```python
import pandas as pd
import numpy as np


df = pd.read_csv('Bank-Loan.csv')


pd.set_option('display.max_columns', None)
df
```

	_INCOME_TOTAL	AI
	202500.0	
	270000.0	
	67500.0	
	135000.0	
	121500.0	
	...	
	270000.0	
	90000.0	
	180000.0	
	90000.0	
	90000.0	

```python
df['NAME_CONTRACT_TYPE'].unique()
```

```
array(['Cash loans', 'Revolving loans'], dtype=object)
```

```python
df['NAME_CONTRACT_TYPE'].replace({'Cash loans':0,'Revolving loans':1},inplace=True)
```

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

	SK_ID_CURR	TARGET	NAME_CONTRACT_TYPE	CODE_GENDER	FLAG_OWN_CAR	FLAG_OWN_REALTY	CNT_CHILDREN	AMT_INCOME_TOTAL	AMT_(
0	100002	1	0	M	N	Y	0	202500.0	40
1	100003	0	0	F	N	N	0	270000.0	129
2	100004	0	1	M	Y	Y	0	67500.0	13
3	100006	0	0	F	N	Y	0	135000.0	31
4	100007	0	0	M	N	Y	0	121500.0	51

```python
df['CODE_GENDER'].replace({'M':0,'F':1},inplace=True)
df.head()
```

| | SK_ID_CURR | TARGET | NAME_CONTRACT_TYPE | CODE_GENDER | FLAG_OWN_CAR | FLAG_OWN_REALTY | CNT_CHILDREN | AMT_INCOME_TOTAL | AMT_CREDIT | A |

```
df['FLAG_OWN_CAR'].replace({'N':0,'Y':1},inplace=True)
df.head()
```

	SK_ID_CURR	TARGET	NAME_CONTRACT_TYPE	CODE_GENDER	FLAG_OWN_CAR	FLAG_OWN_REALTY	CNT_CHILDREN	AMT_INCOME_TOTAL	AMT_CREDIT	A
0	100002	1	0	0	0	Y	0	202500.0	406597.5	
1	100003	0	0	1	0	N	0	270000.0	1293502.5	
2	100004	0	1	0	1	Y	0	67500.0	135000.0	
3	100006	0	0	1	0	Y	0	135000.0	312682.5	
4	100007	0	0	0	0	Y	0	121500.0	513000.0	

```
df['FLAG_OWN_REALTY'].replace({'N':0,'Y':1},inplace=True)
df.head()
```

	SK_ID_CURR	TARGET	NAME_CONTRACT_TYPE	CODE_GENDER	FLAG_OWN_CAR	FLAG_OWN_REALTY	CNT_CHILDREN	AMT_INCOME_TOTAL	AMT_CREDIT	A
0	100002	1	0	0	0	1	0	202500.0	406597.5	
1	100003	0	0	1	0	0	0	270000.0	1293502.5	
2	100004	0	1	0	1	1	0	67500.0	135000.0	
3	100006	0	0	1	0	1	0	135000.0	312682.5	
4	100007	0	0	0	0	1	0	121500.0	513000.0	

```
df['NAME_TYPE_SUITE'].unique()
```

```
array(['Unaccompanied', 'Family', 'Spouse, partner', 'Children',
       'Other_A', nan, 'Other_B', 'Group of people'], dtype=object)
```

```
df['NAME_TYPE_SUITE'].replace({'Unaccompanied':0,'Family':1,'Spouse, partner':2,'Children':3,'Other_A':4,'Other_B':5,'Group of people':6]
df.head()
```

	SK_ID_CURR	TARGET	NAME_CONTRACT_TYPE	CODE_GENDER	FLAG_OWN_CAR	FLAG_OWN_REALTY	CNT_CHILDREN	AMT_INCOME_TOTAL	AMT_CREDIT	A
0	100002	1	0	0	0	1	0	202500.0	406597.5	
1	100003	0	0	1	0	0	0	270000.0	1293502.5	
2	100004	0	1	0	1	1	0	67500.0	135000.0	
3	100006	0	0	1	0	1	0	135000.0	312682.5	
4	100007	0	0	0	0	1	0	121500.0	513000.0	

```
df['NAME_INCOME_TYPE'].unique()
```

```
array(['Working', 'State servant', 'Commercial associate', 'Pensioner',
       'Unemployed', 'Student', 'Businessman', 'Maternity leave'],
      dtype=object)
```

```
df['NAME_INCOME_TYPE'].replace({'Working':0,'State servant':1,'Commercial associate':2,'Pensioner':3,'Unemployed':4,'Student':5,'Busines
df.head()
```

	SK_ID_CURR	TARGET	NAME_CONTRACT_TYPE	CODE_GENDER	FLAG_OWN_CAR	FLAG_OWN_REALTY	CNT_CHILDREN	AMT_INCOME_TOTAL	AMT_CREDIT A
0	100002	1	0	0	0	1	0	202500.0	406597.5
1	100003	0	0	1	0	0	0	270000.0	1293502.5
2	100004	0	1	0	1	1	0	67500.0	135000.0

```
df['NAME_EDUCATION_TYPE'].unique()
```

```
array(['Secondary / secondary special', 'Higher education',
       'Incomplete higher', 'Lower secondary', 'Academic degree'],
      dtype=object)
```

```
df['NAME_EDUCATION_TYPE'].replace({'Secondary / secondary special':0,'Higher education':1,'Incomplete higher':2,'Lower secondary':3,'Aca
df.head()
```

	SK_ID_CURR	TARGET	NAME_CONTRACT_TYPE	CODE_GENDER	FLAG_OWN_CAR	FLAG_OWN_REALTY	CNT_CHILDREN	AMT_INCOME_TOTAL	AMT_CREDIT A
0	100002	1	0	0	0	1	0	202500.0	406597.5
1	100003	0	0	1	0	0	0	270000.0	1293502.5
2	100004	0	1	0	1	1	0	67500.0	135000.0
3	100006	0	0	1	0	1	0	135000.0	312682.5
4	100007	0	0	0	0	1	0	121500.0	513000.0

```
df['NAME_FAMILY_STATUS'].unique()
```

```
array(['Single / not married', 'Married', 'Civil marriage', 'Widow',
       'Separated', 'Unknown'], dtype=object)
```

```
df['NAME_FAMILY_STATUS'].replace({'Single / not married':0,'Married':1,'Civil marriage':2,'Widow':3,'Separated':4,'Unknown':np.nan},inpla
df.head()
```

