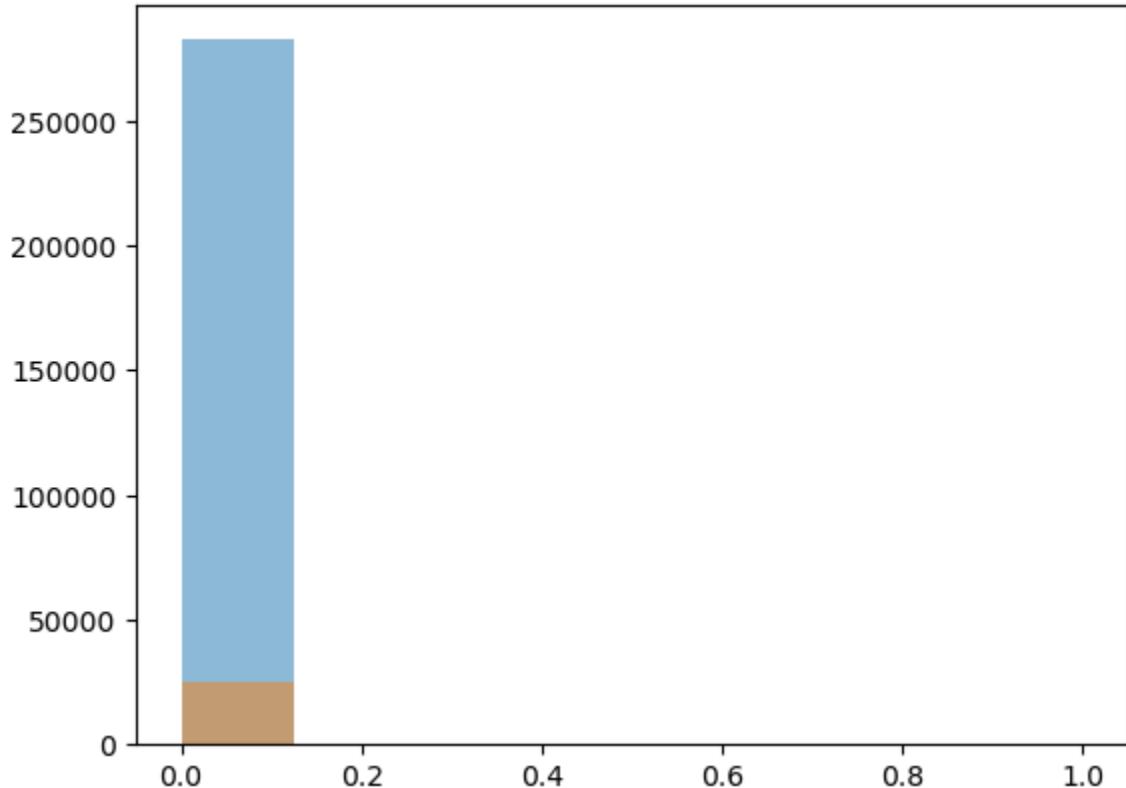
```
Out[86]: <Axes: xlabel='FLAG_DOCUMENT_6', ylabel='Density'>
```



Plot of FLAG_DOCUMENT_7

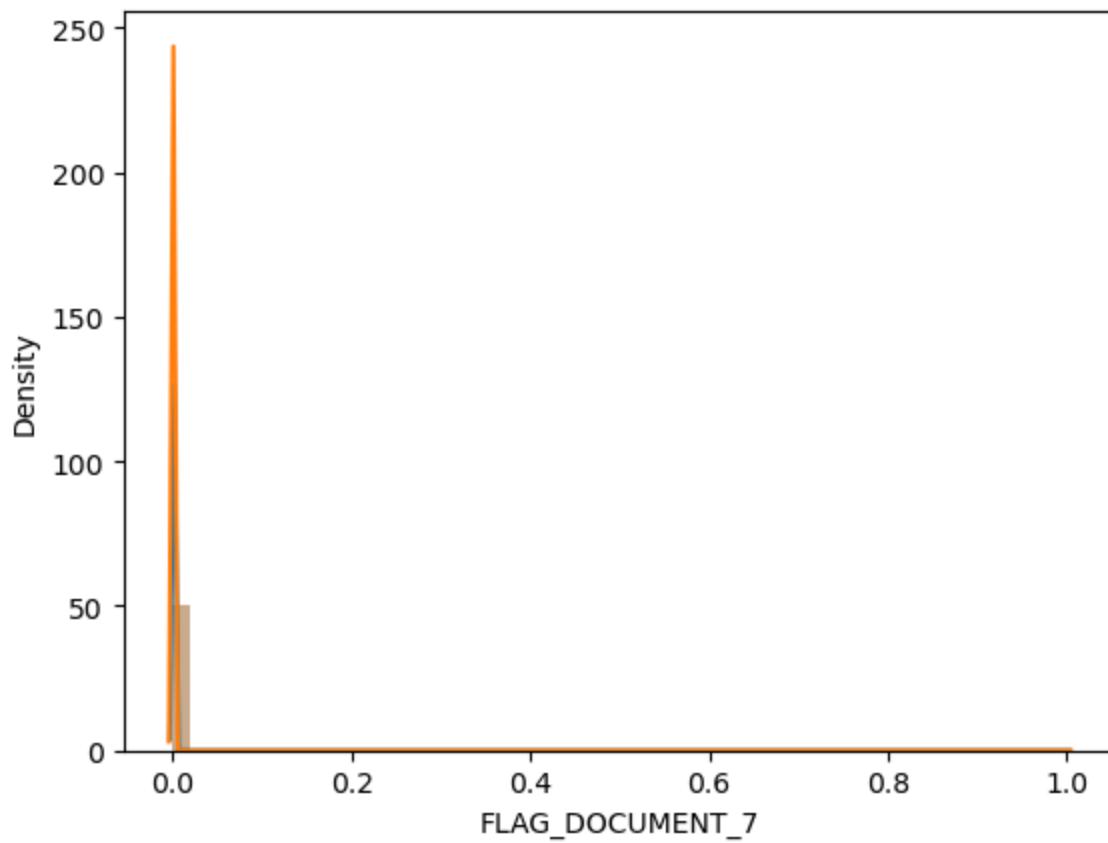
```
Out[86]: (array([2.8263e+05, 0.0000e+00, 0.0000e+00, 0.0000e+00, 0.0000e+00,
       0.0000e+00, 0.0000e+00, 5.6000e+01]),
       array([0.    , 0.125, 0.25 , 0.375, 0.5   , 0.625, 0.75 , 0.875, 1.    ]),
       <BarContainer object of 8 artists>)

Out[86]: (array([2.4822e+04, 0.0000e+00, 0.0000e+00, 0.0000e+00, 0.0000e+00,
       0.0000e+00, 0.0000e+00, 3.0000e+00]),
       array([0.    , 0.125, 0.25 , 0.375, 0.5   , 0.625, 0.75 , 0.875, 1.    ]),
       <BarContainer object of 8 artists>)
```



```
Out[86]: <Axes: xlabel='FLAG_DOCUMENT_7', ylabel='Density'>
```

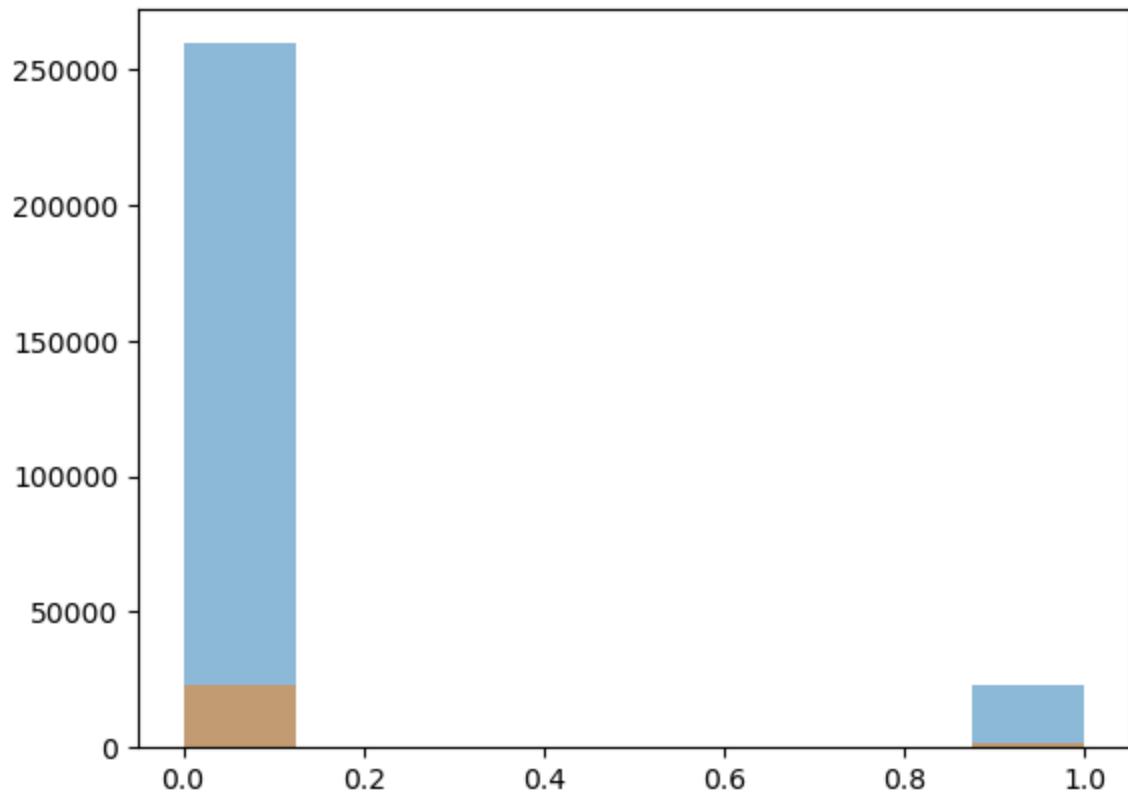
```
Out[86]: <Axes: xlabel='FLAG_DOCUMENT_7', ylabel='Density'>
```



Plot of FLAG_DOCUMENT_8

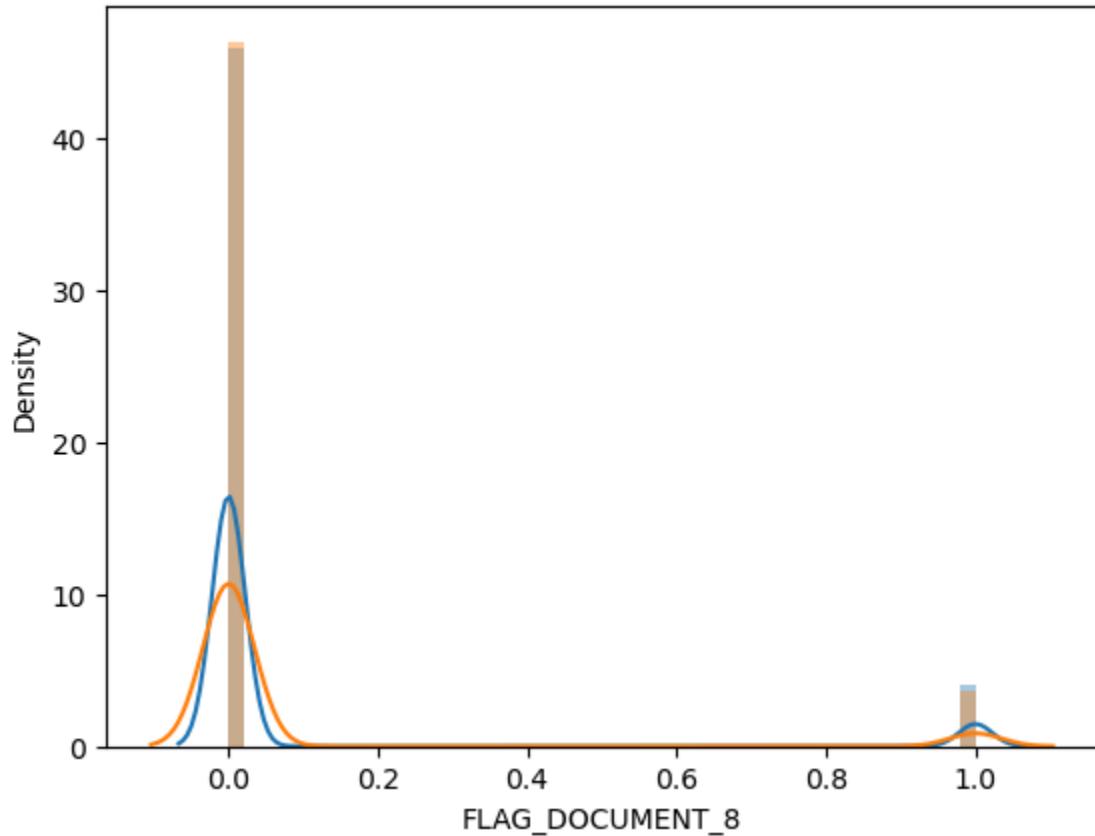
```
Out[86]: (array([259498.,      0.,      0.,      0.,      0.,      0.,      0.,
   23188.]),
 array([0.     , 0.125, 0.25 , 0.375, 0.5   , 0.625, 0.75 , 0.875, 1.     ]),
 <BarContainer object of 8 artists>)

Out[86]: (array([22989.,      0.,      0.,      0.,      0.,      0.,      0.,
  1836.]),
 array([0.     , 0.125, 0.25 , 0.375, 0.5   , 0.625, 0.75 , 0.875, 1.     ]),
 <BarContainer object of 8 artists>)
```



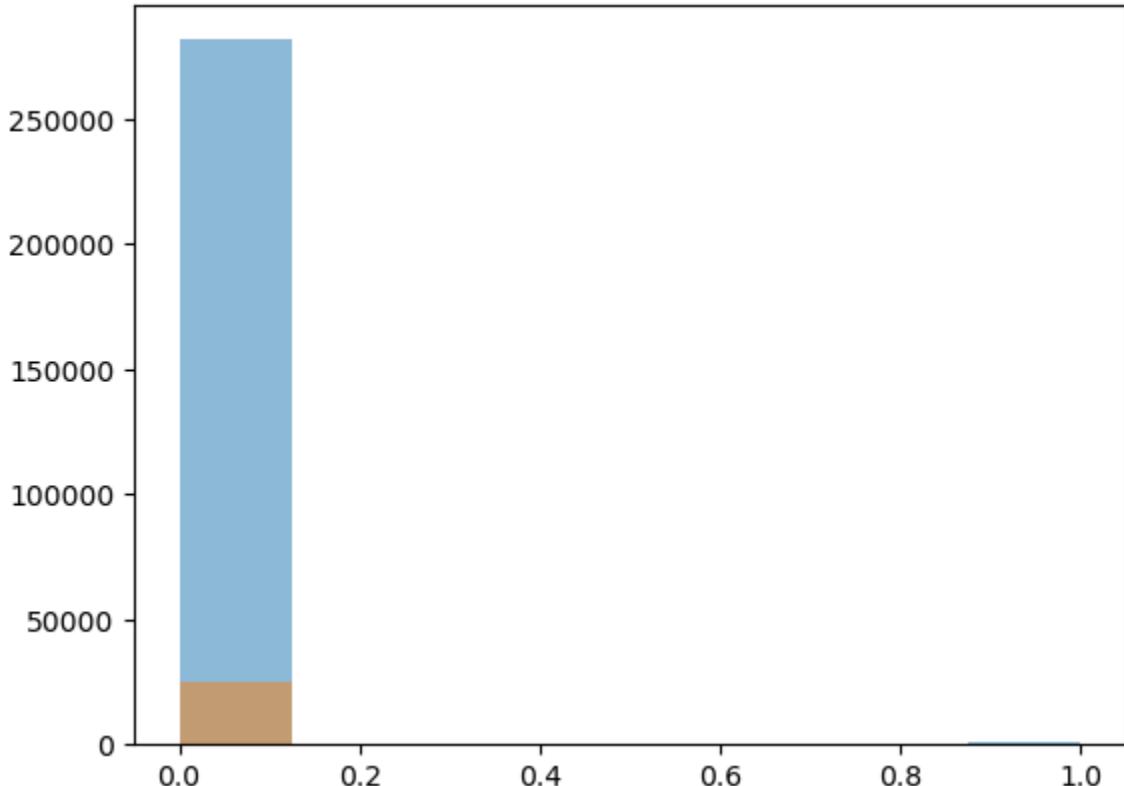
```
Out[86]: <Axes: xlabel='FLAG_DOCUMENT_8', ylabel='Density'>
```

```
Out[86]: <Axes: xlabel='FLAG_DOCUMENT_8', ylabel='Density'>
```



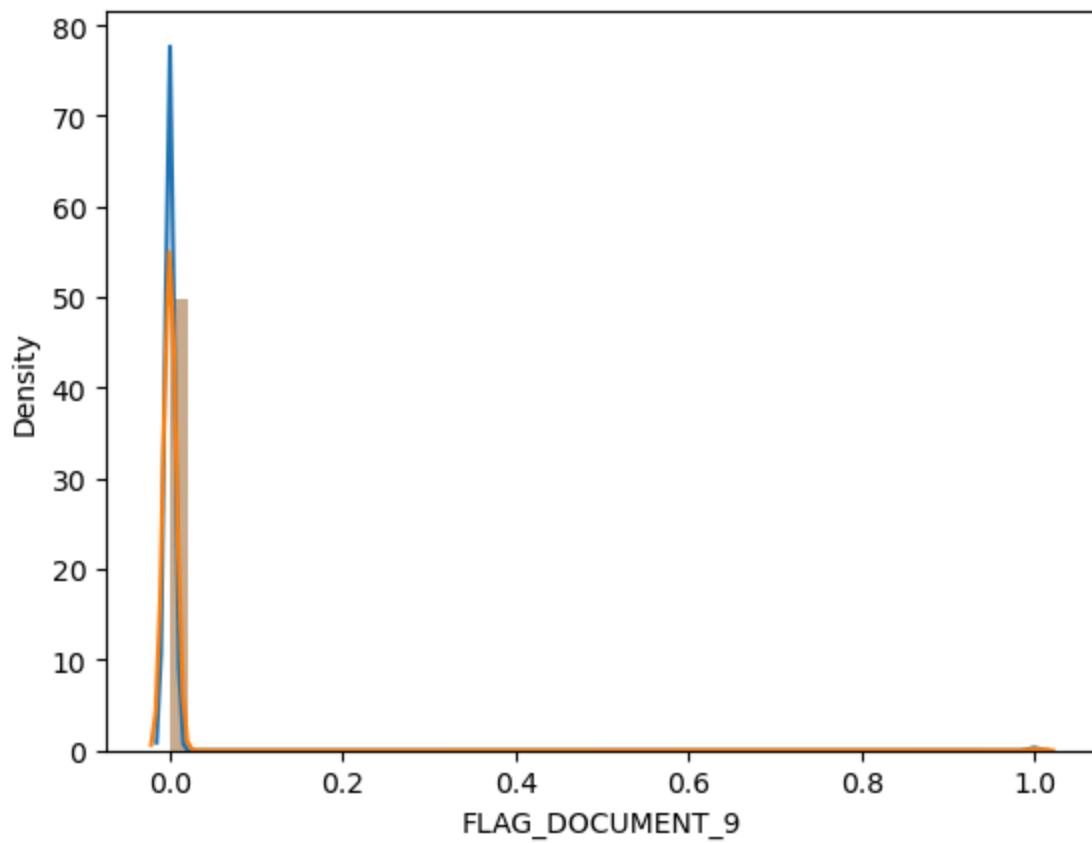
Plot of FLAG_DOCUMENT_9

```
Out[86]: (array([281562.,      0.,      0.,      0.,      0.,      0.,      0.,
   1124.]),
 array([0.    , 0.125, 0.25 , 0.375, 0.5  , 0.625, 0.75 , 0.875, 1.    ]),
 <BarContainer object of 8 artists>
Out[86]: (array([24751.,      0.,      0.,      0.,      0.,      0.,      0.,
  74.]),
 array([0.    , 0.125, 0.25 , 0.375, 0.5  , 0.625, 0.75 , 0.875, 1.    ]),
 <BarContainer object of 8 artists>)
```



```
Out[86]: <Axes: xlabel='FLAG_DOCUMENT_9', ylabel='Density'>
```

