Lab-3-Data_Preprocessing

July 15, 2020

Data Preprocessing The sklearn.preprocessing package provides several common utility functions and transformer classes to change raw feature vectors into a representation that is more suitable for the downstream estimators.

- 1. Data Cleaning
- 1.1. Handling missing data
- 1.1.1.Dropping rows that contain missing value

We can drop the rows that contain null or missing data by using dropna(). If we set inplace to True then the original dataset gets modified

```
In [1]: import numpy as np
        import pandas as pd
       df = pd.read_csv('lending_club_loans.csv', low_memory=False)
       df.info()
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 42538 entries, 0 to 42537
Columns: 115 entries, id to total_il_high_credit_limit
dtypes: float64(86), object(29)
memory usage: 37.3+ MB
In [2]: #dropna will not work on NAN and NAT
       df1 = pd.DataFrame({'A':[1,2,3,4,5],'B':[1,2,'NaN',4,5],'C':[1,2,3,'NaT',5]})
       df1.isnull()
Out[2]:
                     В
              Α
       O False False False
       1 False False False
       2 False False False
       3 False False False
       4 False False False
In [3]: #therefore replace with ap.nanm
       df1.replace(["NaN", 'NaT'], np.nan, inplace = True)
       df1.isnull()
```

```
False
          False
                          False
        1 False
                 False
                          False
        2 False
                   True
                          False
        3 False False
                           True
        4 False False
                         False
In [4]: df2= df1.dropna()
        df2
Out [4]:
           Α
                В
              1.0
           1
                   1.0
        1
           2
              2.0
                   2.0
           5 5.0 5.0
In [5]: df.isnull()
Out[5]:
                   id member_id loan_amnt funded_amnt funded_amnt_inv
                                                                               term \
        0
                           False
                                       False
               False
                                                    False
                                                                      False False
        1
               False
                           False
                                       False
                                                    False
                                                                      False False
        2
               False
                           False
                                       False
                                                    False
                                                                      False False
        3
               False
                           False
                                       False
                                                    False
                                                                      False False
        4
               False
                           False
                                       False
                                                    False
                                                                      False False
        5
               False
                           False
                                       False
                                                    False
                                                                      False False
        6
               False
                           False
                                       False
                                                    False
                                                                      False False
        7
               False
                           False
                                       False
                                                    False
                                                                      False False
        8
               False
                           False
                                       False
                                                                      False False
                                                    False
        9
               False
                           False
                                       False
                                                    False
                                                                      False False
        10
               False
                           False
                                       False
                                                    False
                                                                      False False
        11
               False
                           False
                                       False
                                                    False
                                                                      False False
        12
               False
                           False
                                       False
                                                    False
                                                                      False False
                                       False
        13
               False
                           False
                                                    False
                                                                      False False
        14
               False
                           False
                                       False
                                                    False
                                                                      False False
        15
               False
                           False
                                       False
                                                    False
                                                                      False False
                                       False
                                                                      False False
        16
               False
                           False
                                                    False
        17
               False
                           False
                                       False
                                                    False
                                                                      False False
        18
               False
                           False
                                       False
                                                    False
                                                                      False False
        19
               False
                           False
                                       False
                                                    False
                                                                      False False
        20
               False
                           False
                                       False
                                                    False
                                                                      False False
        21
               False
                           False
                                       False
                                                    False
                                                                      False False
        22
               False
                           False
                                       False
                                                                      False False
                                                    False
        23
               False
                           False
                                       False
                                                    False
                                                                      False False
        24
               False
                           False
                                       False
                                                    False
                                                                      False False
        25
               False
                           False
                                       False
                                                                      False False
                                                    False
        26
               False
                           False
                                       False
                                                    False
                                                                      False False
        27
               False
                           False
                                       False
                                                    False
                                                                      False False
        28
               False
                           False
                                       False
                                                    False
                                                                      False False
        29
               False
                           False
                                       False
                                                    False
                                                                      False False
                  . . .
                             . . .
```

Out[3]:

Α

В

С

| 42508 | False | False | False | False | False | False | |
|-------|----------|-------------|-------|-----------|-------|-------|---|
| 42509 | False | False | False | False | False | False | |
| 42510 | False | False | False | False | False | False | |
| 42511 | False | False | False | False | False | False | |
| 42512 | False | False | False | False | False | False | |
| 42513 | False | False | False | False | False | False | |
| 42514 | False | False | False | False | False | False | |
| 42515 | False | False | False | False | False | False | |
| 42516 | False | False | False | False | False | False | |
| 42517 | False | False | False | False | False | False | |
| 42518 | False | False | False | False | False | False | |
| 42519 | False | False | False | False | False | False | |
| 42520 | False | False | False | False | False | False | |
| 42521 | False | False | False | False | False | False | |
| 42522 | False | False | False | False | False | False | |
| 42523 | False | False | False | False | False | False | |
| 42524 | False | False | False | False | False | False | |
| 42525 | False | False | False | False | False | False | |
| 42526 | False | False | False | False | False | False | |
| 42527 | False | False | False | False | False | False | |
| 42528 | | | False | False | | | |
| | False | False | | | False | False | |
| 42529 | False | False | False | False | False | False | |
| 42530 | False | False | False | False | False | False | |
| 42531 | False | False | False | False | False | False | |
| 42532 | False | False | False | False | False | False | |
| 42533 | False | False | False | False | False | False | |
| 42534 | False | False | False | False | False | False | |
| 42535 | False | False | False | False | False | False | |
| 42536 | False | True - | True | True | True | True | |
| 42537 | False | True | True | True | True | True | |
| | | | , | , , | | | , |
| 0 | int_rate | installment | • | sub_grade | • • • | | \ |
| 0 | False | | False | | • • • | | |
| 1 | False | | False | False | • • • | | |
| 2 | False | False | | False | • • • | | |
| 3 | False | False | | False | • • • | | |
| 4 | False | False | | False | | | |
| 5 | False | False | | False | | | |
| 6 | False | False | | False | | | |
| 7 | False | False | | False | | | |
| 8 | False | False | False | False | • • • | | |
| 9 | False | False | False | False | • • • | | |
| 10 | False | False | False | False | | | |
| 11 | False | False | False | False | • • • | | |
| 12 | False | False | False | False | • • • | | |
| 13 | False | False | False | False | | | |
| 14 | False | False | | False | | | |
| 15 | False | False | False | False | | | |
| | | | | | | | |

```
16
                         False False
          False
                                             False
17
          False
                         False
                                False
                                             False
18
          False
                         False
                                False
                                             False
19
                         False
          False
                                False
                                            False
20
          False
                         False False
                                            False
                         False
21
          False
                                False
                                             False
22
          False
                         False False
                                             False
23
          False
                         False False
                                            False
                                                                 . . .
                         False False
24
          False
                                            False
25
          False
                         False
                                False
                                            False
26
                         False
          False
                                False
                                             False
                                                                 . . .
27
          False
                         False
                                False
                                             False
28
          False
                         False
                                False
                                             False
29
          False
                         False
                                False
                                             False
. . .
             . . .
                           . . .
                                               . . .
                                                                 . . .
42508
                         False
          False
                                False
                                             False
42509
          False
                         False
                                False
                                             False
42510
                         False
          False
                                False
                                            False
42511
                         False
                                False
          False
                                            False
42512
          False
                         False
                                False
                                             False
42513
          False
                         False
                                False
                                             False
                         False
42514
          False
                                False
                                            False
                                                                 . . .
42515
          False
                         False False
                                            False
42516
                         False False
          False
                                            False
42517
          False
                         False
                                False
                                            False
42518
          False
                         False
                                False
                                             False
42519
          False
                         False
                                False
                                             False
42520
          False
                         False
                                False
                                             False
42521
          False
                         False
                                False
                                             False
                                                                 . . .
42522
          False
                         False
                                False
                                             False
42523
                         False
          False
                                False
                                             False
42524
          False
                         False
                                False
                                             False
42525
          False
                         False
                                False
                                            False
42526
                         False
          False
                                False
                                            False
42527
          False
                         False False
                                            False
42528
          False
                         False False
                                             False
                                                                 . . .
                         False False
42529
          False
                                             False
42530
          False
                         False False
                                            False
                         False False
42531
          False
                                            False
42532
          False
                         False
                                False
                                            False
42533
                         False
          False
                                False
                                            False
                         False
42534
          False
                                False
                                             False
42535
                         False
          False
                                False
                                             False
42536
            True
                          True
                                  True
                                              True
42537
            True
                          True
                                  True
                                              True
                                                                 . . .
       num_tl_90g_dpd_24m num_tl_op_past_12m pct_tl_nvr_dlq \
```