	SK_ID_CURR	TARGET	NAME_CONTRACT_TYPE	CODE_GENDER	FLAG_OWN_CAR	FLAG_OWN_REALTY	CNT_CHILDREN	AMT_INCOME_TOTAL	AMT_CREDIT A
0	100002	1	0	0	0	1	0	202500.0	406597.5
1	100003	0	0	1	0	0	0	270000.0	1293502.5
2	100004	0	1	0	1	1	0	67500.0	135000.0
3	100006	0	0	1	0	1	0	135000.0	312682.5
4	100007	0	0	0	0	1	0	121500.0	513000.0

```
df['NAME_HOUSING_TYPE'].unique()
```

```
array(['House / apartment', 'Rented apartment', 'With parents',
       'Municipal apartment', 'Office apartment', 'Co-op apartment'],
      dtype=object)
```

```
df['NAME_HOUSING_TYPE'].replace({'House / apartment':0,'Rented apartment':1,'With parents':2,'Municipal apartment':3,'Offi
df.head()
```

	SK_ID_CURR	TARGET	NAME_CONTRACT_TYPE	CODE_GENDER	FLAG_OWN_CAR	FLAG_OWN_REALTY	CNT_CHILDREN	AMT_INCOME_TOTAL	AMT_CREDIT A
0	100002	1	0	0	0	1	0	202500.0	406597.5
1	100003	0	0	1	0	0	0	270000.0	1293502.5
2	100004	0	1	0	1	1	0	67500.0	135000.0
3	100006	0	0	1	0	1	0	135000.0	312682.5
4	100007	0	0	0	0	1	0	121500.0	513000.0

```
df['OCCUPATION_TYPE'].unique()
```

```
array(['Laborers', 'Core staff', 'Accountants', 'Managers', nan,
       'Drivers', 'Sales staff', 'Cleaning staff', 'Cooking staff',
       'Private service staff', 'Medicine staff', 'Security staff',
       'High skill tech staff', 'Waiters/barmen staff',
       'Low-skill Laborers', 'Realty agents', 'Secretaries', 'IT staff',
       'HR staff'], dtype=object)
```

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

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

	SK_ID_CURR	TARGET	NAME_CONTRACT_TYPE	CODE_GENDER	FLAG_OWN_CAR	FLAG_OWN_REALTY	CNT_CHILDREN	AMT_INCOME_TOTAL	AMT_CREDIT
0	100002	1	0	0	0	1	0	202500.0	406597.5
1	100003	0	0	1	0	0	0	270000.0	1293502.5
2	100004	0	1	0	1	1	0	67500.0	135000.0
3	100006	0	0	1	0	1	0	135000.0	312682.5
4	100007	0	0	0	0	1	0	121500.0	513000.0

```python
df['WEEKDAY_APPR_PROCESS_START'].unique()
```

```
array(['WEDNESDAY', 'MONDAY', 'THURSDAY', 'SUNDAY', 'SATURDAY', 'FRIDAY',
       'TUESDAY'], dtype=object)
```

```python
df['WEEKDAY_APPR_PROCESS_START'].replace({'SUNDAY': 0,
    'MONDAY': 1,
    'TUESDAY': 2,
    'WEDNESDAY': 3,
    'THURSDAY': 4,
    'FRIDAY': 5,
    'SATURDAY': 6},inplace=True)
df.head()
```

	SK_ID_CURR	TARGET	NAME_CONTRACT_TYPE	CODE_GENDER	FLAG_OWN_CAR	FLAG_OWN_REALTY	CNT_CHILDREN	AMT_INCOME_TOTAL	AMT_CREDIT
0	100002	1	0	0	0	1	0	202500.0	406597.5
1	100003	0	0	1	0	0	0	270000.0	1293502.5
2	100004	0	1	0	1	1	0	67500.0	135000.0
3	100006	0	0	1	0	1	0	135000.0	312682.5
4	100007	0	0	0	0	1	0	121500.0	513000.0

```python
df['ORGANIZATION_TYPE'].unique()
```

```
array(['Business Entity Type 3', 'School', 'Government', 'Religion',
       'Other', 'XNA', 'Electricity', 'Medicine',
       'Business Entity Type 2', 'Self-employed', 'Transport: type 2',
       'Construction', 'Housing', 'Kindergarten', 'Trade: type 7',
       'Industry: type 11', 'Military', 'Services', 'Security Ministries',
       'Transport: type 4', 'Industry: type 1', 'Emergency', 'Security',
       'Trade: type 2', 'University', 'Transport: type 3', 'Police',
       'Business Entity Type 1', 'Postal', 'Industry: type 4',
       'Agriculture', 'Restaurant', 'Culture', 'Hotel',
```

```
                'Industry: type 7', 'Trade: type 3', 'Industry: type 3', 'Bank',
                'Industry: type 9', 'Insurance', 'Trade: type 6',
                'Industry: type 2', 'Transport: type 1', 'Industry: type 12',
                'Mobile', 'Trade: type 1', 'Industry: type 5', 'Industry: type 10',
                'Legal Services', 'Advertising', 'Trade: type 5', 'Cleaning',
                'Industry: type 13', 'Trade: type 4', 'Telecom',
                'Industry: type 8', 'Realtor', 'Industry: type 6'], dtype=object)
```

```python
df['ORGANIZATION_TYPE'].replace({
    'Business Entity Type 3': 0,
    'School': 1,
    'Government': 2,
    'Religion': 3,
    'Other': 4,
    'XNA': 5,
    'Electricity': 6,
    'Medicine': 7,
    'Business Entity Type 2': 8,
    'Self-employed': 9,
    'Transport: type 2': 10,
    'Construction': 11,
    'Housing': 12,
    'Kindergarten': 13,
    'Trade: type 7': 14,
    'Industry: type 11': 15,
    'Military': 16,
    'Services': 17,
    'Security Ministries': 18,
    'Transport: type 4': 19,
    'Industry: type 1': 20,
    'Emergency': 21,
    'Security': 22,
    'Trade: type 2': 23,
    'University': 24,
    'Transport: type 3': 25,
    'Police': 26,
    'Business Entity Type 1': 27,
    'Postal': 28,
    'Industry: type 4': 29,
    'Agriculture': 30,
    'Restaurant': 31,
    'Culture': 32,
    'Hotel': 33,
    'Industry: type 7': 34,
    'Trade: type 3': 35,
    'Industry: type 3': 36,
    'Bank': 37,
    'Industry: type 9': 38,
    'Insurance': 39,
    'Trade: type 6': 40,
    'Industry: type 2': 41,
    'Transport: type 1': 42,
    'Industry: type 12': 43,
    'Mobile': 44,
    'Trade: type 1': 45,
    'Industry: type 5': 46,
    'Industry: type 10': 47,
    'Legal Services': 48,
    'Advertising': 49,
    'Trade: type 5': 50,
    'Cleaning': 51,
    'Industry: type 13': 52,
    'Trade: type 4': 53,
    'Telecom': 54,
    'Industry: type 8': 55,
    'Realtor': 56,
    'Industry: type 6': 57
},inplace=True)
```