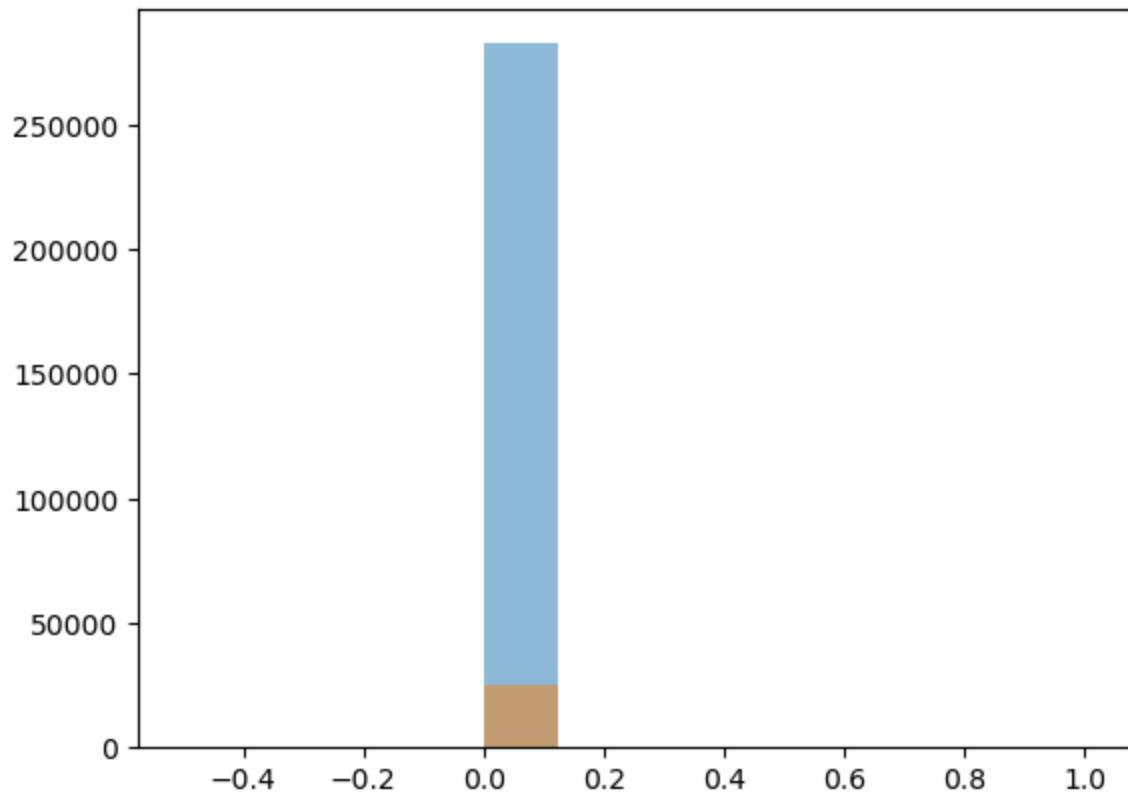
```
Out[86]: <Axes: xlabel='FLAG_DOCUMENT_9', ylabel='Density'>
```



Plot of `FLAG_DOCUMENT_10`

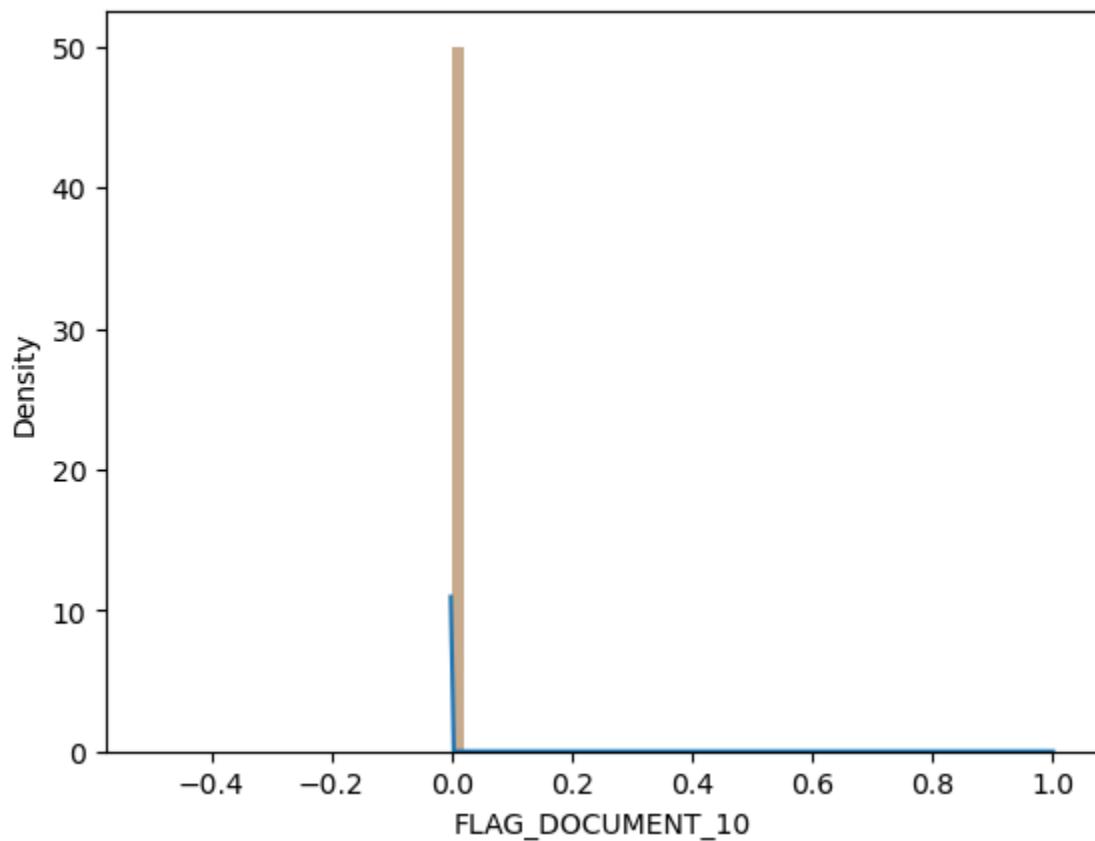
```
Out[86]: (array([2.82679e+05, 0.00000e+00, 0.00000e+00, 0.00000e+00, 0.00000e+00,
       0.00000e+00, 0.00000e+00, 7.00000e+00]),
       array([0.    , 0.125, 0.25 , 0.375, 0.5   , 0.625, 0.75 , 0.875, 1.    ]),
       <BarContainer object of 8 artists>)

Out[86]: (array([ 0.,  0.,  0.,  0., 24825.,  0.,  0.,  0.]),
       array([-0.5  , -0.375, -0.25 , -0.125,  0.    ,  0.125,  0.25 ,  0.375,
              0.5  ]),
       <BarContainer object of 8 artists>)
```



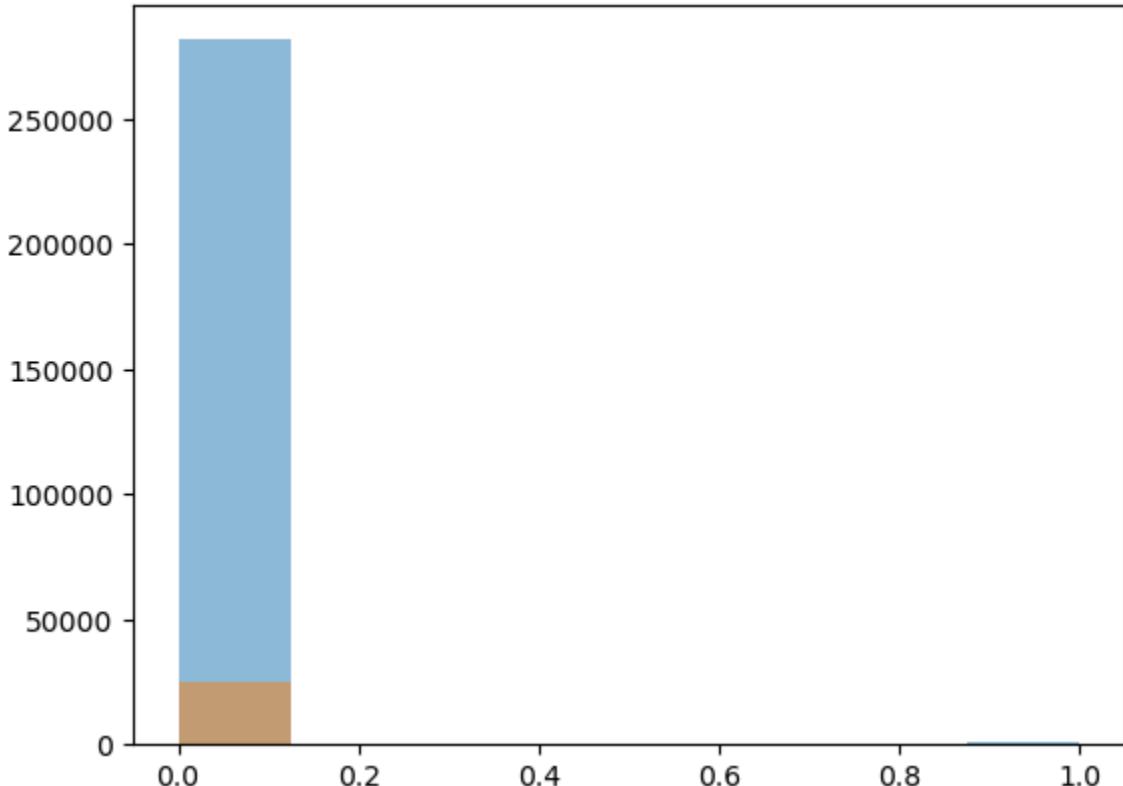
```
Out[86]: <Axes: xlabel='FLAG_DOCUMENT_10', ylabel='Density'>
```

```
Out[86]: <Axes: xlabel='FLAG_DOCUMENT_10', ylabel='Density'>
```



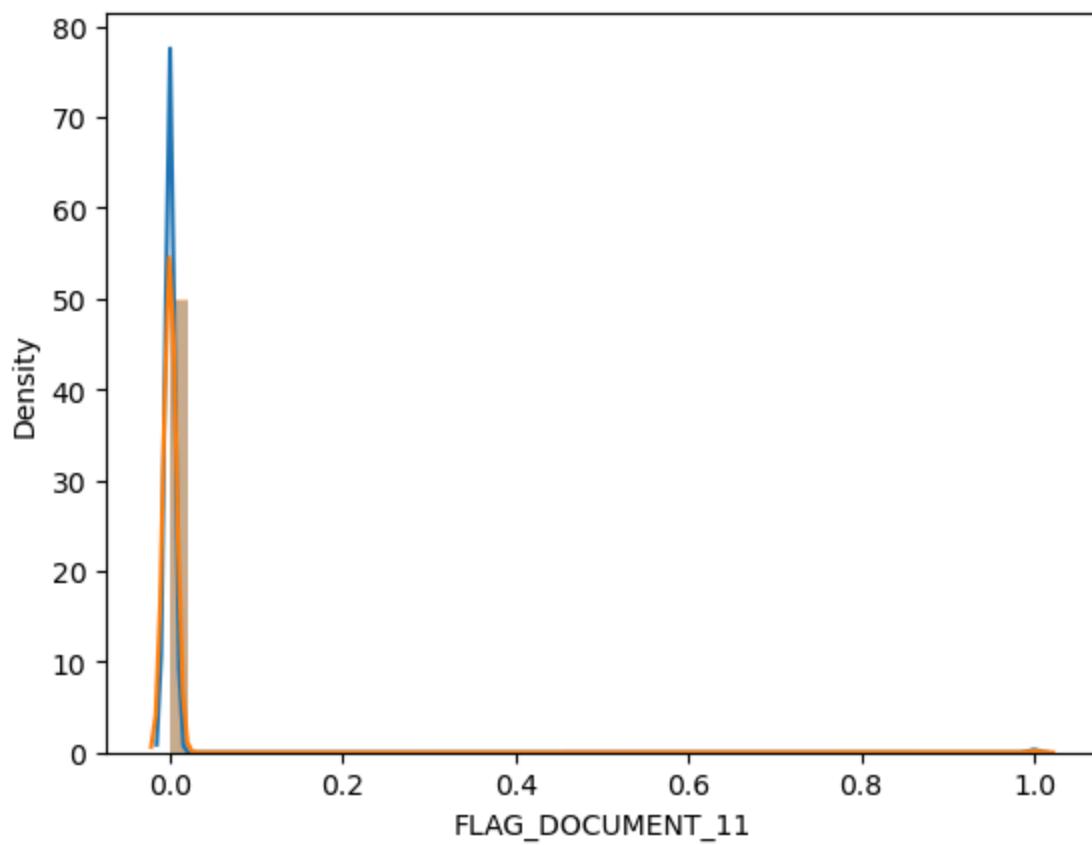
Plot of FLAG_DOCUMENT_11

```
Out[86]: (array([281558.,      0.,      0.,      0.,      0.,      0.,
       1128.]),
 array([0.    , 0.125, 0.25 , 0.375, 0.5   , 0.625, 0.75 , 0.875, 1.    ]),
 <BarContainer object of 8 artists>
Out[86]: (array([24750.,      0.,      0.,      0.,      0.,      0.,
       0.,      75.]),
 array([0.    , 0.125, 0.25 , 0.375, 0.5   , 0.625, 0.75 , 0.875, 1.    ]),
 <BarContainer object of 8 artists>)
```



```
Out[86]: <Axes: xlabel='FLAG_DOCUMENT_11', ylabel='Density'>
```

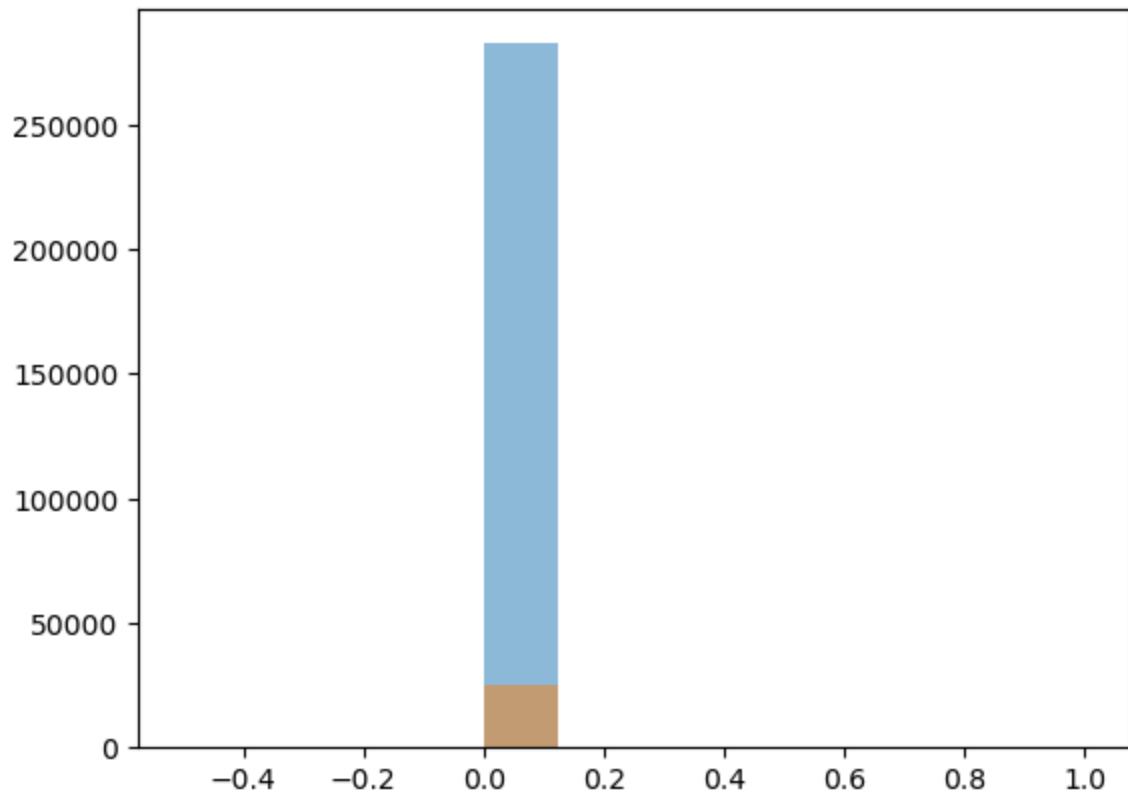
```
Out[86]: <Axes: xlabel='FLAG_DOCUMENT_11', ylabel='Density'>
```