True

True

True

0

| 1 | True | True | True |
|-------|------|------|------|
| 2 | True | True | True |
| 3 | True | True | True |
| 4 | True | True | True |
| 5 | True | True | True |
| 6 | True | True | True |
| 7 | True | True | True |
| 8 | True | True | True |
| 9 | True | True | True |
| 10 | True | True | True |
| 11 | True | True | True |
| 12 | True | True | True |
| 13 | True | True | True |
| 14 | True | True | True |
| 15 | True | True | True |
| 16 | True | True | True |
| 17 | True | True | True |
| 18 | True | True | True |
| 19 | True | True | True |
| 20 | True | True | True |
| 21 | True | True | True |
| 22 | True | True | True |
| 23 | True | True | True |
| 24 | True | True | True |
| 25 | True | True | True |
| 26 | True | True | True |
| 27 | True | True | True |
| 28 | True | True | True |
| 29 | True | True | True |
| | | | |
| 42508 | True | True | True |
| 42509 | True | True | True |
| 42510 | True | True | True |
| 42511 | True | True | True |
| 42512 | True | True | True |
| 42513 | True | True | True |
| 42514 | True | True | True |
| 42515 | True | True | True |
| 42516 | True | True | True |
| 42517 | True | True | True |
| 42518 | True | True | True |
| 42519 | True | True | True |
| 42520 | True | True | True |
| 42521 | True | True | True |
| 42522 | True | True | True |
| 42523 | True | True | True |
| 42524 | True | True | True |
| 42525 | True | True | True |
| | | | |

| 42526 | Tru | e True | | True | |
|-------|------------------|---------------------------------|-----------|-----------------|---|
| 42527 | Tru | e True | | True | |
| 42528 | Tru | e True | | True | |
| 42529 | Tru | e True | | True | |
| 42530 | Tru | e True | | True | |
| 42531 | Tru | e True | | True | |
| 42532 | Tru | e True | | True | |
| 42533 | Tru | e True | | True | |
| 42534 | Tru | e True | | True | |
| 42535 | Tru | e True | | True | |
| 42536 | Tru | e True | | True | |
| 42537 | Tru | | | True | |
| | | | | | |
| | percent_bc_gt_75 | <pre>pub_rec_bankruptcies</pre> | tax_liens | tot_hi_cred_lim | \ |
| 0 | True | False | False | True | |
| 1 | True | False | False | True | |
| 2 | True | False | False | True | |
| 3 | True | False | False | True | |
| 4 | True | False | False | True | |
| 5 | True | False | False | True | |
| 6 | True | False | False | True | |
| 7 | True | False | False | True | |
| 8 | True | False | False | True | |
| 9 | True | False | False | True | |
| 10 | True | False | False | True | |
| 11 | True | False | False | True | |
| 12 | True | False | False | True | |
| 13 | True | False | False | True | |
| 14 | True | False | False | True | |
| 15 | True | False | False | True | |
| 16 | True | False | False | True | |
| 17 | True | False | False | True | |
| 18 | True | False | False | True | |
| 19 | True | False | False | True | |
| 20 | True | False | False | True | |
| 21 | True | False | False | True | |
| 22 | True | False | False | True | |
| 23 | True | False | False | True | |
| 24 | True | False | False | True | |
| 25 | True | False | False | True | |
| 26 | True | False | False | True | |
| 27 | True | False | False | True | |
| 28 | True | False | False | True | |
| 29 | True | False | False | True | |
| | | | | | |
| 42508 | True | True | True | True | |
| 42509 | True | True | True | True | |
| 42510 | True | True | True | True | |
| | | | | | |