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

	SK_ID_CURR	TARGET	NAME_CONTRACT_TYPE	CODE_GENDER	FLAG_OWN_CAR	FLAG_OWN_REALTY	CNT_CHILDREN	AMT_INCOME_TOTAL	AMT_CREDIT	A
0	100002	1	0	0	0	1	0	202500.0	406597.5	
1	100003	0	0	1	0	0	0	270000.0	1293502.5	
2	100004	0	1	0	1	1	0	67500.0	135000.0	
3	100006	0	0	1	0	1	0	135000.0	312682.5	
4	100007	0	0	0	0	1	0	121500.0	513000.0	

```python
df['FONDKAPREMONT_MODE'].unique()
```

```
array(['reg oper account', nan, 'org spec account',
       'reg oper spec account', 'not specified'], dtype=object)
```

```python
df['FONDKAPREMONT_MODE'].replace({'reg oper account': 0,
    'org spec account': 1,
    'reg oper spec account': 2,
    'not specified': np.nan},inplace=True)
df.head()
```

	SK_ID_CURR	TARGET	NAME_CONTRACT_TYPE	CODE_GENDER	FLAG_OWN_CAR	FLAG_OWN_REALTY	CNT_CHILDREN	AMT_INCOME_TOTAL	AMT_CREDIT
0	100002	1	0	0	0	1	0	202500.0	406597.5
1	100003	0	0	1	0	0	0	270000.0	1293502.5
2	100004	0	1	0	1	1	0	67500.0	135000.0
3	100006	0	0	1	0	1	0	135000.0	312682.5
4	100007	0	0	0	0	1	0	121500.0	513000.0

```python
df['HOUSETYPE_MODE'].unique()
```

```
array(['block of flats', nan, 'terraced house', 'specific housing'],
      dtype=object)
```

```python
df['HOUSETYPE_MODE'].replace({
    'block of flats':0,
    'terraced house':1,
    'specific housing':2
},inplace=True)
df.head()
```

	SK_ID_CURR	TARGET	NAME_CONTRACT_TYPE	CODE_GENDER	FLAG_OWN_CAR	FLAG_OWN_REALTY	CNT_CHILDREN	AMT_INCOME_TOTAL	AMT_CREDIT
0	100002	1	0	0	0	1	0	202500.0	406597.5
1	100003	0	0	1	0	0	0	270000.0	1293502.5
2	100004	0	1	0	1	1	0	67500.0	135000.0
3	100006	0	0	1	0	1	0	135000.0	312682.5
4	100007	0	0	0	0	1	0	121500.0	513000.0

```python
df['WALLSMATERIAL_MODE'].unique()
```

```
array(['Stone, brick', 'Block', nan, 'Panel', 'Mixed', 'Wooden', 'Others',
       'Monolithic'], dtype=object)
```

```python
df['WALLSMATERIAL_MODE'].replace({
    'Stone, brick': 0,
    'Block': 1,
    'Panel': 2,
    'Mixed': 3,
    'Wooden': 4,
    'Others': 5,
    'Monolithic': 6
},inplace=True)
df.head()
```

	SK_ID_CURR	TARGET	NAME_CONTRACT_TYPE	CODE_GENDER	FLAG_OWN_CAR	FLAG_OWN_REALTY	CNT_CHILDREN	AMT_INCOME_TOTAL	AMT_CREDIT
0	100002	1	0	0	0	1	0	202500.0	406597.5
1	100003	0	0	1	0	0	0	270000.0	1293502.5
2	100004	0	1	0	1	1	0	67500.0	135000.0
3	100006	0	0	1	0	1	0	135000.0	312682.5
4	100007	0	0	0	0	1	0	121500.0	513000.0

```python
df['AMT_INCOME_TOTAL'].min()
```

```
25650.0
```

```python
df['AMT_INCOME_TOTAL'].max()
```

```
117000000.0
```

```python
_TOTAL']=df['AMT_INCOME_TOTAL'].apply(lambda v:(v-df['AMT_INCOME_TOTAL'].min()))/(df['AMT_INCOME_TOTAL'].max()-df['AMT_INCO
```

```python
df['AMT_CREDIT'].min()
```

```
45000.0
```

```python
df['AMT_CREDIT'].max()
```

```
4050000.0
```

```python
df['AMT_CREDIT']=df['AMT_CREDIT'].apply(lambda v:(v-df['AMT_CREDIT'].min()))/(df['AMT_CREDIT'].max()-df['AMT_CREDIT'].min())
df.head()
```