Plot of `FLAG_DOCUMENT_12`

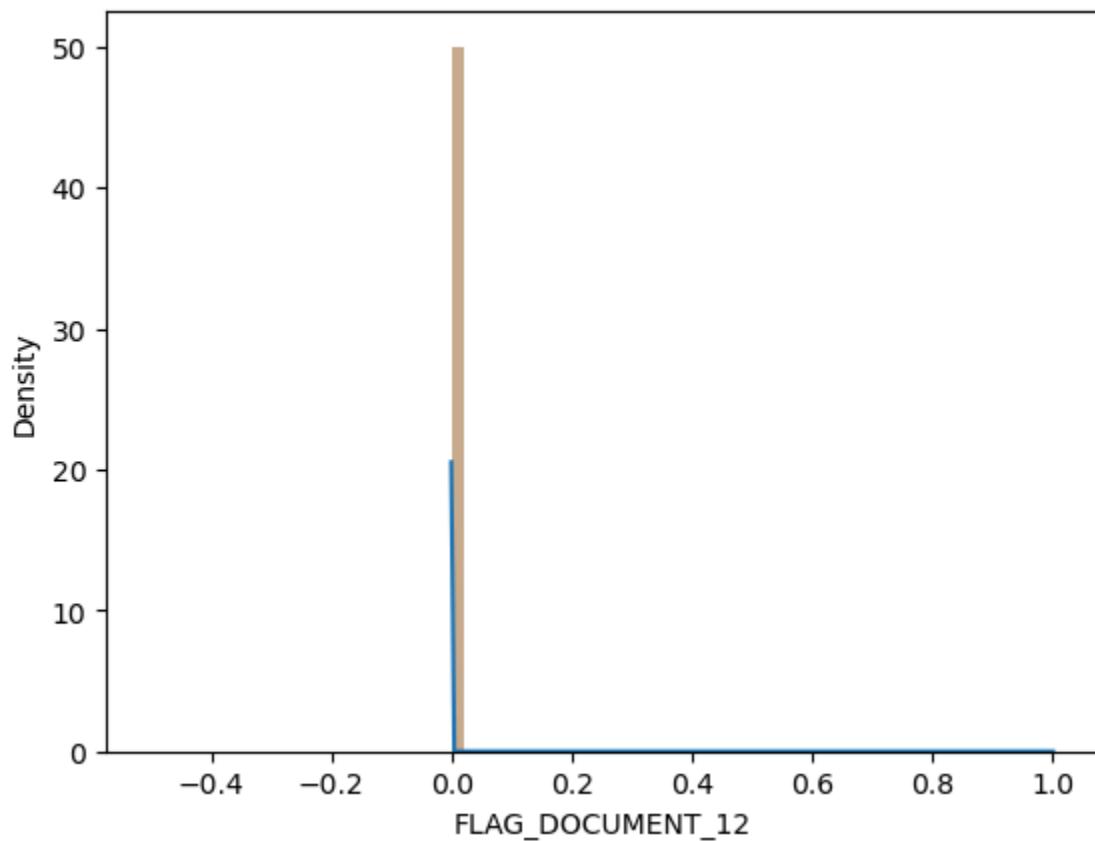
```
Out[86]: (array([2.82684e+05, 0.00000e+00, 0.00000e+00, 0.00000e+00, 0.00000e+00,
       0.00000e+00, 0.00000e+00, 2.00000e+00]),
       array([0.    , 0.125, 0.25 , 0.375, 0.5   , 0.625, 0.75 , 0.875, 1.    ]),
       <BarContainer object of 8 artists>)

Out[86]: (array([ 0.,  0.,  0.,  0., 24825.,  0.,  0.,  0.]),
       array([-0.5  , -0.375, -0.25 , -0.125,  0.    ,  0.125,  0.25 ,  0.375,
              0.5  ]),
       <BarContainer object of 8 artists>)
```



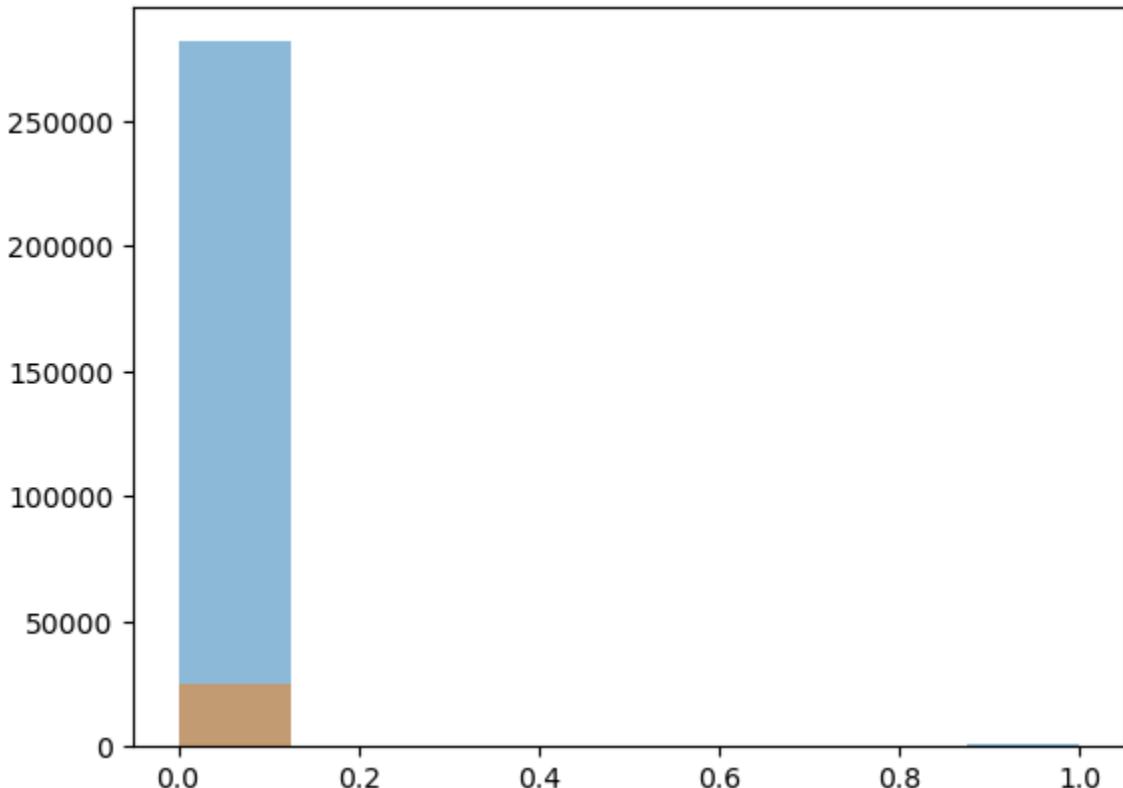
```
Out[86]: <Axes: xlabel='FLAG_DOCUMENT_12', ylabel='Density'>
```

```
Out[86]: <Axes: xlabel='FLAG_DOCUMENT_12', ylabel='Density'>
```



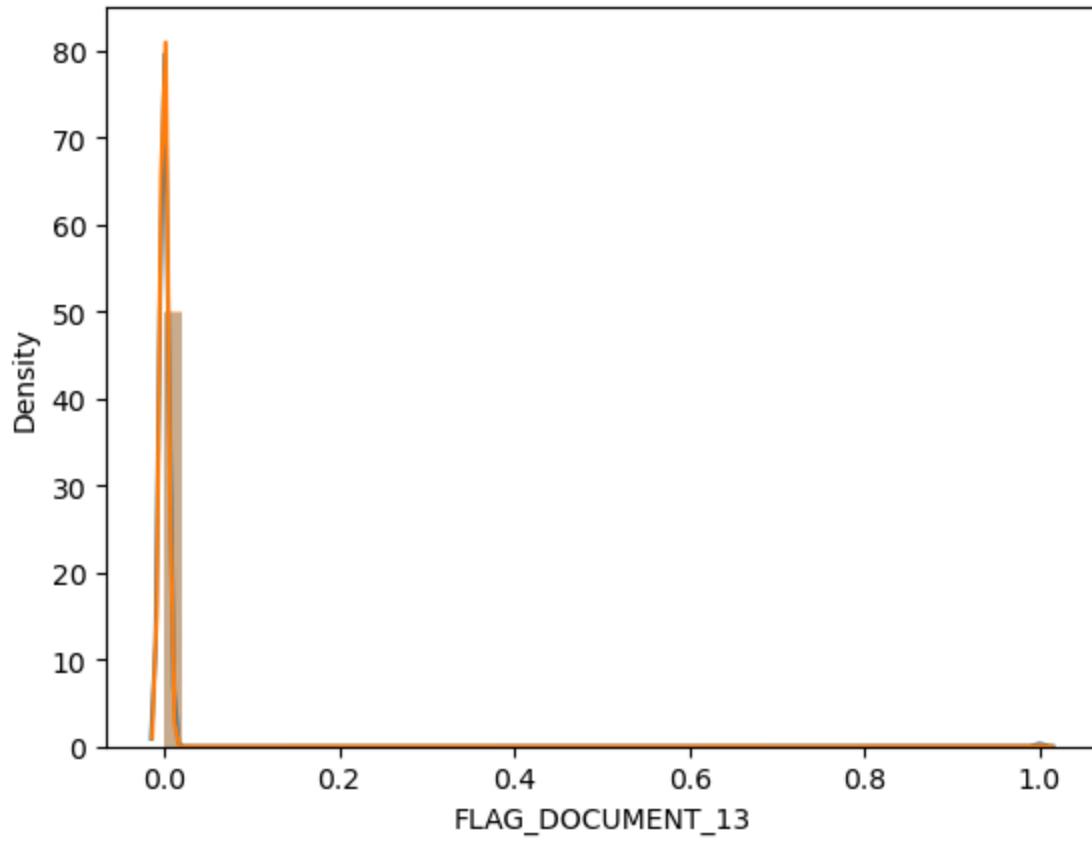
Plot of FLAG_DOCUMENT_13

```
Out[86]: (array([281632.,      0.,      0.,      0.,      0.,      0.,      0.,
   1054.]),
 array([0.    , 0.125, 0.25 , 0.375, 0.5  , 0.625, 0.75 , 0.875, 1.    ]),
 <BarContainer object of 8 artists>
Out[86]: (array([24795.,      0.,      0.,      0.,      0.,      0.,      0.,      30.]),
 array([0.    , 0.125, 0.25 , 0.375, 0.5  , 0.625, 0.75 , 0.875, 1.    ]),
 <BarContainer object of 8 artists>)
```



```
Out[86]: <Axes: xlabel='FLAG_DOCUMENT_13', ylabel='Density'>
```

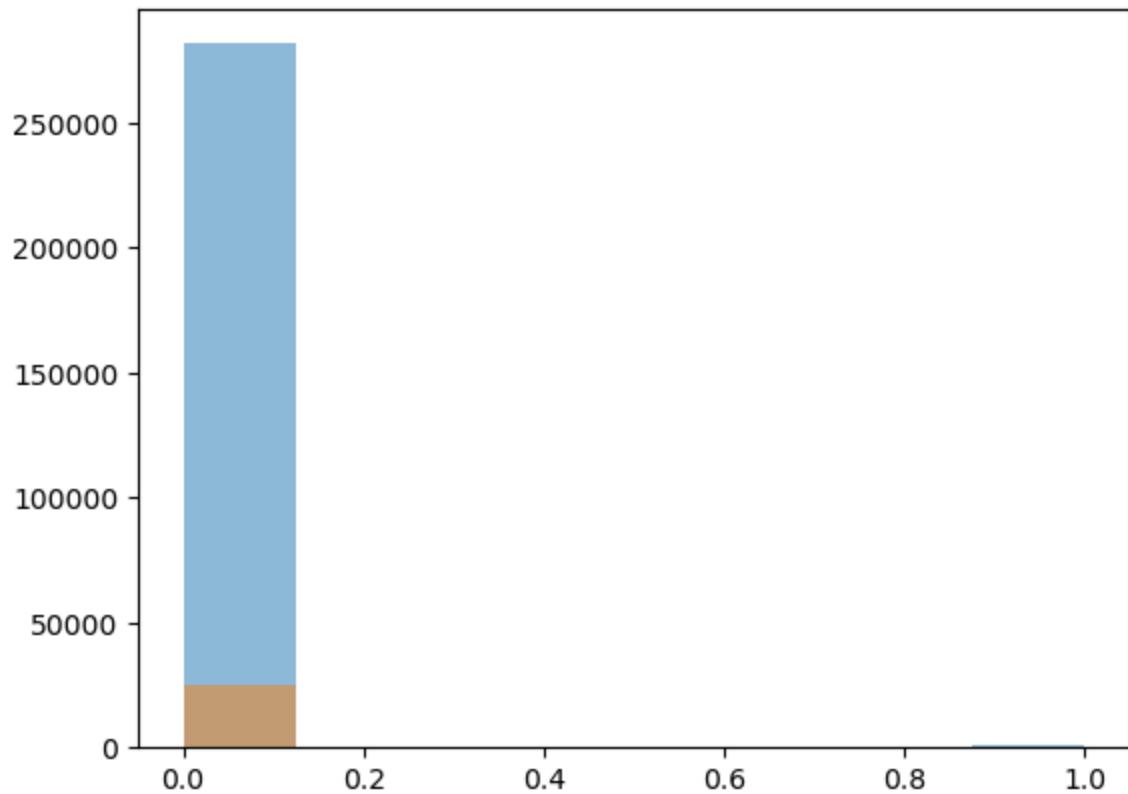
```
Out[86]: <Axes: xlabel='FLAG_DOCUMENT_13', ylabel='Density'>
```



Plot of `FLAG_DOCUMENT_14`

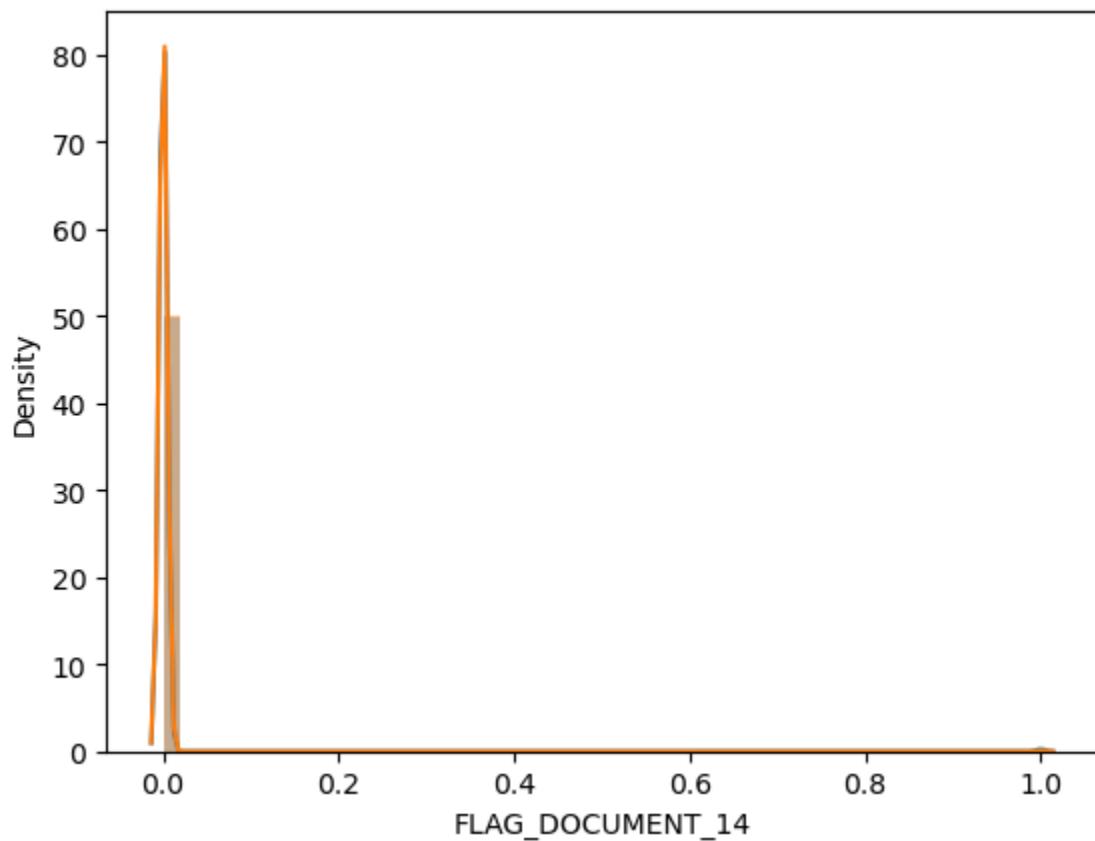
```
Out[86]: (array([281813.,      0.,      0.,      0.,      0.,      0.,      0.,      0.,
   873.]),
 array([0.    , 0.125, 0.25 , 0.375, 0.5   , 0.625, 0.75 , 0.875, 1.    ]),
 <BarContainer object of 8 artists>)

Out[86]: (array([24795.,      0.,      0.,      0.,      0.,      0.,      0.,      30.]),
 array([0.    , 0.125, 0.25 , 0.375, 0.5   , 0.625, 0.75 , 0.875, 1.    ]),
 <BarContainer object of 8 artists>)
```



```
Out[86]: <Axes: xlabel='FLAG_DOCUMENT_14', ylabel='Density'>
```