| 42511 | True | | True | True | True |
|---|--|--|----------|--------------|---|
| 42512 | True | | True | True | True |
| 42513 | True | | True | True | True |
| 42514 | True | | True | True | True |
| 42515 | True | | True | True | True |
| 42516 | True | | True | True | True |
| | | | | | |
| 42517 | True | | True | True | True |
| 42518 | True | | True | True | True |
| 42519 | True | | True | True | True |
| 42520 | True | | True | True | True |
| 42521 | True | | True | True | True |
| 42522 | True | | True | True | True |
| 42523 | True | | True | True | True |
| 42524 | True | | True | True | True |
| 42525 | True | | True | True | True |
| 42526 | True | | True | True | True |
| 42527 | True | | True | True | True |
| 42528 | True | | True | True | True |
| 42529 | True | | True | True | True |
| 42529 | | | | | |
| | True | | True | True | True |
| 42531 | True | | True | True | True |
| 42532 | True | | True | True | True |
| 42533 | True | | True | True | True |
| 42534 | True | | True | True | True |
| 42535 | True | | True | True | True |
| 42536 | True | | True | True | True |
| 42537 | True | | True | True | True |
| | | | | | |
| | total_bal_ex_mort | total_bc_limit | total_il | _high_credit | _limit |
| 0 | True | True | _ | | True |
| 1 | True | True | | | True |
| 2 | True | True | | | |
| 3 | | | | | True |
| | True | | | | True True |
| Δ | True | True | | | True |
| 4 | True | True True | | | True True |
| 5 | True True | True True True | | | True True True |
| 5 6 | True True True | True True True True | | | True True True True |
| 5 6 7 | True True True True | True True True True True | | | True True True True True |
| 5 6 7 8 | True True True True True | True True True True True True | | | True True True True True True True |
| 5 6 7 8 9 | True True True True True True | True True True True True True | | | True True True True True True True True |
| 5 6 7 8 9 10 | True True True True True True True | True True True True True True True | | | True True True True True True True True |
| 5 6 7 8 9 10 11 | True True True True True True | True True True True True True | | | True True True True True True True True |
| 5 6 7 8 9 10 | True True True True True True True | True True True True True True True | | | True True True True True True True True |
| 5 6 7 8 9 10 11 | True True True True True True True True | True True True True True True True True | | | True True True True True True True True |
| 5 6 7 8 9 10 11 12 | True True True True True True True True | True True True True True True True True | | | True True True True True True True True |
| 5 6 7 8 9 10 11 12 13 | True True True True True True True True | True True True True True True True True | | | True True True True True True True True |
| 5 6 7 8 9 10 11 12 13 14 | True True True True True True True True | True True True True True True True True | | | True True True True True True True True |
| 5 6 7 8 9 10 11 12 13 14 15 16 | True True True True True True True True | True True True True True True True True | | | True True True True True True True True |
| 5 6 7 8 9 10 11 12 13 14 15 | True True True True True True True True | True True True True True True True True | | | True True True True True True True True |

| 20 True True 21 True True | True |
|---------------------------|------|
| 21 True True | |
| | True |
| 22 True True | True |
| 23 True True | True |
| 24 True True | True |
| 25 True True | True |
| 26 True True | True |
| 27 True True | True |
| 28 True True | True |
| 29 True True | True |
| | |
| 42508 True True | True |
| 42509 True True | True |
| 42510 True True | True |
| 42511 True True | True |
| 42512 True True | True |
| 42513 True True | True |
| 42514 True True | True |
| 42515 True True | True |
| 42516 True True | True |
| 42517 True True | True |
| 42518 True True | True |
| 42519 True True | True |
| 42520 True True | True |
| 42521 True True | True |
| 42522 True True | True |
| 42523 True True | True |
| 42524 True True | True |
| 42525 True True | True |
| 42526 True True | True |
| 42527 True True | True |
| 42528 True True | True |
| 42529 True True | True |
| 42530 True True | True |
| 42531 True True | True |
| 42532 True True | True |
| 42533 True True | True |
| 42534 True True | True |
| 42535 True True | True |
| 42536 True True | True |
| 42537 True True | True |

[42538 rows x 115 columns]

| Out[6]: | | id | member_id | loan_amnt | funded_amnt | funded_amnt_inv | term | \ |
|---------|-------|-------|-----------|-----------|-------------|-----------------|-------|---|
| outloj. | 0 | False | False | False | False | False | | ` |
| | 1 | False | False | False | False | False | False | |
| | 2 | False | False | False | False | False | False | |
| | 3 | False | False | False | False | False | | |
| | 4 | False | False | False | False | False | | |
| | 5 | False | False | False | False | False | | |
| | 6 | False | False | False | False | False | | |
| | 7 | False | False | False | False | False | | |
| | 8 | False | False | False | False | False | | |
| | 9 | False | False | False | False | False | | |
| | 10 | False | False | False | False | False | | |
| | 11 | False | False | False | False | False | | |
| | 12 | False | False | False | False | False | | |
| | 13 | False | False | False | False | False | | |
| | 14 | False | False | False | False | False | | |
| | 15 | False | False | False | False | False | | |
| | 16 | False | False | False | False | False | | |
| | 17 | False | False | False | False | False | False | |
| | 18 | False | False | False | False | False | | |
| | 19 | False | False | False | False | False | | |
| | 20 | False | False | False | False | False | | |
| | 21 | False | False | False | False | False | | |
| | 22 | False | False | False | False | False | False | |
| | 23 | False | False | False | False | False | False | |
| | 24 | False | False | False | False | False | | |
| | 25 | False | False | False | False | False | | |
| | 26 | False | False | False | False | False | | |
| | 27 | False | False | False | False | False | | |
| | 28 | False | False | False | False | False | | |
| | 29 | False | False | False | False | False | | |
| | | raise | raise | raise | raise | raise | raise | |
| | 42508 | False | False | False | False | False | False | |
| | | False | False | False | False | False | | |
| | | False | False | False | False | False | | |
| | | False | False | False | | False | | |
| | | False | False | False | False | False | | |
| | | False | False | False | False | False | | |
| | | False | False | False | False | False | | |
| | | False | False | False | | False | | |
| | | False | False | False | | False | | |
| | | False | False | False | False | False | | |
| | 42517 | False | False | False | False | False | | |
| | 42519 | False | False | False | False | False | | |
| | 42520 | False | False | False | False | False | | |
| | | False | False | False | False | False | | |
| | | False | False | False | | False | | |
| | | False | False | False | False | False | | |
| | +2020 | raise | гатъе | гатъе | гатае | raise | raise | |