	SK_ID_CURR	TARGET	NAME_CONTRACT_TYPE	CODE_GENDER	FLAG_OWN_CAR	FLAG_OWN_REALTY	CNT_CHILDREN	AMT_INCOME_TOTAL	AMT_CREDIT	A
0	100002	1	0	0	0	1	0	0.001512	0.090287	
1	100003	0	0	1	0	0	0	0.002089	0.311736	
2	100004	0	1	0	1	1	0	0.000358	0.022472	
3	100006	0	0	1	0	1	0	0.000935	0.066837	
4	100007	0	0	0	0	1	0	0.000819	0.116854	

```python
df['AMT_ANNUITY'].min()
```

```
1615.5
```

```python
df['AMT_ANNUITY'].max()
```

```
258025.5
```

```python
df['AMT_ANNUITY']=df['AMT_ANNUITY'].apply(lambda v:(v-df['AMT_ANNUITY'].min()))/(df['AMT_ANNUITY'].max()-df['AMT_ANNUITY'].min())
df.head()
```

	SK_ID_CURR	TARGET	NAME_CONTRACT_TYPE	CODE_GENDER	FLAG_OWN_CAR	FLAG_OWN_REALTY	CNT_CHILDREN	AMT_INCOME_TOTAL	AMT_CREDIT	A
0	100002	1	0	0	0	1	0	0.001512	0.090287	
1	100003	0	0	1	0	0	0	0.002089	0.311736	
2	100004	0	1	0	1	1	0	0.000358	0.022472	
3	100006	0	0	1	0	1	0	0.000935	0.066837	
4	100007	0	0	0	0	1	0	0.000819	0.116854	

```python
df['AMT_GOODS_PRICE'].min()
```

```
40500.0
```

```python
df['AMT_GOODS_PRICE'].max()
```

```
4050000.0
```

```
df['AMT_GOODS_PRICE']=df['AMT_GOODS_PRICE'].apply(lambda v:(v-df['AMT_GOODS_PRICE'].min()))/(df['AMT_GOODS_PRICE'].max()-df['AMT_GOODS_PR
df.head()
```

	SK_ID_CURR	TARGET	NAME_CONTRACT_TYPE	CODE_GENDER	FLAG_OWN_CAR	FLAG_OWN_REALTY	CNT_CHILDREN	AMT_INCOME_TOTAL	AMT_CREDIT
0	100002	1	0	0	0	1	0	0.001512	0.090287
1	100003	0	0	1	0	0	0	0.002089	0.311736
2	100004	0	1	0	1	1	0	0.000358	0.022472
3	100006	0	0	1	0	1	0	0.000935	0.066837
4	100007	0	0	0	0	1	0	0.000819	0.116854

```
df['DAYS_BIRTH']=df['DAYS_BIRTH'].apply(lambda v:(v-df['DAYS_BIRTH'].min()))/(df['DAYS_BIRTH'].max()-df['DAYS_BIRTH'].min(
df.head()
```

	SK_ID_CURR	TARGET	NAME_CONTRACT_TYPE	CODE_GENDER	FLAG_OWN_CAR	FLAG_OWN_REALTY	CNT_CHILDREN	AMT_INCOME_TOTAL	AMT_CREDIT
0	100002	1	0	0	0	1	0	0.001512	0.090287
1	100003	0	0	1	0	0	0	0.002089	0.311736
2	100004	0	1	0	1	1	0	0.000358	0.022472
3	100006	0	0	1	0	1	0	0.000935	0.066837
4	100007	0	0	0	0	1	0	0.000819	0.116854

```
df['DAYS_EMPLOYED']=df['DAYS_EMPLOYED'].apply(lambda v:(v-df['DAYS_EMPLOYED'].min()))/(df['DAYS_EMPLOYED'].max()-df['DAYS_
df.head()
```

	SK_ID_CURR	TARGET	NAME_CONTRACT_TYPE	CODE_GENDER	FLAG_OWN_CAR	FLAG_OWN_REALTY	CNT_CHILDREN	AMT_INCOME_TOTAL	AMT_CREDIT
0	100002	1	0	0	0	1	0	0.001512	0.090287
1	100003	0	0	1	0	0	0	0.002089	0.311736
2	100004	0	1	0	1	1	0	0.000358	0.022472
3	100006	0	0	1	0	1	0	0.000935	0.066837
4	100007	0	0	0	0	1	0	0.000819	0.116854

```
df['DAYS_REGISTRATION'].min()
```

```
-24672.0
```

```
df['DAYS_REGISTRATION'].max()
```

```
0.0
```

```
df['DAYS_REGISTRATION']=df['DAYS_REGISTRATION'].apply(lambda v:(v-df['DAYS_REGISTRATION'].min()))/(df['DAYS_REGISTRATION'].max()-df['DAYS
df.head()
```

	SK_ID_CURR	TARGET	NAME_CONTRACT_TYPE	CODE_GENDER	FLAG_OWN_CAR	FLAG_OWN_REALTY	CNT_CHILDREN	AMT_INCOME_TOTAL	AMT_C
0	100002	1	0	0	0	1	0	0.001512	0.0
1	100003	0	0	1	0	0	0	0.002089	0.3
2	100004	0	1	0	1	1	0	0.000358	0.0
3	100006	0	0	1	0	1	0	0.000935	0.0
4	100007	0	0	0	0	1	0	0.000819	0.

```
df['DAYS_ID_PUBLISH'].min()
```

```
-7197
```

```
df['DAYS_ID_PUBLISH'].max()
```

```
0
```

```
df['DAYS_ID_PUBLISH']=df['DAYS_ID_PUBLISH'].apply(lambda v:(v-df['DAYS_ID_PUBLISH'].min()))/(df['DAYS_ID_PUBLISH'].max()-df['DAYS_ID_PUBI
df.head()
```

	SK_ID_CURR	TARGET	NAME_CONTRACT_TYPE	CODE_GENDER	FLAG_OWN_CAR	FLAG_OWN_REALTY	CNT_CHILDREN	AMT_INCOME_TOTAL	AMT_CREDIT	A
0	100002	1	0	0	0	1	0	0.001512	0.090287	
1	100003	0	0	1	0	0	0	0.002089	0.311736	
2	100004	0	1	0	1	1	0	0.000358	0.022472	
3	100006	0	0	1	0	1	0	0.000935	0.066837	
4	100007	0	0	0	0	1	0	0.000819	0.116854	

```
df['OWN_CAR_AGE'].min()
```

```
0.0
```

```
df['OWN_CAR_AGE'].max()
```

```
91.0
```

```
df['OWN_CAR_AGE']=df['OWN_CAR_AGE'].apply(lambda v:(v-df['OWN_CAR_AGE'].min()))/(df['OWN_CAR_AGE'].max()-df['OWN_CAR_AGE'].min())
df.head()
```