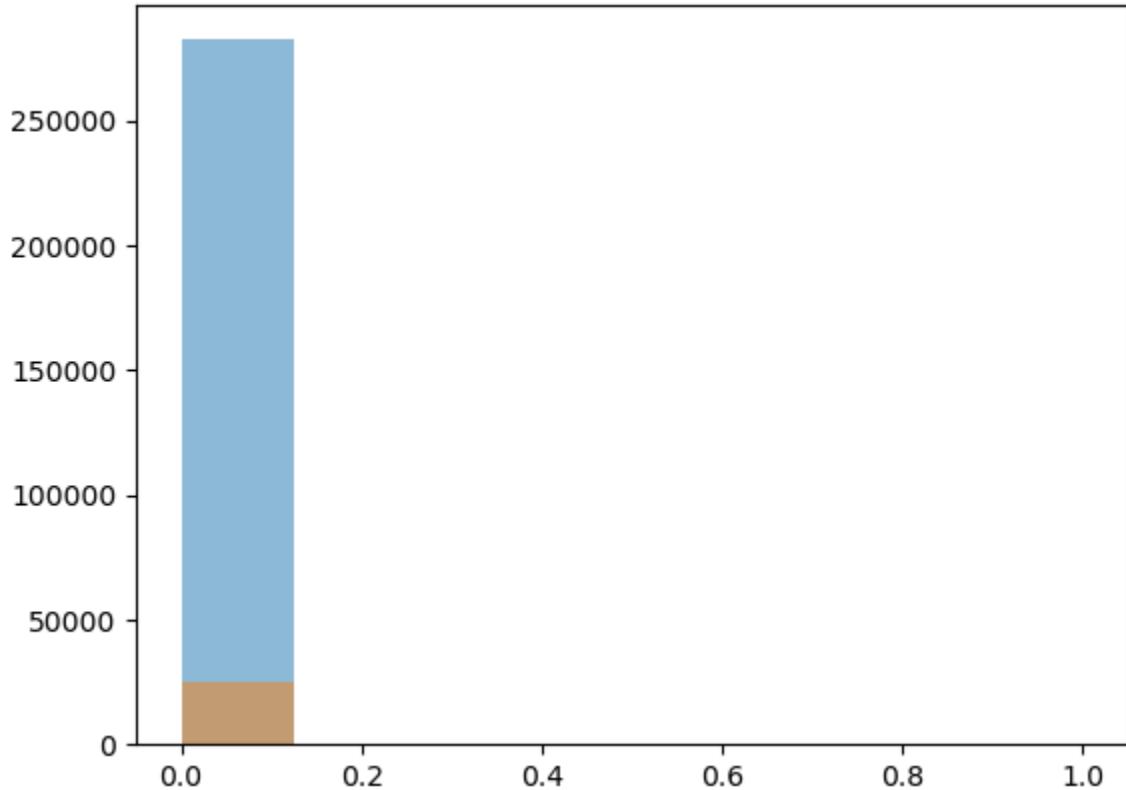
```
Out[86]: <Axes: xlabel='FLAG_DOCUMENT_14', ylabel='Density'>
```



Plot of FLAG_DOCUMENT_15

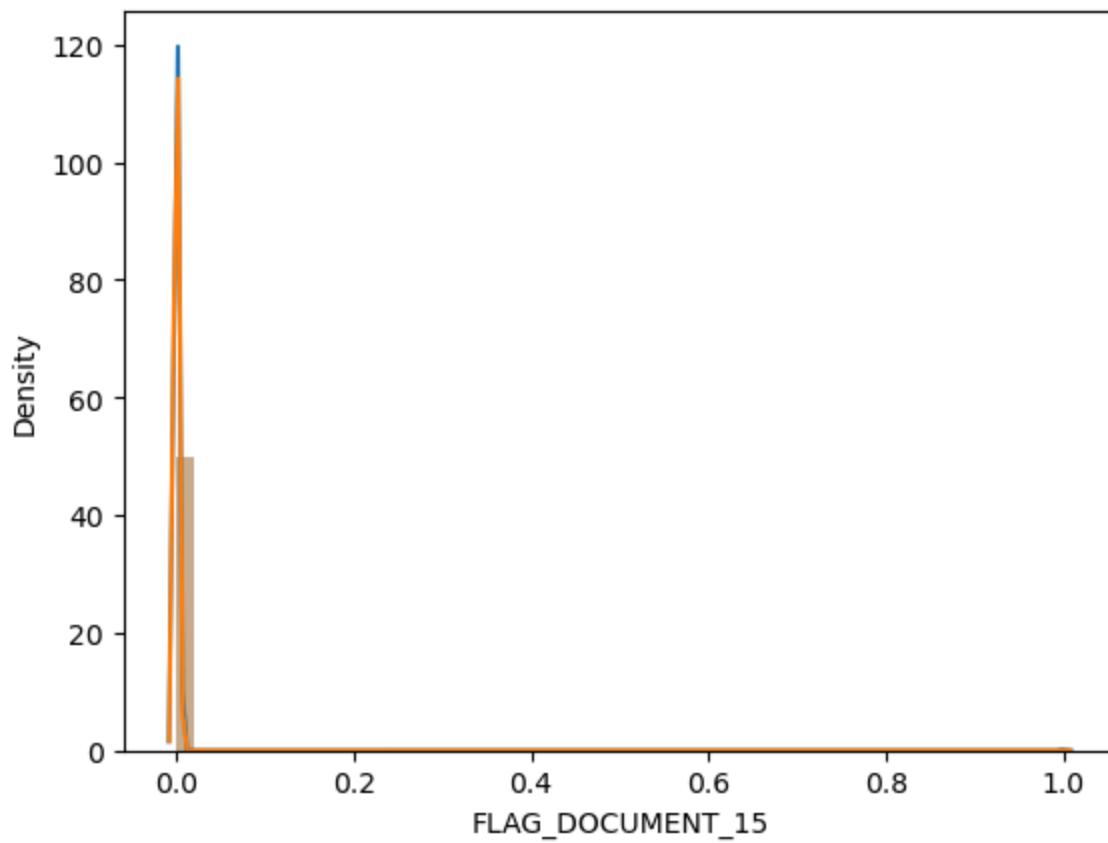
```
Out[86]: (array([282325.,      0.,      0.,      0.,      0.,      0.,      0.,
   361.]),
 array([0.    , 0.125, 0.25 , 0.375, 0.5   , 0.625, 0.75 , 0.875, 1.    ]),
 <BarContainer object of 8 artists>)

Out[86]: (array([2.4814e+04, 0.0000e+00, 0.0000e+00, 0.0000e+00, 0.0000e+00,
 0.0000e+00, 0.0000e+00, 1.1000e+01]),
 array([0.    , 0.125, 0.25 , 0.375, 0.5   , 0.625, 0.75 , 0.875, 1.    ]),
 <BarContainer object of 8 artists>)
```



```
Out[86]: <Axes: xlabel='FLAG_DOCUMENT_15', ylabel='Density'>
```

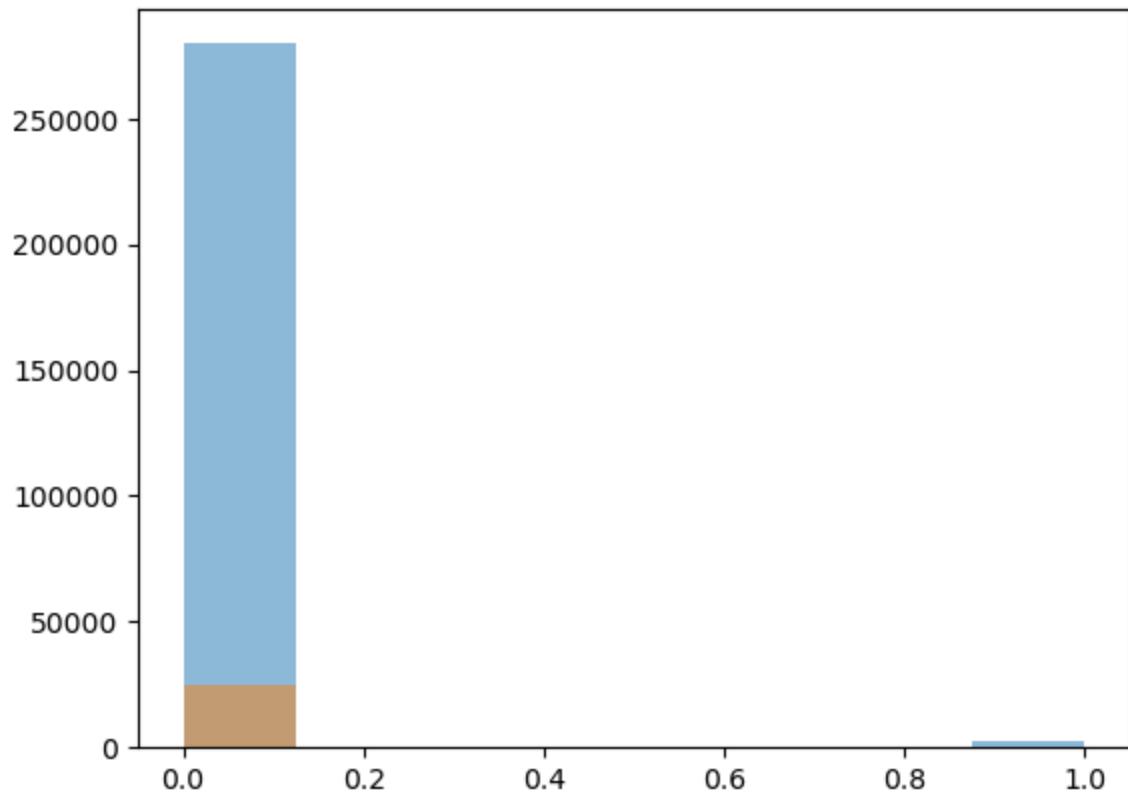
```
Out[86]: <Axes: xlabel='FLAG_DOCUMENT_15', ylabel='Density'>
```



Plot of FLAG_DOCUMENT_16

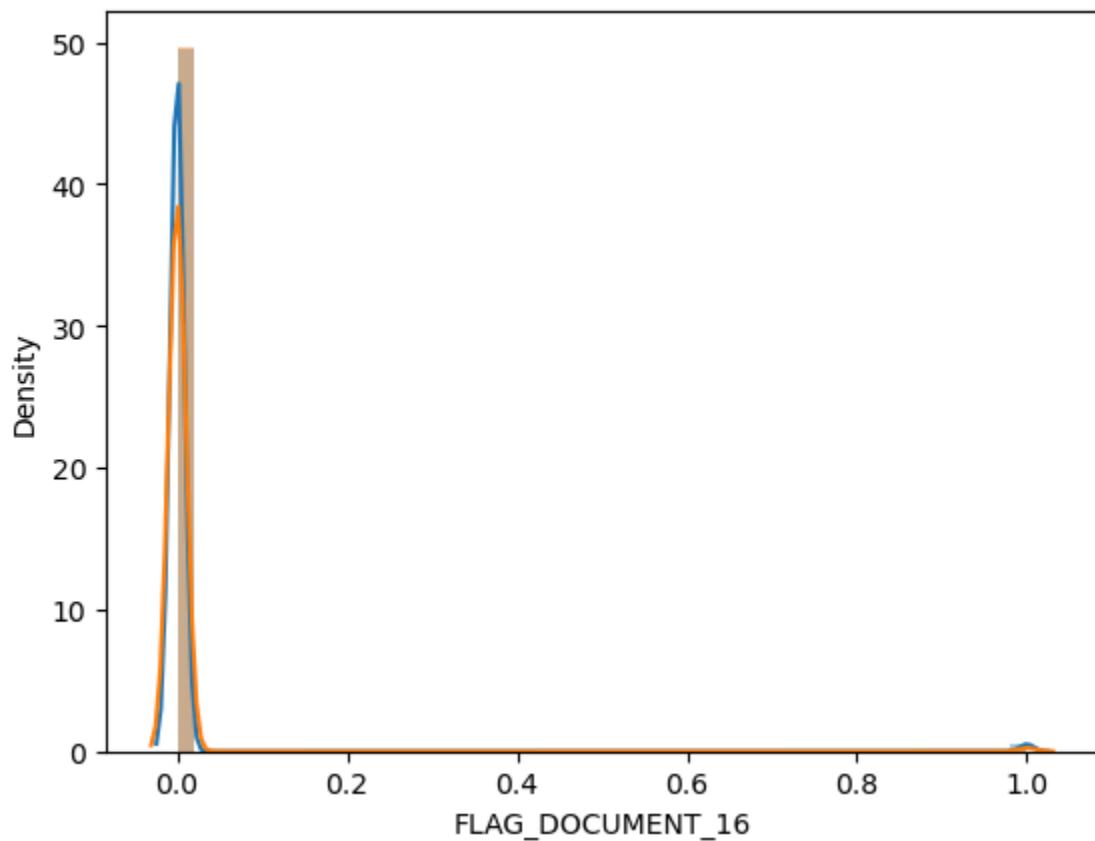
```
Out[86]: (array([279783.,      0.,      0.,      0.,      0.,      0.,      0.,      0.,
   2903.]),
 array([0.     , 0.125, 0.25 , 0.375, 0.5   , 0.625, 0.75 , 0.875, 1.     ]),
 <BarContainer object of 8 artists>)

Out[86]: (array([24675.,      0.,      0.,      0.,      0.,      0.,      0., 150.]),
 array([0.     , 0.125, 0.25 , 0.375, 0.5   , 0.625, 0.75 , 0.875, 1.     ]),
 <BarContainer object of 8 artists>)
```



```
Out[86]: <Axes: xlabel='FLAG_DOCUMENT_16', ylabel='Density'>
```

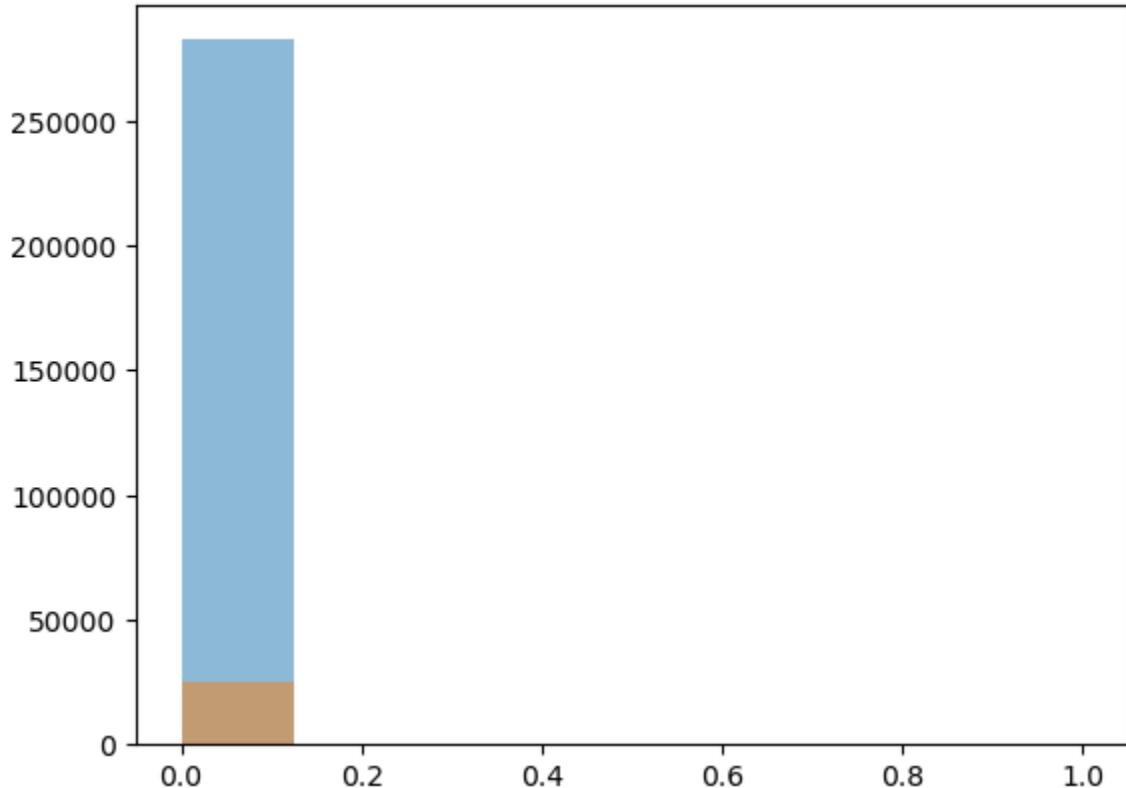
```
Out[86]: <Axes: xlabel='FLAG_DOCUMENT_16', ylabel='Density'>
```



Plot of FLAG_DOCUMENT_17

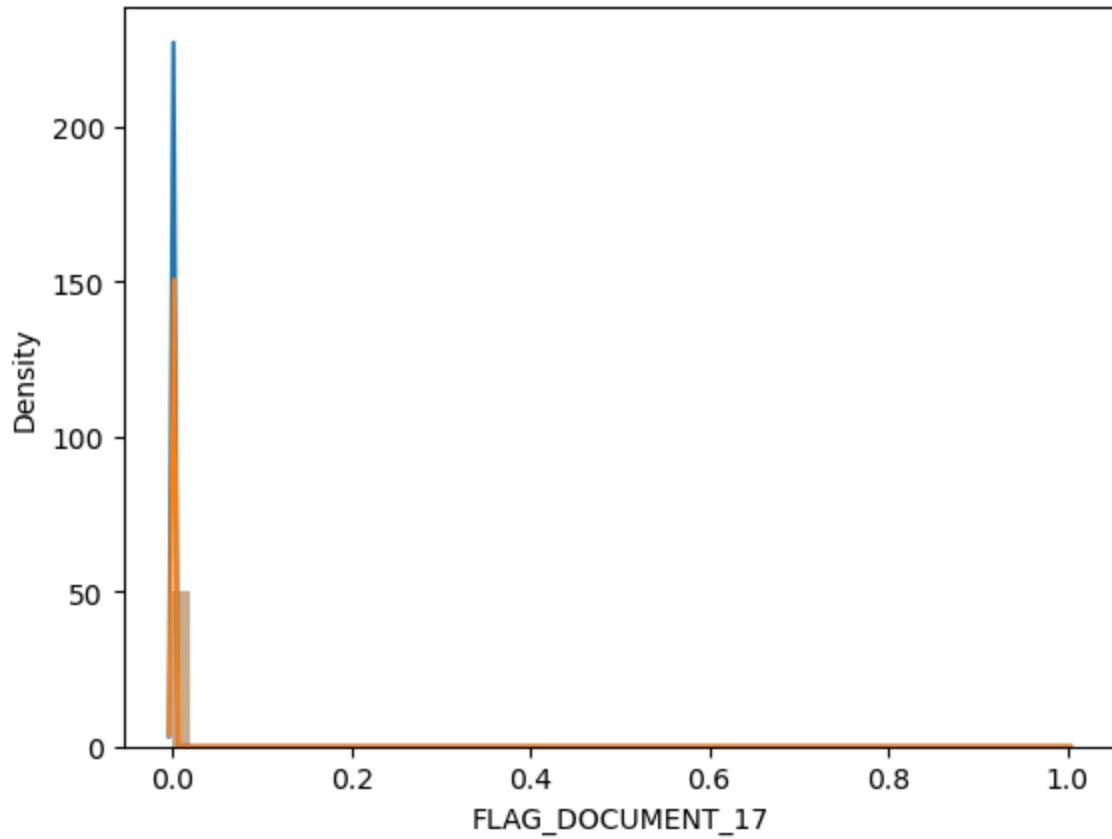
```
Out[86]: (array([2.82606e+05, 0.00000e+00, 0.00000e+00, 0.00000e+00, 0.00000e+00,
       0.00000e+00, 0.00000e+00, 8.00000e+01]),
       array([0.    , 0.125, 0.25 , 0.375, 0.5   , 0.625, 0.75 , 0.875, 1.    ]),
       <BarContainer object of 8 artists>)

Out[86]: (array([2.4823e+04, 0.0000e+00, 0.0000e+00, 0.0000e+00, 0.0000e+00,
       0.0000e+00, 0.0000e+00, 2.0000e+00]),
       array([0.    , 0.125, 0.25 , 0.375, 0.5   , 0.625, 0.75 , 0.875, 1.    ]),
       <BarContainer object of 8 artists>)
```



```
Out[86]: <Axes: xlabel='FLAG_DOCUMENT_17', ylabel='Density'>
```

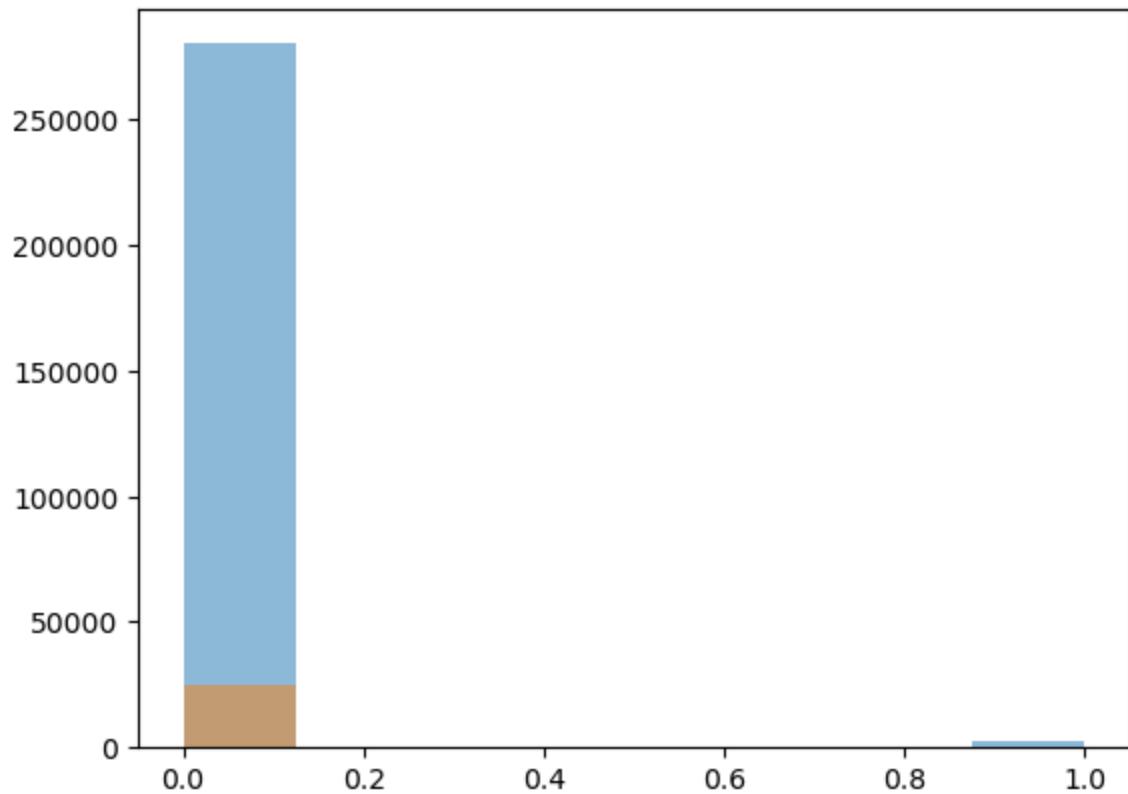
```
Out[86]: <Axes: xlabel='FLAG_DOCUMENT_17', ylabel='Density'>
```



Plot of `FLAG_DOCUMENT_18`

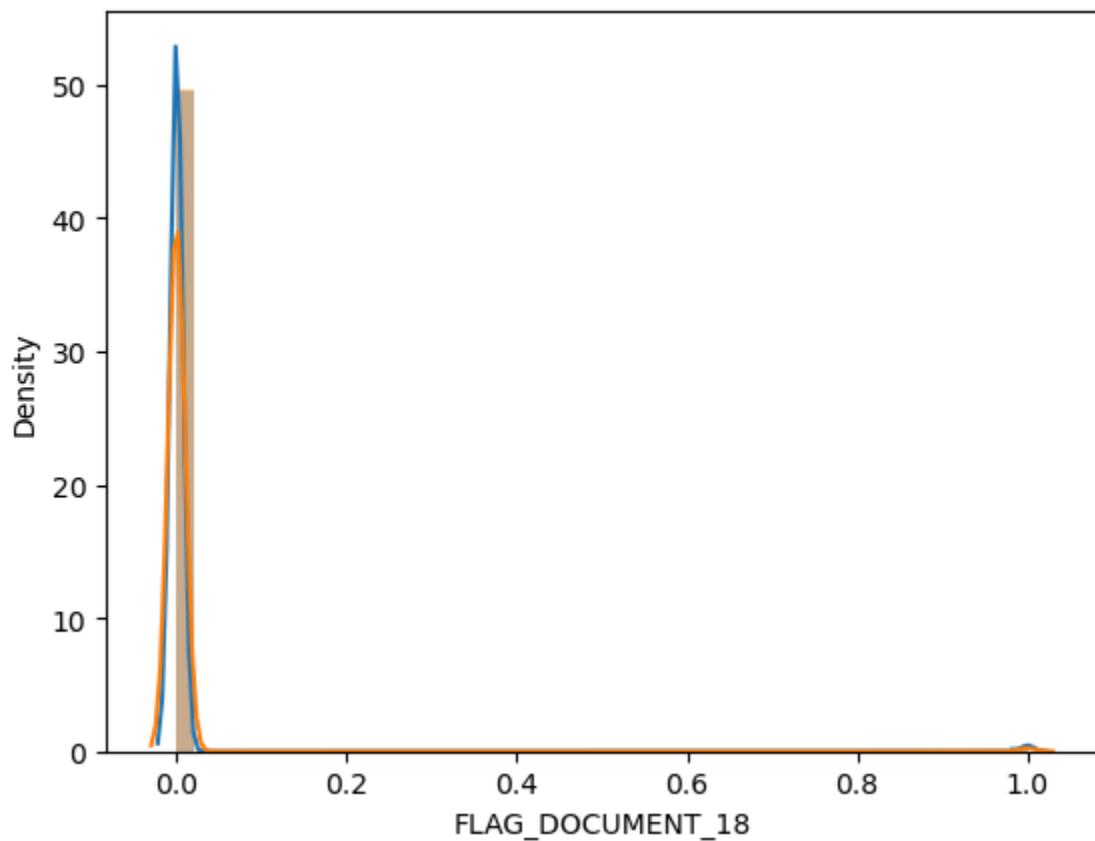
```
Out[86]: (array([280328.,      0.,      0.,      0.,      0.,      0.,      0.,
   2358.]),
 array([0.     , 0.125, 0.25 , 0.375, 0.5   , 0.625, 0.75 , 0.875, 1.     ]),
 <BarContainer object of 8 artists>)

Out[86]: (array([24683.,      0.,      0.,      0.,      0.,      0.,      0.,
  142.]),
 array([0.     , 0.125, 0.25 , 0.375, 0.5   , 0.625, 0.75 , 0.875, 1.     ]),
 <BarContainer object of 8 artists>)
```



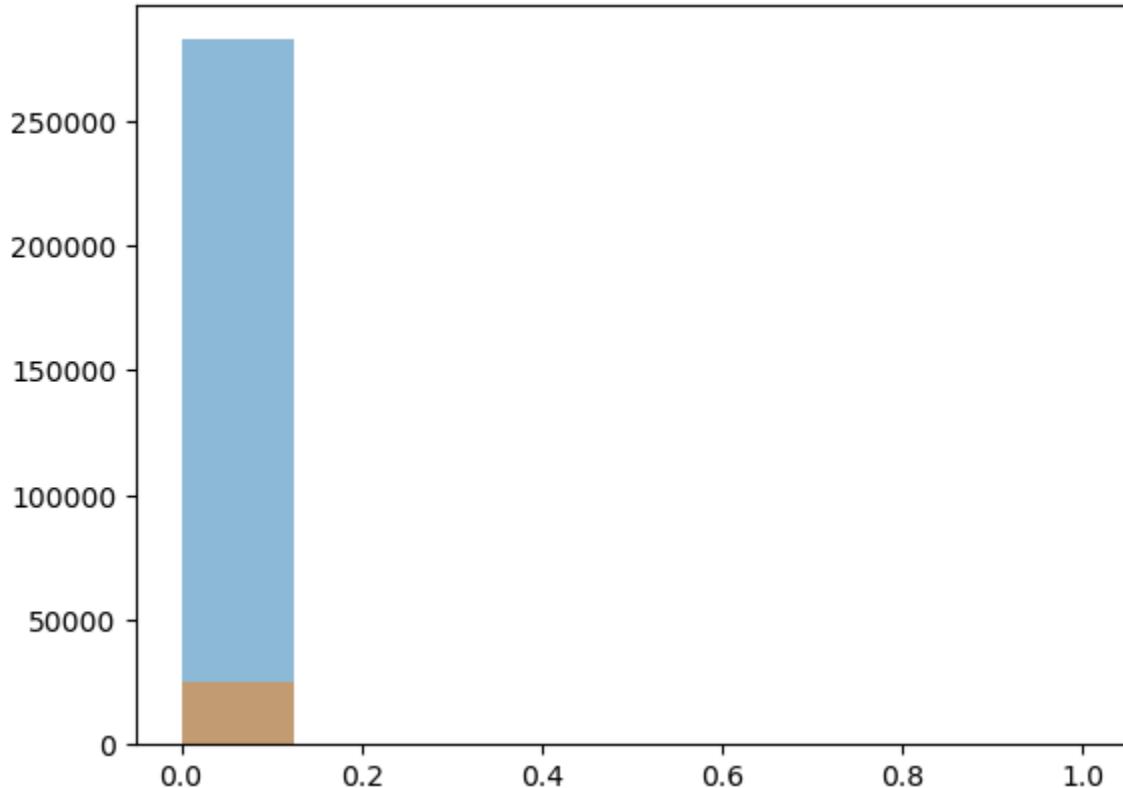
```
Out[86]: <Axes: xlabel='FLAG_DOCUMENT_18', ylabel='Density'>
```

```
Out[86]: <Axes: xlabel='FLAG_DOCUMENT_18', ylabel='Density'>
```



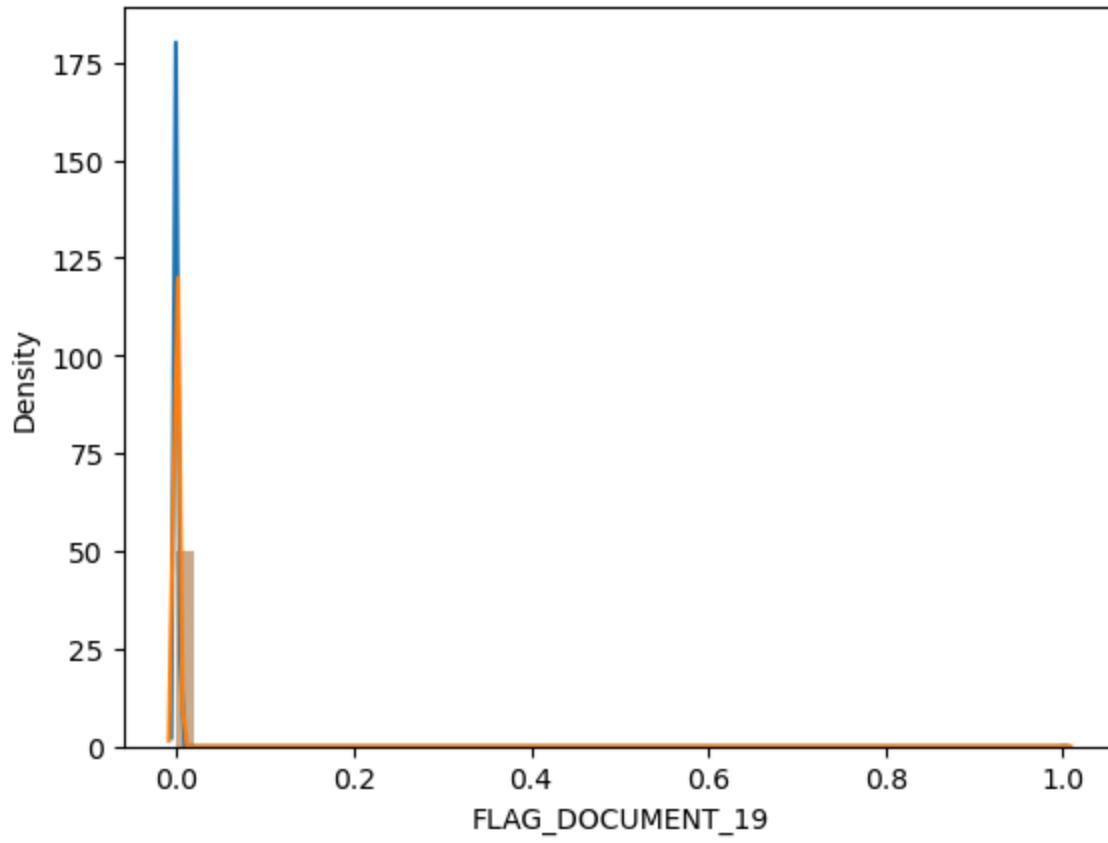
Plot of FLAG_DOCUMENT_19

```
Out[86]: (array([2.82515e+05, 0.00000e+00, 0.00000e+00, 0.00000e+00, 0.00000e+00,
       0.00000e+00, 0.00000e+00, 1.71000e+02]),  
 array([0.    , 0.125, 0.25 , 0.375, 0.5   , 0.625, 0.75 , 0.875, 1.    ]),  
 <BarContainer object of 8 artists>)  
  
Out[86]: (array([2.4813e+04, 0.0000e+00, 0.0000e+00, 0.0000e+00, 0.0000e+00,  
       0.0000e+00, 0.0000e+00, 1.2000e+01]),  
 array([0.    , 0.125, 0.25 , 0.375, 0.5   , 0.625, 0.75 , 0.875, 1.    ]),  
 <BarContainer object of 8 artists>)
```



```
Out[86]: <Axes: xlabel='FLAG_DOCUMENT_19', ylabel='Density'>
```

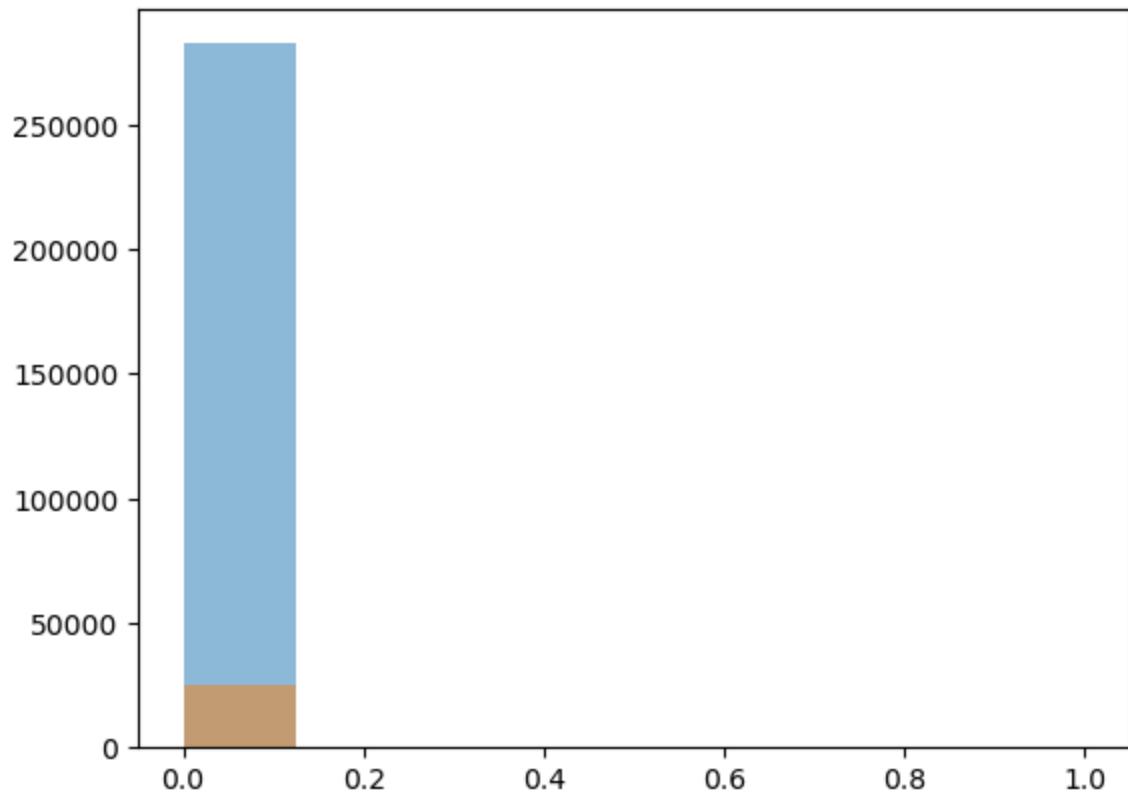
```
Out[86]: <Axes: xlabel='FLAG_DOCUMENT_19', ylabel='Density'>
```