	SK_ID_CURR	TARGET	NAME_CONTRACT_TYPE	CODE_GENDER	FLAG_OWN_CAR	FLAG_OWN_REALTY	CNT_CHILDREN	AMT_INCOME_TOTAL	AMT_CREDIT	A
0	100002	1	0	0	0	1	0	0.001512	0.090287	
1	100003	0	0	1	0	0	0	0.002089	0.311736	
2	100004	0	1	0	1	1	0	0.000358	0.022472	
3	100006	0	0	1	0	1	0	0.000935	0.066837	
4	100007	0	0	0	0	1	0	0.000819	0.116854	

```
df['HOUR_APPR_PROCESS_START'].min()
```

```
0
```

```
df['HOUR_APPR_PROCESS_START'].max()
```

```
23
```

```
df['HOUR_APPR_PROCESS_START']=df['HOUR_APPR_PROCESS_START'].apply(lambda v:(v-df['HOUR_APPR_PROCESS_START'].min()))/(df['HOUR_APPR_PROCES
df.head()
```

	SK_ID_CURR	TARGET	NAME_CONTRACT_TYPE	CODE_GENDER	FLAG_OWN_CAR	FLAG_OWN_REALTY	CNT_CHILDREN	AMT_INCOME_TOTAL	AMT_CREDIT	A
0	100002	1	0	0	0	1	0	0.001512	0.090287	
1	100003	0	0	1	0	0	0	0.002089	0.311736	
2	100004	0	1	0	1	1	0	0.000358	0.022472	
3	100006	0	0	1	0	1	0	0.000935	0.066837	
4	100007	0	0	0	0	1	0	0.000819	0.116854	

```
df['EMERGENCYSTATE_MODE'].unique()
```

```
array(['No', nan, 'Yes'], dtype=object)
```

```
df['EMERGENCYSTATE_MODE'].replace({'No':0,'Yes':1},inplace=True)
df.head()
```

	SK_ID_CURR	TARGET	NAME_CONTRACT_TYPE	CODE_GENDER	FLAG_OWN_CAR	FLAG_OWN_REALTY	CNT_CHILDREN	AMT_INCOME_TOTAL	AMT_CREDIT	

```
df['OBS_30_CNT_SOCIAL_CIRCLE'].min()
```

    0.0

2	100004	0	1	0	1	1	0	0.000358	0.022472

```
df['OBS_30_CNT_SOCIAL_CIRCLE'].max()
```

    348.0

```
df['OBS_30_CNT_SOCIAL_CIRCLE']=df['OBS_30_CNT_SOCIAL_CIRCLE'].apply(lambda v:(v-df['OBS_30_CNT_SOCIAL_CIRCLE'].min()))/(df['OBS_30_CNT_S(
df.head()
```

	SK_ID_CURR	TARGET	NAME_CONTRACT_TYPE	CODE_GENDER	FLAG_OWN_CAR	FLAG_OWN_REALTY	CNT_CHILDREN	AMT_INCOME_TOTAL	AMT_CREDIT	A
0	100002	1	0	0	0	1	0	0.001512	0.090287	
1	100003	0	0	1	0	0	0	0.002089	0.311736	
2	100004	0	1	0	1	1	0	0.000358	0.022472	
3	100006	0	0	1	0	1	0	0.000935	0.066837	
4	100007	0	0	0	0	1	0	0.000819	0.116854	

```
df['DEF_30_CNT_SOCIAL_CIRCLE'].min()
```

    0.0

```
df['DEF_30_CNT_SOCIAL_CIRCLE'].max()
```

    34.0

```
df['DEF_30_CNT_SOCIAL_CIRCLE']=df['DEF_30_CNT_SOCIAL_CIRCLE'].apply(lambda v:(v-df['DEF_30_CNT_SOCIAL_CIRCLE'].min()))/(df['DEF_30_CNT_S(
df.head()
```

	SK_ID_CURR	TARGET	NAME_CONTRACT_TYPE	CODE_GENDER	FLAG_OWN_CAR	FLAG_OWN_REALTY	CNT_CHILDREN	AMT_INCOME_TOTAL	AMT_CREDIT	A
0	100002	1	0	0	0	1	0	0.001512	0.090287	
1	100003	0	0	1	0	0	0	0.002089	0.311736	
2	100004	0	1	0	1	1	0	0.000358	0.022472	
3	100006	0	0	1	0	1	0	0.000935	0.066837	
4	100007	0	0	0	0	1	0	0.000819	0.116854	

```
df['OBS_60_CNT_SOCIAL_CIRCLE'].min()
```

    0.0

```
df['OBS_60_CNT_SOCIAL_CIRCLE'].max()
```

    344.0

```
df['OBS_60_CNT_SOCIAL_CIRCLE']=df['OBS_60_CNT_SOCIAL_CIRCLE'].apply(lambda v:(v-df['OBS_60_CNT_SOCIAL_CIRCLE'].min()))/(df['OBS_60_CNT_S(
df.head()
```

	SK_ID_CURR	TARGET	NAME_CONTRACT_TYPE	CODE_GENDER	FLAG_OWN_CAR	FLAG_OWN_REALTY	CNT_CHILDREN	AMT_INCOME_TOTAL	AMT_CREDIT	A
0	100002	1	0	0	0	1	0	0.001512	0.090287	
1	100003	0	0	1	0	0	0	0.002089	0.311736	
2	100004	0	1	0	1	1	0	0.000358	0.022472	
3	100006	0	0	1	0	1	0	0.000935	0.066837	
4	100007	0	0	0	0	1	0	0.000819	0.116854	

```
df['DEF_60_CNT_SOCIAL_CIRCLE'].min()
```

    0.0

```
df['DEF_60_CNT_SOCIAL_CIRCLE'].max()
```

```
24.0
```

```
df['DEF_60_CNT_SOCIAL_CIRCLE']=df['DEF_60_CNT_SOCIAL_CIRCLE'].apply(lambda v:(v-df['DEF_60_CNT_SOCIAL_CIRCLE'].min()))/(df['DEF_60_CNT_S(
df.head()
```