Plot of `FLAG_DOCUMENT_20`

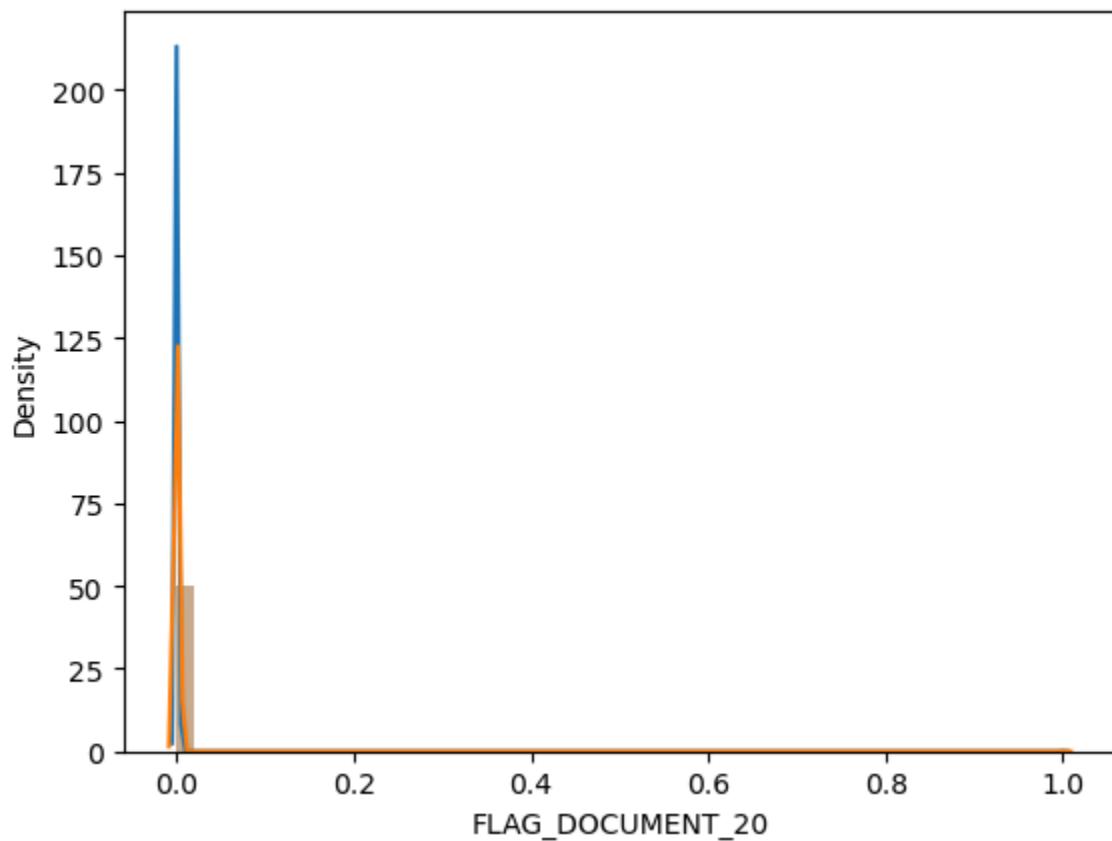
```
Out[86]: (array([2.82543e+05, 0.00000e+00, 0.00000e+00, 0.00000e+00, 0.00000e+00,
       0.00000e+00, 0.00000e+00, 1.43000e+02]),
       array([0.    , 0.125, 0.25 , 0.375, 0.5   , 0.625, 0.75 , 0.875, 1.    ]),
       <BarContainer object of 8 artists>)
```

```
Out[86]: (array([2.4812e+04, 0.0000e+00, 0.0000e+00, 0.0000e+00, 0.0000e+00,
       0.0000e+00, 0.0000e+00, 1.3000e+01]),
       array([0.    , 0.125, 0.25 , 0.375, 0.5   , 0.625, 0.75 , 0.875, 1.    ]),
       <BarContainer object of 8 artists>)
```



```
Out[86]: <Axes: xlabel='FLAG_DOCUMENT_20', ylabel='Density'>
```

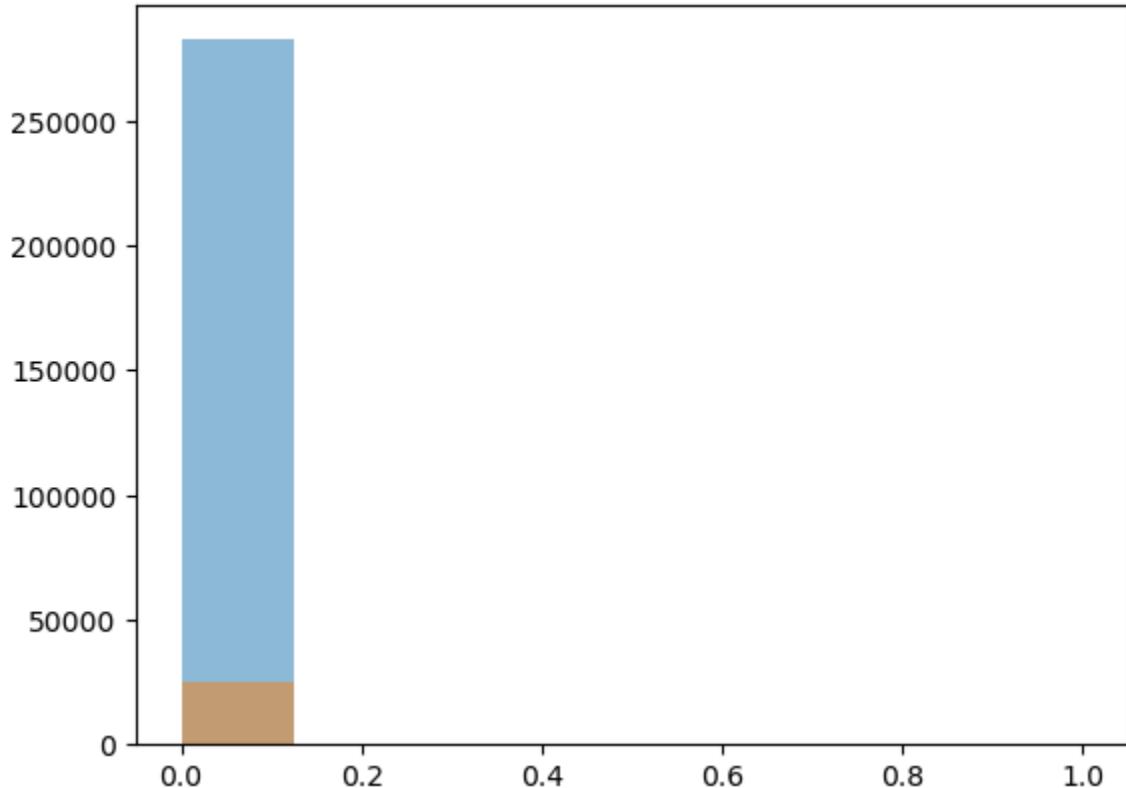
```
Out[86]: <Axes: xlabel='FLAG_DOCUMENT_20', ylabel='Density'>
```



Plot of FLAG_DOCUMENT_21

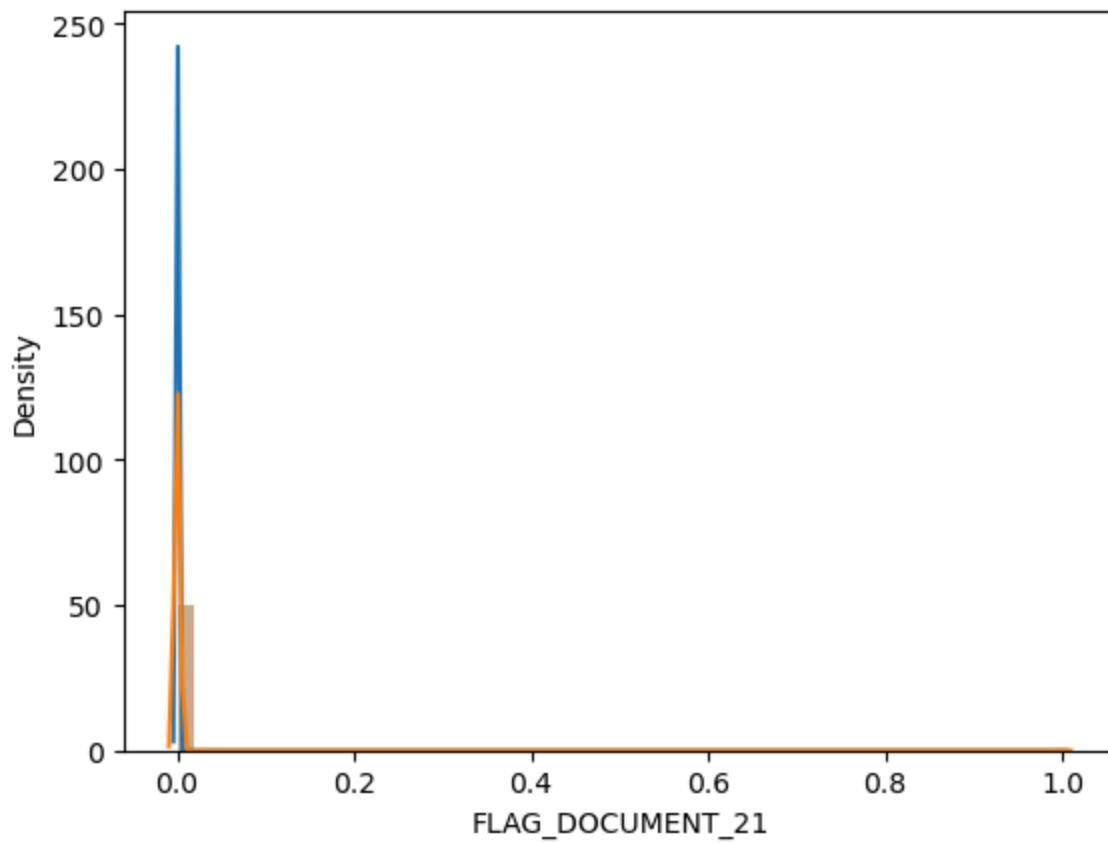
```
Out[86]: (array([2.82597e+05, 0.00000e+00, 0.00000e+00, 0.00000e+00, 0.00000e+00,
       0.00000e+00, 0.00000e+00, 8.90000e+01]),
       array([0.    , 0.125, 0.25 , 0.375, 0.5   , 0.625, 0.75 , 0.875, 1.    ]),
       <BarContainer object of 8 artists>)

Out[86]: (array([2.4811e+04, 0.0000e+00, 0.0000e+00, 0.0000e+00, 0.0000e+00,
       0.0000e+00, 0.0000e+00, 1.4000e+01]),
       array([0.    , 0.125, 0.25 , 0.375, 0.5   , 0.625, 0.75 , 0.875, 1.    ]),
       <BarContainer object of 8 artists>)
```



```
Out[86]: <Axes: xlabel='FLAG_DOCUMENT_21', ylabel='Density'>
```

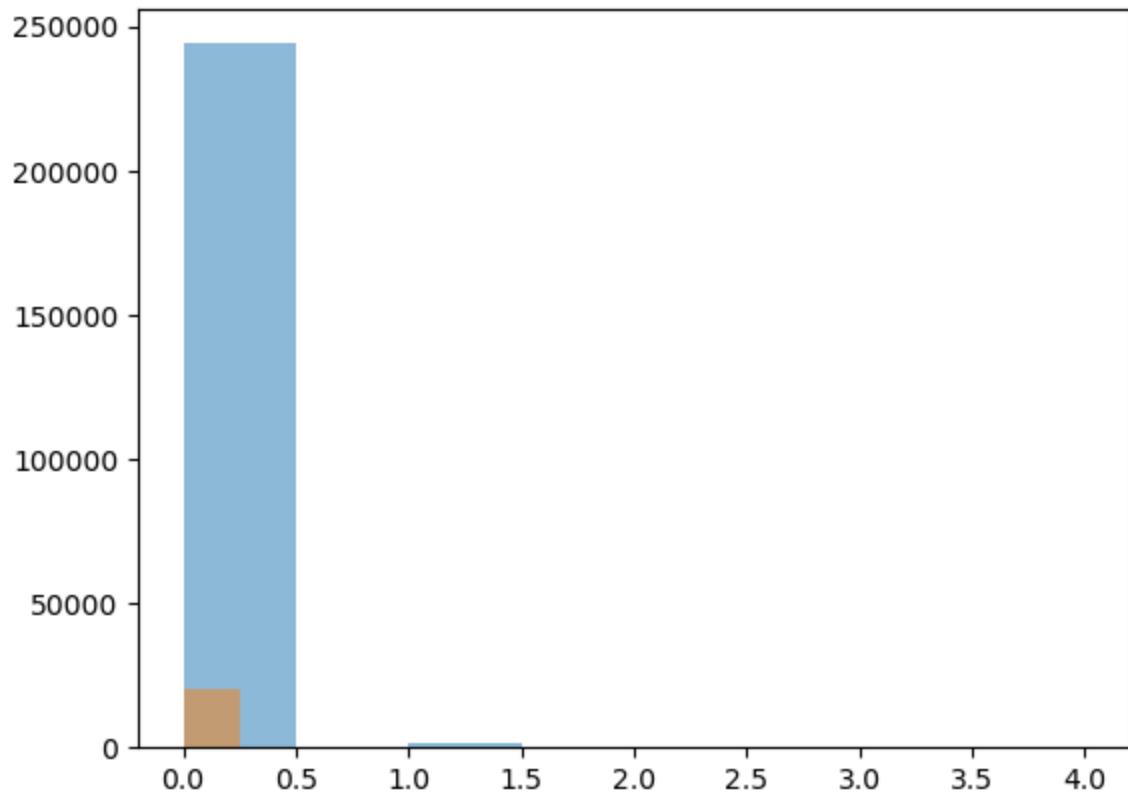
```
Out[86]: <Axes: xlabel='FLAG_DOCUMENT_21', ylabel='Density'>
```



Plot of `AMT_REQ_CREDIT_BUREAU_HOUR`

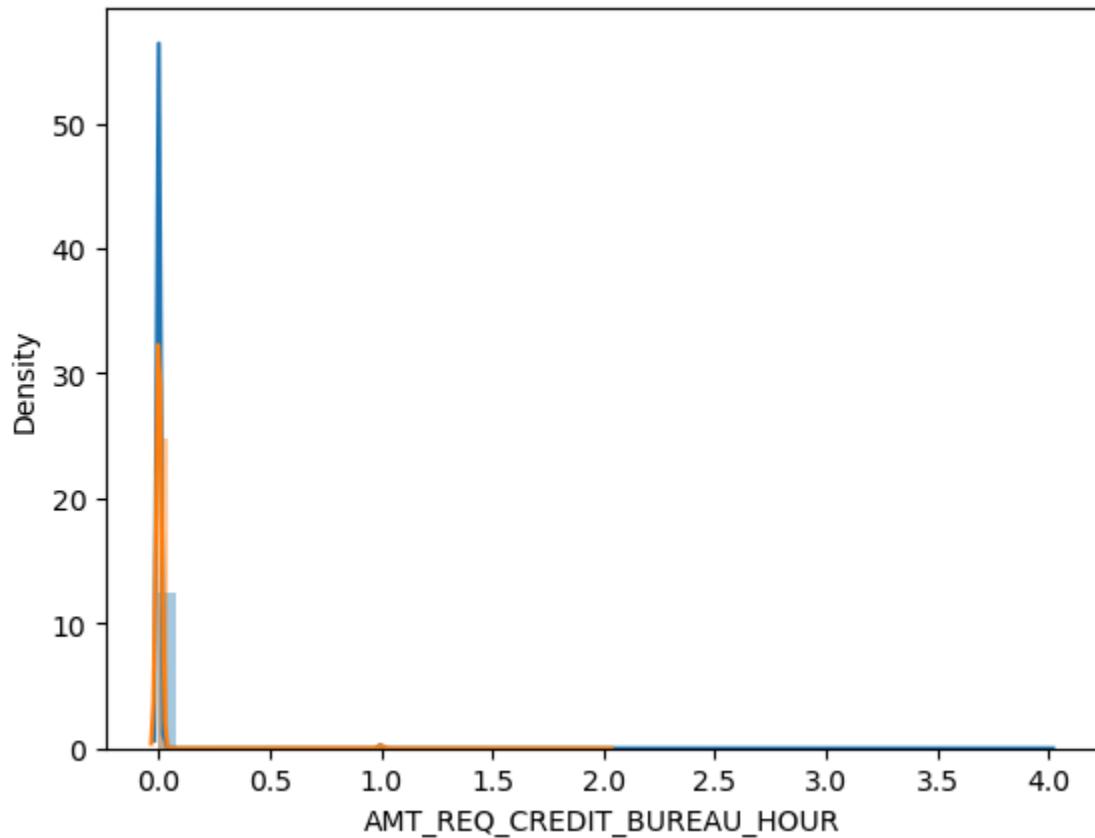
```
Out[86]: (array([2.43964e+05, 0.00000e+00, 1.43500e+03, 0.00000e+00, 5.00000e+01,
   0.00000e+00, 9.00000e+00, 1.00000e+00]),
 array([0. , 0.5, 1. , 1.5, 2. , 2.5, 3. , 3.5, 4. ]),
 <BarContainer object of 8 artists>)
```

```
Out[86]: (array([2.0402e+04, 0.0000e+00, 0.0000e+00, 0.0000e+00, 1.2500e+02,
   0.0000e+00, 0.0000e+00, 6.0000e+00]),
 array([0. , 0.25, 0.5 , 0.75, 1. , 1.25, 1.5 , 1.75, 2. ]),
 <BarContainer object of 8 artists>)
```



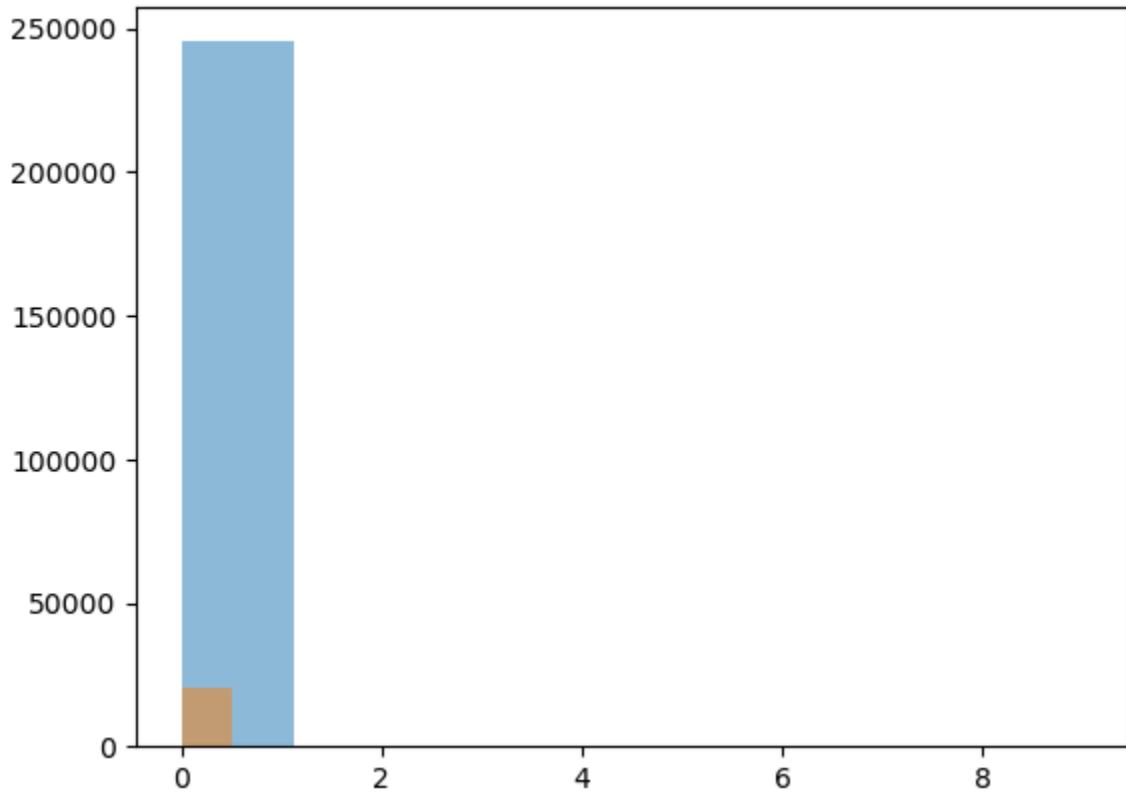
```
Out[86]: <Axes: xlabel='AMT_REQ_CREDIT_BUREAU_HOUR', ylabel='Density'>
```

```
Out[86]: <Axes: xlabel='AMT_REQ_CREDIT_BUREAU_HOUR', ylabel='Density'>
```



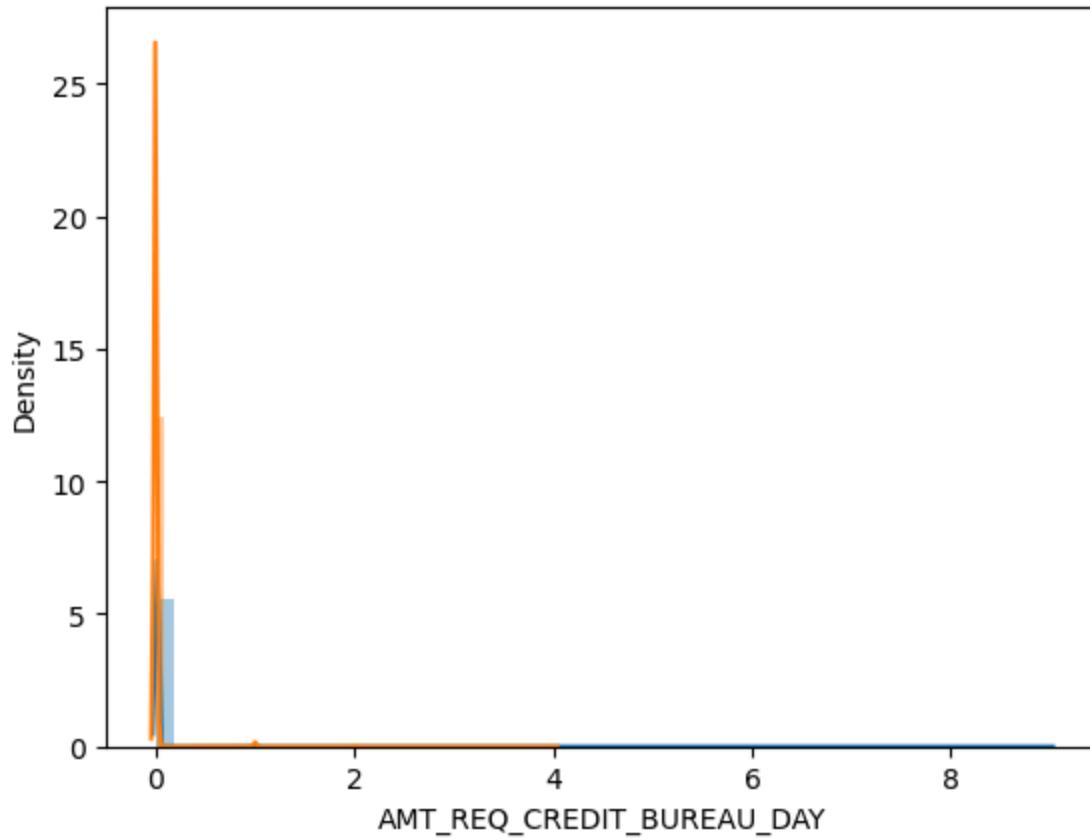
Plot of AMT_REQ_CREDIT_BUREAU_DAY

```
Out[86]: (array([2.45278e+05, 9.50000e+01, 4.30000e+01, 2.30000e+01, 9.00000e+00,
   8.00000e+00, 0.00000e+00, 3.00000e+00]),  
 array([0.    , 1.125, 2.25 , 3.375, 4.5   , 5.625, 6.75 , 7.875, 9.    ]),  
 <BarContainer object of 8 artists>)  
  
Out[86]: (array([2.0392e+04, 0.0000e+00, 1.2500e+02, 0.0000e+00, 1.1000e+01,  
 0.0000e+00, 2.0000e+00, 3.0000e+00]),  
 array([0.  , 0.5, 1.  , 1.5, 2.  , 2.5, 3.  , 3.5, 4.  ]),  
 <BarContainer object of 8 artists>)
```



```
Out[86]: <Axes: xlabel='AMT_REQ_CREDIT_BUREAU_DAY', ylabel='Density'>
```

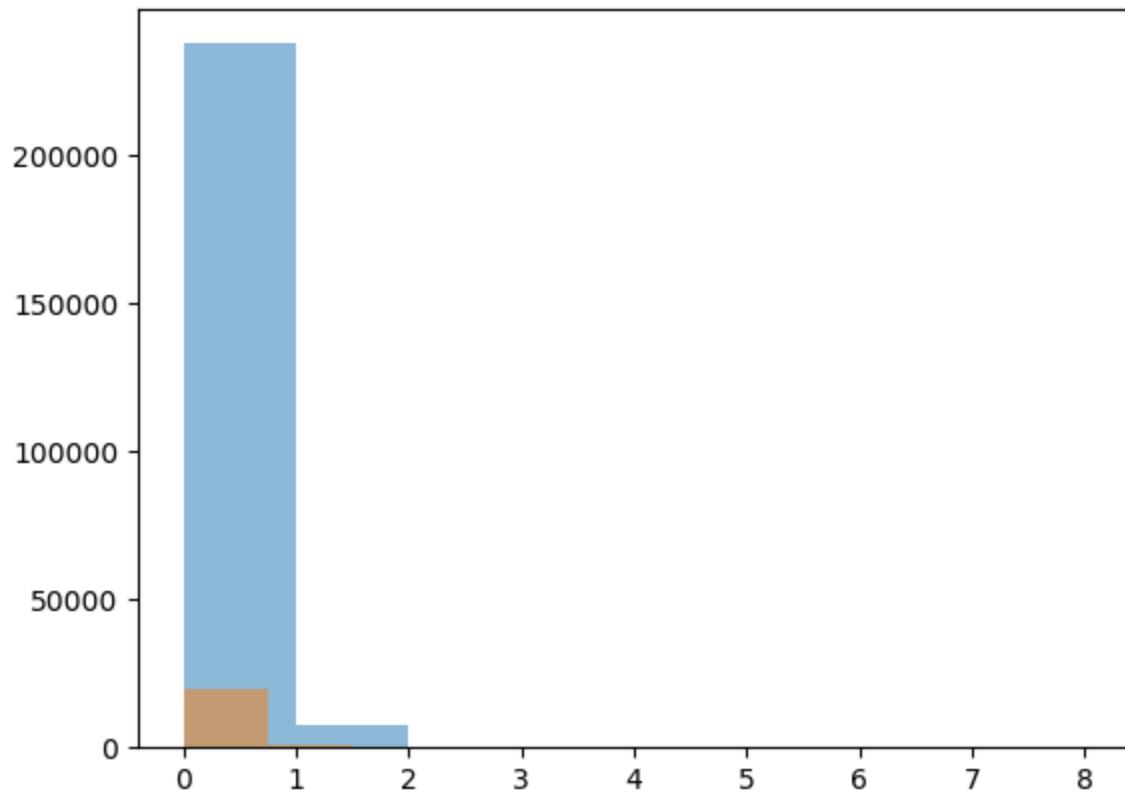
```
Out[86]: <Axes: xlabel='AMT_REQ_CREDIT_BUREAU_DAY', ylabel='Density'>
```



Plot of `AMT_REQ_CREDIT_BUREAU_WEEK`

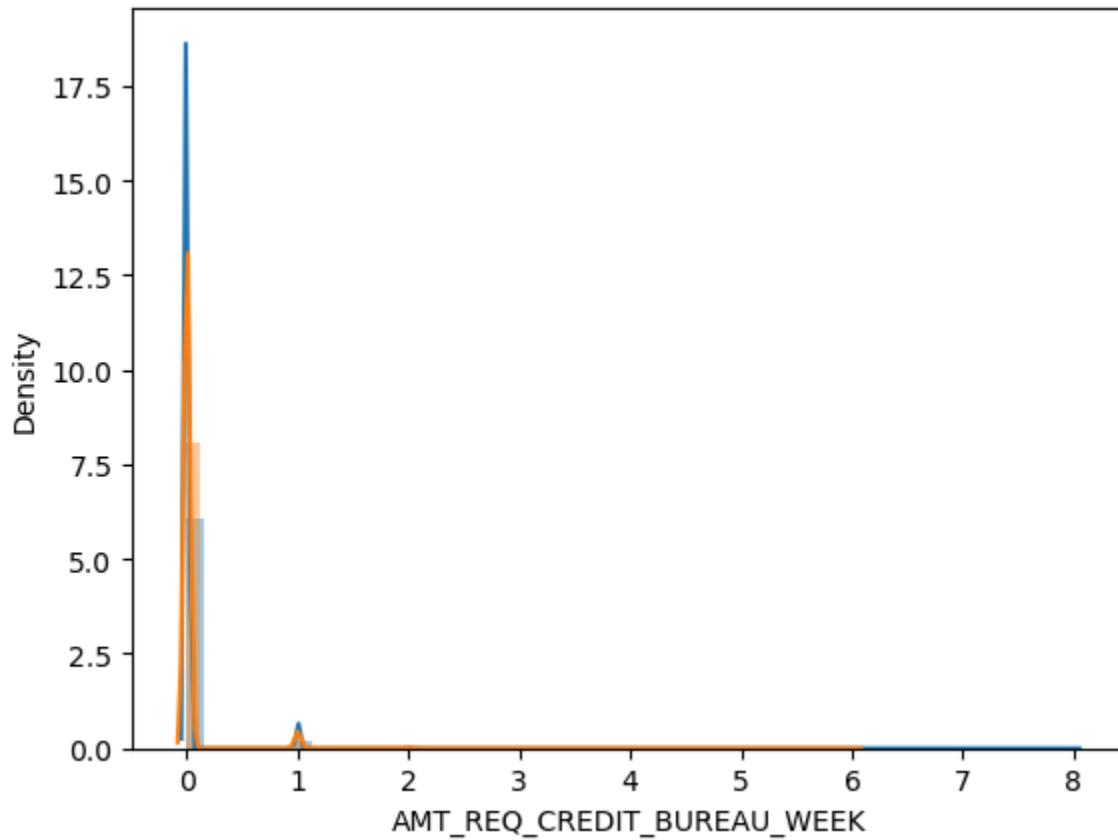
```
Out[86]: (array([2.37591e+05, 7.57000e+03, 1.79000e+02, 5.40000e+01, 3.00000e+01,
   9.00000e+00, 1.90000e+01, 7.00000e+00]),  
 array([0., 1., 2., 3., 4., 5., 6., 7., 8.]),  
 <BarContainer object of 8 artists>)
```

```
Out[86]: (array([1.9865e+04, 6.3800e+02, 2.0000e+01, 0.0000e+00, 4.0000e+00,
   4.0000e+00, 1.0000e+00, 1.0000e+00]),  
 array([0. , 0.75, 1.5 , 2.25, 3. , 3.75, 4.5 , 5.25, 6. ]),  
 <BarContainer object of 8 artists>)
```



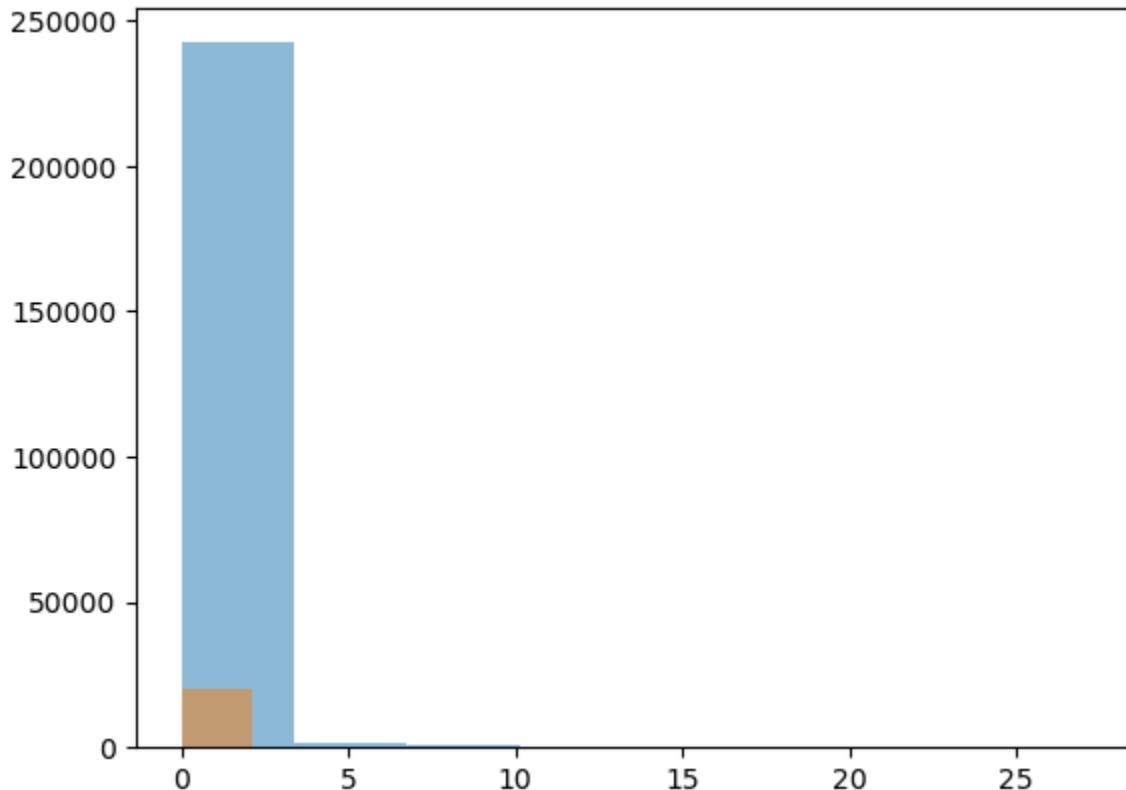
```
Out[86]: <Axes: xlabel='AMT_REQ_CREDIT_BUREAU_WEEK', ylabel='Density'>
```

```
Out[86]: <Axes: xlabel='AMT_REQ_CREDIT_BUREAU_WEEK', ylabel='Density'>
```



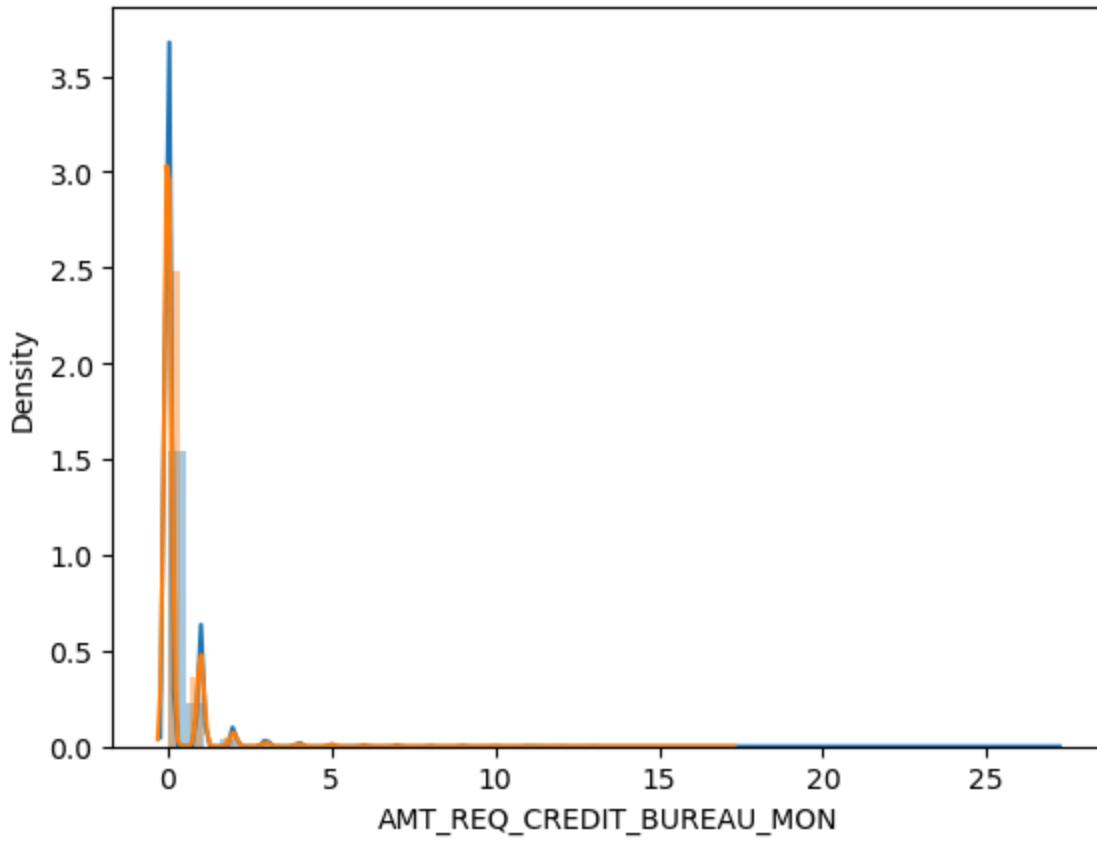
Plot of AMT_REQ_CREDIT_BUREAU_MON

```
Out[86]: (array([2.42395e+05, 1.90300e+03, 7.79000e+02, 2.60000e+02, 9.60000e+01,
   2.20000e+01, 2.00000e+00, 2.00000e+00]),  
 array([ 0.    ,  3.375,  6.75  , 10.125, 13.5  , 16.875, 20.25  , 23.625,  
 27.    ]),  
<BarContainer object of 8 artists>)  
Out[86]: (array([2.0254e+04, 1.6800e+02, 5.8000e+01, 2.3000e+01, 1.9000e+01,  
 5.0000e+00, 4.0000e+00, 2.0000e+00]),  
 array([ 0.    ,  2.125,  4.25  ,  6.375,  8.5  , 10.625, 12.75  , 14.875,  
 17.    ]),  
<BarContainer object of 8 artists>)
```



```
Out[86]: <Axes: xlabel='AMT_REQ_CREDIT_BUREAU_MON', ylabel='Density'>
```