	SK_ID_CURR	TARGET	NAME_CONTRACT_TYPE	CODE_GENDER	FLAG_OWN_CAR	FLAG_OWN_REALTY	CNT_CHILDREN	AMT_INCOME_TOTAL	AMT_CREDIT A
0	100002	1	0	0	0	1	0	0.001512	0.090287
1	100003	0	0	1	0	0	0	0.002089	0.311736
2	100004	0	1	0	1	1	0	0.000358	0.022472
3	100006	0	0	1	0	1	0	0.000935	0.066837
4	100007	0	0	0	0	1	0	0.000819	0.116854

```
df['DAYS_LAST_PHONE_CHANGE'].min()
```

```
-4292.0
```

```
df['DAYS_LAST_PHONE_CHANGE'].max()
```

```
0.0
```

```
df['DAYS_LAST_PHONE_CHANGE']=df['DAYS_LAST_PHONE_CHANGE'].apply(lambda v:(v-df['DAYS_LAST_PHONE_CHANGE'].min()))/(df['DAYS_LAST_PHONE_CH/
df.head()
```

	SK_ID_CURR	TARGET	NAME_CONTRACT_TYPE	CODE_GENDER	FLAG_OWN_CAR	FLAG_OWN_REALTY	CNT_CHILDREN	AMT_INCOME_TOTAL	AMT_(
0	100002	1	0	0	0	1	0	0.001512	0.(
1	100003	0	0	1	0	0	0	0.002089	0.3
2	100004	0	1	0	1	1	0	0.000358	0.(
3	100006	0	0	1	0	1	0	0.000935	0.(
4	100007	0	0	0	0	1	0	0.000819	0.

```
df['AMT_REQ_CREDIT_BUREAU_HOUR'].min()
```

```
0.0
```

```
df['AMT_REQ_CREDIT_BUREAU_HOUR'].max()
```

```
4.0
```

```
df['AMT_REQ_CREDIT_BUREAU_HOUR']=df['AMT_REQ_CREDIT_BUREAU_HOUR'].apply(lambda v:(v-df['AMT_REQ_CREDIT_BUREAU_HOUR'].min()))/(df['AMT_RE(
df.head()
```

	SK_ID_CURR	TARGET	NAME_CONTRACT_TYPE	CODE_GENDER	FLAG_OWN_CAR	FLAG_OWN_REALTY	CNT_CHILDREN	AMT_INCOME_TOTAL	AMT_(
0	100002	1	0	0	0	1	0	0.001512	0.(
1	100003	0	0	1	0	0	0	0.002089	0.3
2	100004	0	1	0	1	1	0	0.000358	0.(
3	100006	0	0	1	0	1	0	0.000935	0.(
4	100007	0	0	0	0	1	0	0.000819	0.

```
df['AMT_REQ_CREDIT_BUREAU_DAY'].min()
```

```
0.0
```

```
df['AMT_REQ_CREDIT_BUREAU_DAY'].max()
```

```
9.0
```

```
df['AMT_REQ_CREDIT_BUREAU_DAY']=df['AMT_REQ_CREDIT_BUREAU_DAY'].apply(lambda v:(v-df['AMT_REQ_CREDIT_BUREAU_DAY'].min()))/(df['AMT_REQ_CI
df.head()
```

	SK_ID_CURR	TARGET	NAME_CONTRACT_TYPE	CODE_GENDER	FLAG_OWN_CAR	FLAG_OWN_REALTY	CNT_CHILDREN	AMT_INCOME_TOTAL	AMT_C
0	100002	1	0	0	0	1	0	0.001512	0.0
1	100003	0	0	1	0	0	0	0.002089	0.3
2	100004	0	1	0	1	1	0	0.000358	0.0
3	100006	0	0	1	0	1	0	0.000935	0.0
4	100007	0	0	0	0	1	0	0.000819	0.

```
df['AMT_REQ_CREDIT_BUREAU_WEEK'].min()
```

```
0.0
```

```
df['AMT_REQ_CREDIT_BUREAU_WEEK'].max()
```

```
8.0
```

```
df['AMT_REQ_CREDIT_BUREAU_WEEK']=df['AMT_REQ_CREDIT_BUREAU_WEEK'].apply(lambda v:(v-df['AMT_REQ_CREDIT_BUREAU_WEEK'].min()))/(df['AMT_REC
df.head()
```

	SK_ID_CURR	TARGET	NAME_CONTRACT_TYPE	CODE_GENDER	FLAG_OWN_CAR	FLAG_OWN_REALTY	CNT_CHILDREN	AMT_INCOME_TOTAL	AMT_C
0	100002	1	0	0	0	1	0	0.001512	0.0
1	100003	0	0	1	0	0	0	0.002089	0.
2	100004	0	1	0	1	1	0	0.000358	0.0
3	100006	0	0	1	0	1	0	0.000935	0.0
4	100007	0	0	0	0	1	0	0.000819	0.

```
df['AMT_REQ_CREDIT_BUREAU_MON'].min()
```

```
0.0
```

```
df['AMT_REQ_CREDIT_BUREAU_MON'].max()
```

```
27.0
```

```
df['AMT_REQ_CREDIT_BUREAU_MON']=df['AMT_REQ_CREDIT_BUREAU_MON'].apply(lambda v:(v-df['AMT_REQ_CREDIT_BUREAU_MON'].min()))/(df['AMT_REQ_CF
df.head()
```

	SK_ID_CURR	TARGET	NAME_CONTRACT_TYPE	CODE_GENDER	FLAG_OWN_CAR	FLAG_OWN_REALTY	CNT_CHILDREN	AMT_INCOME_TOTAL	AMT_C
0	100002	1	0	0	0	1	0	0.001512	0.0
1	100003	0	0	1	0	0	0	0.002089	0.
2	100004	0	1	0	1	1	0	0.000358	0.0
3	100006	0	0	1	0	1	0	0.000935	0.0
4	100007	0	0	0	0	1	0	0.000819	0.

```
df['AMT_REQ_CREDIT_BUREAU_QRT'].min()
```

```
0.0
```

```
df['AMT_REQ_CREDIT_BUREAU_QRT'].max()
```

```
261.0
```

```
df['AMT_REQ_CREDIT_BUREAU_QRT']=df['AMT_REQ_CREDIT_BUREAU_QRT'].apply(lambda v:(v-df['AMT_REQ_CREDIT_BUREAU_QRT'].min()))/(df['AMT_REQ_CF
df.head()
```

	SK_ID_CURR	TARGET	NAME_CONTRACT_TYPE	CODE_GENDER	FLAG_OWN_CAR	FLAG_OWN_REALTY	CNT_CHILDREN	AMT_INCOME_TOTAL	AMT_(
0	100002	1	0	0	0	1	0	0.001512	0.0

```
df['AMT_REQ_CREDIT_BUREAU_YEAR'].min()
```

```
0.0
```

					.		.		

```
df['AMT_REQ_CREDIT_BUREAU_YEAR'].max()
```

```
25.0
```

```
df['AMT_REQ_CREDIT_BUREAU_YEAR']=df['AMT_REQ_CREDIT_BUREAU_YEAR'].apply(lambda v:(v-df['AMT_REQ_CREDIT_BUREAU_YEAR'].min()))/(df['AMT_RE(
df.head()
```

	SK_ID_CURR	TARGET	NAME_CONTRACT_TYPE	CODE_GENDER	FLAG_OWN_CAR	FLAG_OWN_REALTY	CNT_CHILDREN	AMT_INCOME_TOTAL	AMT_(
0	100002	1	0	0	0	1	0	0.001512	0.0
1	100003	0	0	1	0	0	0	0.002089	0.:
2	100004	0	1	0	1	1	0	0.000358	0.0
3	100006	0	0	1	0	1	0	0.000935	0.0
4	100007	0	0	0	0	1	0	0.000819	0.

```
df.to_csv('Bank_Preprocessed.csv',index=False)
```

```
new_df=pd.read_csv('Bank_Preprocessed.csv',low_memory=False)
new_df.head()
```

	SK_ID_CURR	TARGET	NAME_CONTRACT_TYPE	CODE_GENDER	FLAG_OWN_CAR	FLAG_OWN_REALTY	CNT_CHILDREN	AMT_INCOME_TOTAL	AMT_(
0	100002	1	0	0	0	1	0	0.001512	0.0
1	100003	0	0	1	0	0	0	0.002089	0.:
2	100004	0	1	0	1	1	0	0.000358	0.0
3	100006	0	0	1	0	1	0	0.000935	0.0
4	100007	0	0	0	0	1	0	0.000819	0.

```
pd.set_option('display.max_rows', None)
new_df.corr(numeric_only=True)
```

	SK_ID_CURR	TARGET	NAME_CONTRACT_TYPE	FLAG_O\
SK_ID_CURR	1.000000	-0.002108	0.001654	0.0
TARGET	-0.002108	1.000000	-0.030896	-0.0
NAME_CONTRACT_TYPE	0.001654	-0.030896	1.000000	0.0
FLAG_OWN_CAR	0.001216	-0.021851	0.004022	1.0
FLAG_OWN_REALTY	0.000703	-0.006148	0.067177	-0.0
CNT_CHILDREN	-0.001129	0.019187	0.029998	0.1
AMT_INCOME_TOTAL	-0.001820	-0.003982	-0.003531	0.0
AMT_CREDIT	-0.000343	-0.030369	-0.221648	0.1
AMT_ANNUITY	-0.000433	-0.012817	-0.241544	0.1
AMT_GOODS_PRICE	-0.000232	-0.039645	-0.185850	0.1
NAME_TYPE_SUITE	0.000750	-0.002819	-0.007731	0.0
NAME_INCOME_TYPE	0.002155	-0.057020	-0.032167	-0.1
NAME_EDUCATION_TYPE	0.002452	-0.032434	0.045057	0.0
NAME_FAMILY_STATUS	0.000577	-0.010689	-0.024518	-0.0
NAME_HOUSING_TYPE	0.000909	0.018335	0.011685	-0.0
REGION_POPULATION_RELATIVE	0.000849	-0.037227	0.026867	0.0
DAYS_BIRTH	-0.001500	0.078239	0.086364	0.1
DAYS_EMPLOYED	0.001366	-0.044932	-0.054450	-0.1
DAYS_REGISTRATION	-0.000973	0.041975	0.020557	0.0
DAYS_ID_PUBLISH	-0.000384	0.051457	0.052847	0.0
OWN_CAR_AGE	0.001818	0.037612	0.010602	
FLAG_MOBIL	0.002804	0.000534	0.000585	-0.0
FLAG_EMP_PHONE	-0.001337	0.045982	0.054938	0.1
FLAG_WORK_PHONE	-0.000415	0.028524	-0.033647	0.0
FLAG_CONT_MOBILE	0.002815	0.000370	-0.097116	-0.0
FLAG_PHONE	0.002753	-0.023806	-0.021378	-0.0
FLAG_EMAIL	0.000281	-0.001758	-0.010811	0.0
OCCUPATION_TYPE	0.001332	-0.005134	0.005920	-0.0
CNT_FAM_MEMBERS	-0.002895	0.009308	0.011369	0.1
REGION_RATING_CLIENT	-0.001075	0.058899	-0.021593	-0.0
REGION_RATING_CLIENT_W_CITY	-0.001138	0.060893	-0.023116	-0.0
WEEKDAY_APPR_PROCESS_START	-0.001782	0.000952	0.003542	-0.0
HOUR_APPR_PROCESS_START	0.000350	-0.024166	0.035897	0.0
REG_REGION_NOT_LIVE_REGION	-0.000283	0.005576	0.018561	-0.0
REG_REGION_NOT_WORK_REGION	0.001097	0.006942	0.015774	0.0
LIVE_REGION_NOT_WORK_REGION	0.002903	0.002819	0.007941	0.0
REG_CITY_NOT_LIVE_CITY	-0.001885	0.044395	0.013586	0.0
REG_CITY_NOT_WORK_CITY	-0.001582	0.050994	0.004487	0.0
LIVE_CITY_NOT_WORK_CITY	0.000067	0.032518	-0.004999	0.0
ORGANIZATION_TYPE	0.000379	0.004532	0.019825	0.0
EXT_SOURCE_1	0.000082	-0.155317	-0.013736	-0.0
EXT_SOURCE_2	0.002342	-0.160472	0.015964	0.0
EXT_SOURCE_3	0.000222	-0.178919	-0.005762	-0.0
APARTMENTS_AVG	0.001556	-0.029498	0.013261	0.0
BASEMENTAREA_AVG	-0.002070	-0.022746	0.004230	0.0
YEARS_BEGINEXPLUATATION_AVG	0.001551	-0.009728	-0.001050	0.0
YEARS_BUILD_AVG	0.005900	-0.022149	0.006306	0.0
COMMONAREA_AVG	-0.001463	-0.018550	0.008796	0.0
ELEVATORS_AVG	0.004862	-0.034199	0.016811	0.0