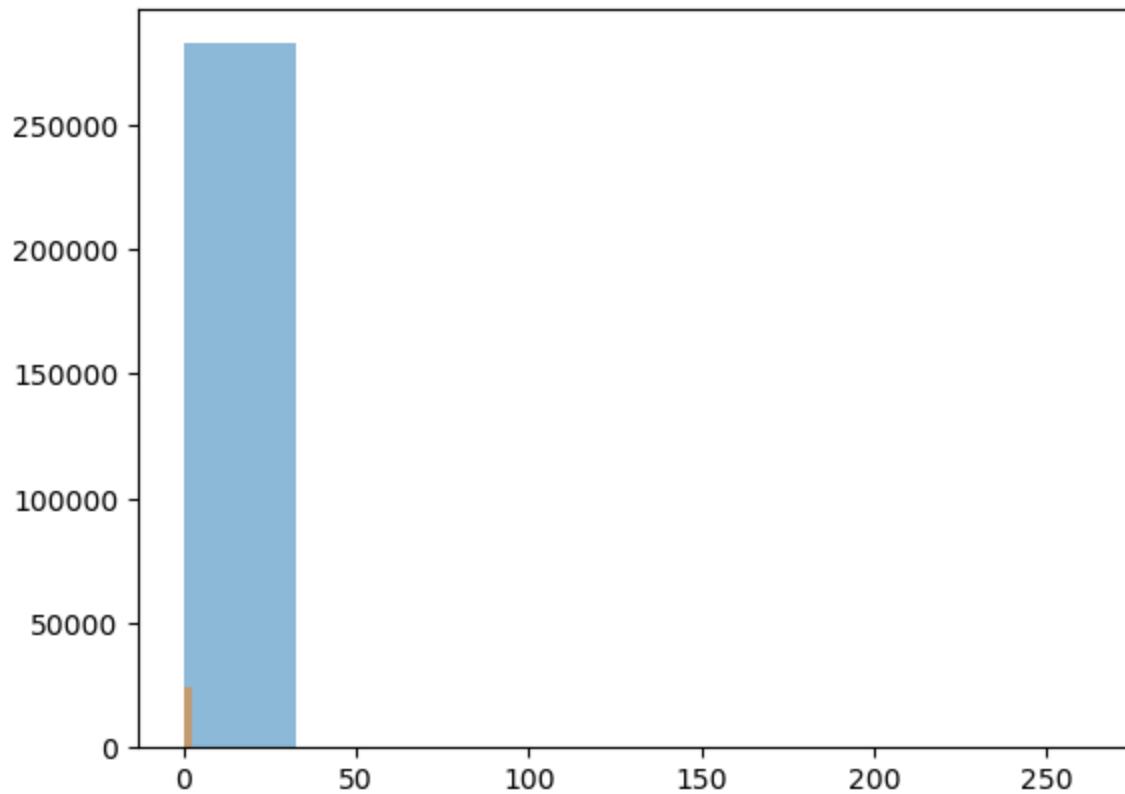
```
Out[86]: <Axes: xlabel='AMT_REQ_CREDIT_BUREAU_MON', ylabel='Density'>
```



Plot of `AMT_REQ_CREDIT_BUREAU_QRT`

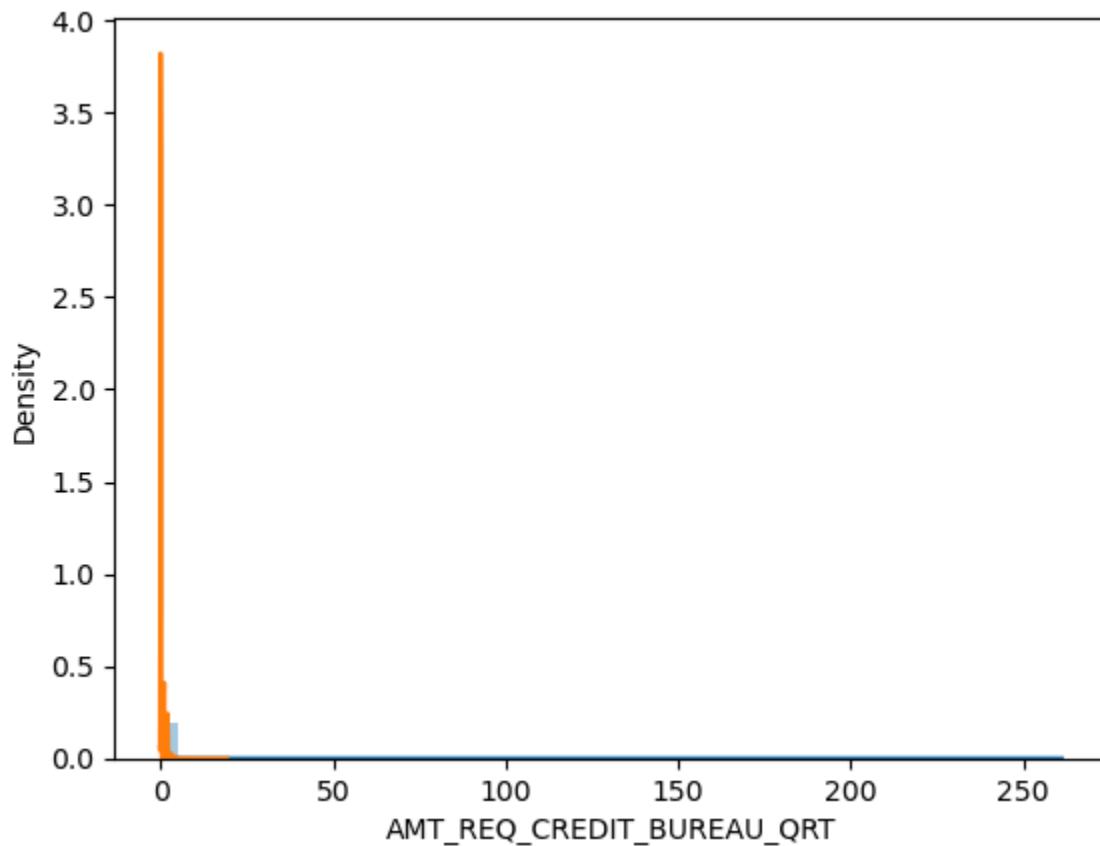
```
Out[86]: (array([2.82685e+05, 0.00000e+00, 0.00000e+00, 0.00000e+00, 0.00000e+00,
       0.00000e+00, 0.00000e+00, 1.00000e+00]),
       array([ 0.    , 32.625, 65.25 , 97.875, 130.5   , 163.125, 195.75 ,
              228.375, 261.    ]),
       <BarContainer object of 8 artists>)

Out[86]: (array([2.4627e+04, 1.8300e+02, 1.4000e+01, 0.0000e+00, 0.0000e+00,
       0.0000e+00, 0.0000e+00, 1.0000e+00]),
       array([ 0.    , 2.375, 4.75 , 7.125, 9.5   , 11.875, 14.25 , 16.625,
              19.    ]),
       <BarContainer object of 8 artists>)
```



```
Out[86]: <Axes: xlabel='AMT_REQ_CREDIT_BUREAU_QRT', ylabel='Density'>
```

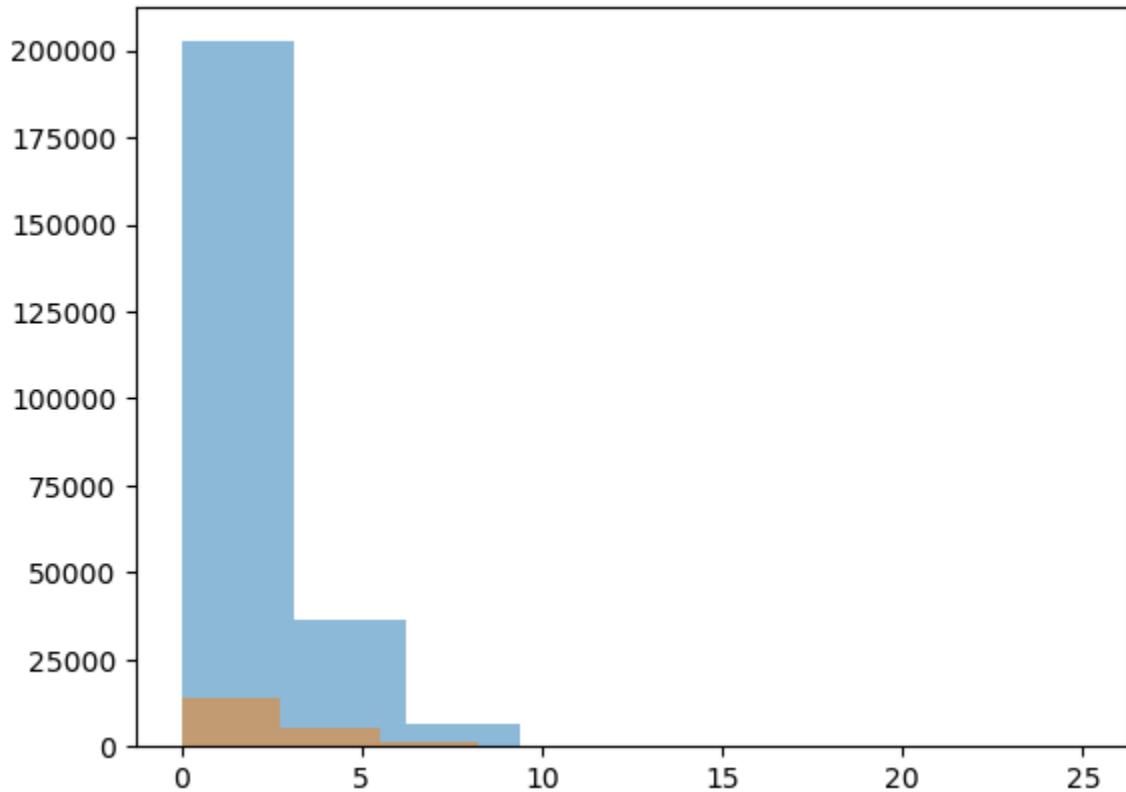
```
Out[86]: <Axes: xlabel='AMT_REQ_CREDIT_BUREAU_QRT', ylabel='Density'>
```



Plot of AMT_REQ_CREDIT_BUREAU_YEAR

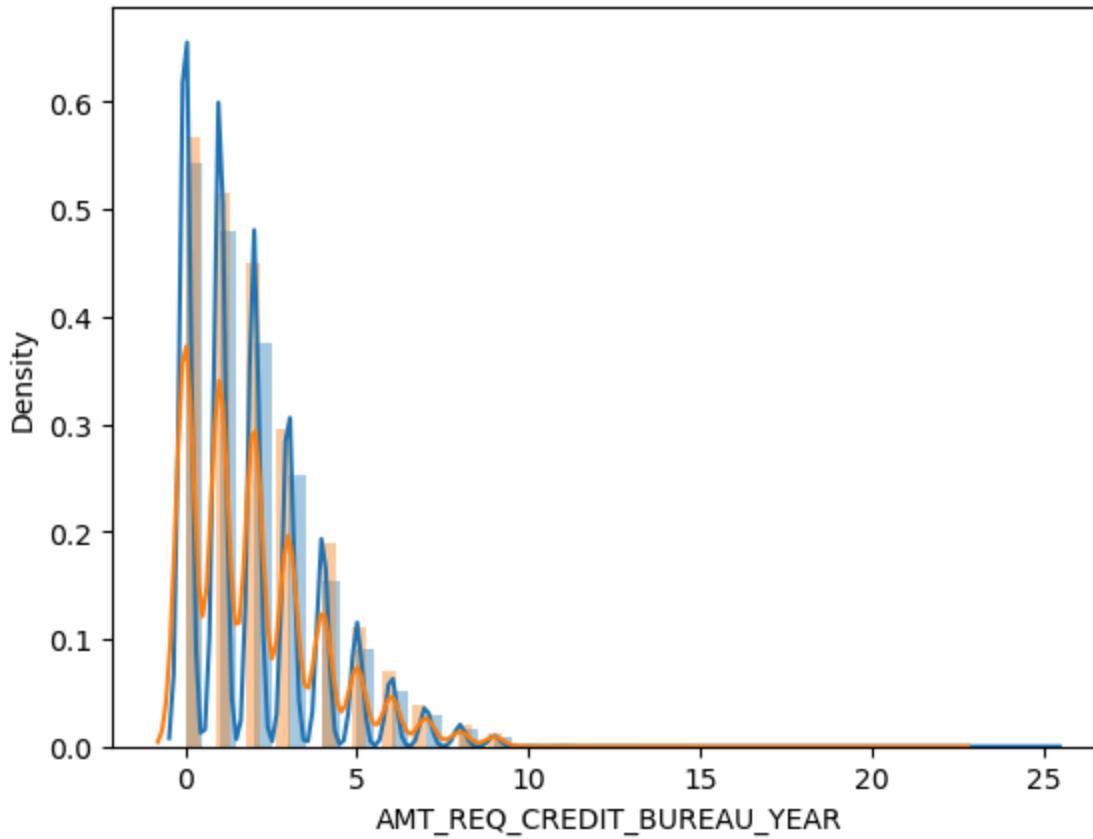
```
Out[86]: (array([2.02509e+05, 3.63880e+04, 6.43400e+03, 7.60000e+01, 3.10000e+01,
   1.30000e+01, 6.00000e+00, 2.00000e+00]),  
 array([ 0. ,  3.125,  6.25 ,  9.375, 12.5 , 15.625, 18.75 , 21.875,  
 25. ]),  
<BarContainer object of 8 artists>)
```

```
Out[86]: (array([1.3841e+04, 5.3890e+03, 1.1710e+03, 1.2200e+02, 5.0000e+00,  
 4.0000e+00, 0.0000e+00, 1.0000e+00]),  
 array([ 0. ,  2.75,  5.5 ,  8.25, 11. , 13.75, 16.5 , 19.25, 22. ]),  
<BarContainer object of 8 artists>)
```



```
Out[86]: <Axes: xlabel='AMT_REQ_CREDIT_BUREAU_YEAR', ylabel='Density'>
```

```
Out[86]: <Axes: xlabel='AMT_REQ_CREDIT_BUREAU_YEAR', ylabel='Density'>
```



```
In [87]: prev_appl = appl_data.merge(prev_app, left_on='SK_ID_CURR', right_on='SK_ID_CURR')
```

```
In [88]: prev_appl.shape
```

```
Out[88]: (1413701, 117)
```

```
In [89]: prev_appl.SK_ID_CURR.value_counts().head()
```

```
Out[89]: SK_ID_CURR
265681    73
173680    72
242412    68
206783    67
389950    64
Name: count, dtype: int64
```

```
In [90]: train_0 = appl_data.loc[appl_data['TARGET'] == 0]
train_1 = appl_data.loc[appl_data['TARGET'] == 1]
```

```
In [91]: ptrain_0 = prev_appl.loc[prev_appl['TARGET'] == 0]
ptrain_1 = prev_appl.loc[prev_appl['TARGET'] == 1]
```

```
In [94]: def plotting(column, hue):
    col = column
    hue = hue
```

```
fig = plt.figure(figsize=(13,10))

ax1 = plt.subplot(221)
appl_data[col].value_counts().plot.pie(autopct = "%1.0f%%", ax=ax1)
plt.title('Plotting data for the column: '+ column)

ax2 = plt.subplot(222)
df = pd.DataFrame()
df['0']= ((train_0[col].value_counts())/len(train_0))
df['1']= ((train_1[col].value_counts())/len(train_1))
df.plot.bar(ax=ax2)
plt.title('Plotting data for target in terms of total count')

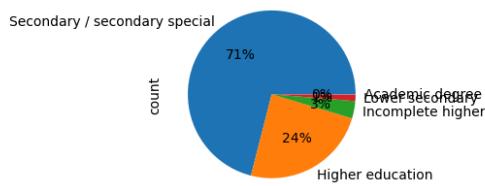
ax3 = plt.subplot(223)
sns.countplot(x=col, hue=hue, data=ptrain_0, ax = ax3)
plt.xticks(rotation=90)
plt.title('Plotting data for Target=0 in terms of percentage')

ax4 = plt.subplot(224)
sns.countplot(x=col, hue=hue, data=ptrain_1, ax = ax4)
plt.xticks(rotation=90)
plt.title('Plotting data for Target=1 in terms of percentage')

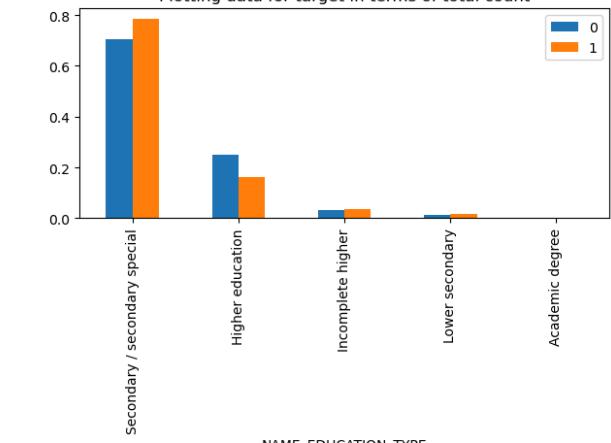
fig.tight_layout() # Or equivalently, "plt.tight_layout()"
plt.show()
```

```
In [95]: plotting('NAME_EDUCATION_TYPE', 'NAME_CONTRACT_STATUS')
```

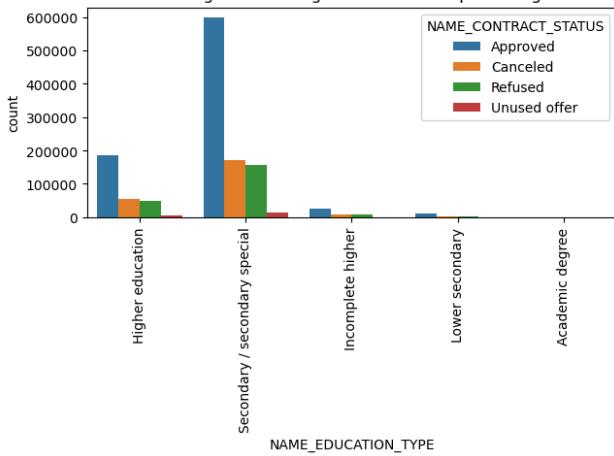
Plotting data for the column: NAME_EDUCATION_TYPE



Plotting data for target in terms of total count



Plotting data for Target=0 in terms of percentage



Plotting data for Target=1 in terms of percentage