ENTRANCES_AVG	-0.002879	-0.019172	-0.000210	0.0
FLOORSMAX_AVG	0.004851	-0.044003	0.025687	0.0
FLOORSMIN_AVG	0.003083	-0.033614	0.025839	0.0
LANDAREA_AVG	0.001465	-0.010885	0.002087	0.0
LIVINGAPARTMENTS_AVG	0.003119	-0.025031	0.014618	0.0
LIVINGAREA_AVG	0.001770	-0.032997	0.017285	0.0
NONLIVINGAPARTMENTS_AVG	-0.002575	-0.003176	0.006182	0.0
NONLIVINGAREA_AVG	0.003042	-0.013578	0.016373	0.0
APARTMENTS_MODE	0.001961	-0.027284	0.011301	0.0
BASEMENTAREA_MODE	-0.001411	-0.019952	0.003041	0.0
YEARS_BEGINEXPLUATATION_MODE	0.001890	-0.009036	-0.000410	0.0
YEARS_BUILD_MODE	0.005245	-0.022068	0.005701	0.0
COMMONAREA_MODE	-0.001058	-0.016340	0.007634	0.0
ELEVATORS_MODE	0.005017	-0.032131	0.015781	0.0
ENTRANCES_MODE	-0.002844	-0.017387	-0.001150	0.0
FLOORSMAX_MODE	0.004386	-0.043226	0.024052	0.0
FLOORSMIN_MODE	0.002081	-0.032698	0.024382	0.0
LANDAREA_MODE	0.001548	-0.010174	0.001160	0.0
LIVINGAPARTMENTS_MODE	0.003589	-0.023393	0.012839	0.0
LIVINGAREA_MODE	0.002156	-0.030685	0.015886	0.0
NONLIVINGAPARTMENTS_MODE	-0.001923	-0.001557	0.004626	0.0
NONLIVINGAREA_MODE	0.001920	-0.012711	0.013913	0.0
APARTMENTS_MEDI	0.001988	-0.029184	0.012405	0.0
BASEMENTAREA_MEDI	-0.001647	-0.022081	0.003767	0.0
YEARS_BEGINEXPLUATATION_MEDI	0.001366	-0.009993	-0.000585	0.0
YEARS_BUILD_MEDI	0.005777	-0.022326	0.006172	0.0
COMMONAREA_MEDI	-0.001036	-0.018573	0.008707	0.0
ELEVATORS_MEDI	0.005067	-0.033863	0.016455	0.0
ENTRANCES_MEDI	-0.002583	-0.019025	-0.000261	0.0
FLOORSMAX_MEDI	0.004588	-0.043768	0.025363	0.0
FLOORSMIN_MEDI	0.002837	-0.033394	0.025358	0.0
LANDAREA_MEDI	0.001699	-0.011256	0.001798	0.0
LIVINGAPARTMENTS_MEDI	0.003272	-0.024621	0.014088	0.0
LIVINGAREA_MEDI	0.002205	-0.032739	0.016887	0.0
NONLIVINGAPARTMENTS_MEDI	-0.003020	-0.002757	0.005177	0.0
NONLIVINGAREA_MEDI	0.002440	-0.013337	0.015928	0.0
FONDKAPREMONT_MODE	0.002032	-0.008041	-0.001441	-0.0
HOUSETYPE_MODE	0.001874	0.013429	-0.003142	-0.0
TOTALAREA_MODE	0.002288	-0.032596	0.017175	0.0
WALLSMATERIAL_MODE	0.001071	-0.006990	0.000474	0.0
EMERGENCYSTATE_MODE	0.003415	0.012201	-0.003483	-0.0
OBS_30_CNT_SOCIAL_CIRCLE	-0.001409	0.009131	-0.018963	0.0
DEF_30_CNT_SOCIAL_CIRCLE	-0.000082	0.032248	-0.008607	-0.0
OBS_60_CNT_SOCIAL_CIRCLE	-0.001423	0.009022	-0.019141	0.0
DEF_60_CNT_SOCIAL_CIRCLE	0.001187	0.031276	-0.006546	-0.0
DAYS_LAST_PHONE_CHANGE	-0.000858	0.055218	0.061830	-0.0
FLAG_DOCUMENT_2	0.000700	0.005417	-0.002109	-0.0
FLAG_DOCUMENT_3	-0.003411	0.044346	-0.479724	-0.0
FLAG_DOCUMENT_4	-0.004139	-0.002672	0.004448	-0.0
FLAG_DOCUMENT_5	-0.001097	-0.000316	0.029372	-0.0

FLAG_DOCUMENT_6	0.002121	-0.028602	-0.098808	-0.1
FLAG_DOCUMENT_7	-0.002694	-0.001520	0.042703	0.0
FLAG_DOCUMENT_8	0.001809	-0.008040	-0.084961	0.2
FLAG_DOCUMENT_9	0.001505	-0.004352	-0.006592	-0.0

```
import matplotlib.pyplot as plt
%matplotlib inline
```

FLAG_DOCUMENT_11	-0.002012	-0.004229	0.023993	-0.0

```
plt.matshow(new_df.corr(numeric_only=True),cmap='summer')
plt.colorbar()
```

<matplotlib.colorbar.Colorbar at 0x21361b13290>

AMT_REQ_CREDIT_BUREAU_YEAR	0.004659	0.019936	-0.032494	-0.0