| 42524 | False | False | False | False | False | False | |
|-------|----------|-------------|-------|-----------|-------|-------|---|
| 42525 | False | False | False | False | False | False | |
| 42526 | False | False | False | False | False | False | |
| 42527 | False | False | False | False | False | False | |
| 42528 | False | False | False | False | False | False | |
| 42529 | False | False | False | False | False | False | |
| 42530 | False | False | False | False | False | False | |
| 42531 | False | False | False | False | False | False | |
| 42532 | False | False | False | False | False | False | |
| 42533 | False | False | False | False | False | False | |
| 42534 | False | False | False | False | False | False | |
| 42535 | False | False | False | False | False | False | |
| 42536 | False | True | True | True | True | True | |
| 42537 | False | True | True | True | True | True | |
| | | | | | | | |
| | int_rate | installment | grade | sub_grade | | | \ |
| 0 | False | False | False | False | | | |
| 1 | False | False | False | False | | | |
| 2 | False | False | False | False | | | |
| 3 | False | False | False | False | | | |
| 4 | False | False | False | False | | | |
| 5 | False | False | False | False | | | |
| 6 | False | False | False | False | | | |
| 7 | False | False | False | False | | | |
| 8 | False | False | False | False | • • • | | |
| 9 | False | False | False | False | | | |
| 10 | False | False | False | False | • • • | | |
| 11 | False | False | False | False | | | |
| 12 | False | False | False | False | | | |
| 13 | False | False | False | False | | | |
| 14 | False | False | False | False | | | |
| 15 | False | False | False | False | | | |
| 16 | False | False | False | False | | | |
| 17 | False | False | False | False | | | |
| 18 | False | False | False | False | | | |
| 19 | False | False | False | False | | | |
| 20 | False | False | False | False | | | |
| 21 | False | False | False | False | | | |
| 22 | False | False | False | False | | | |
| 23 | False | False | False | False | | | |
| 24 | False | False | False | False | | | |
| 25 | False | False | | False | | | |
| 26 | False | False | False | False | • • • | | |
| 27 | False | False | | False | • • • | | |
| 28 | False | False | False | False | • • • | | |
| 29 | False | False | False | False | • • • | | |
| • • • | • • • | | | • • • | • • • | | |
| 42508 | False | False | False | False | • • • | | |
| | | | | | | | |

| 42509 | False | False | False | False | | |
|-------|--------------|------------|---------|-----------|----------------|---|
| 42510 | False | False | False | False | | |
| 42511 | False | False | False | False | | |
| 42512 | False | False | False | False | | |
| 42513 | False | False | False | False | | |
| 42514 | False | False | False | False | | |
| 42515 | False | False | False | False | | |
| 42516 | False | False | False | False | | |
| 42517 | False | False | False | False | | |
| 42518 | False | False | False | False | | |
| 42519 | False | False | False | False | | |
| 42520 | False | False | False | False | | |
| 42521 | False | False | False | False | | |
| 42522 | False | False | False | False | | |
| 42523 | False | False | False | False | | |
| 42524 | False | False | False | False | | |
| 42525 | False | False | False | False | | |
| 42526 | False | False | False | False | | |
| 42527 | False | False | False | False | | |
| 42528 | False | False | False | False | | |
| 42529 | False | False | False | False | | |
| 42530 | False | False | False | False | | |
| 42531 | False | False | False | False | | |
| 42532 | False | False | False | False | | |
| 42533 | False | False | False | False | | |
| 42534 | False | False | False | False | | |
| 42535 | False | False | False | False | | |
| 42536 | True | True | True | True | | |
| 42537 | True | True | True | True | | |
| | | | | | | |
| | num_t1_90g_d | dpd_24m nu | m_tl_op | _past_12m | pct_tl_nvr_dlq | \ |
| 0 | | True | | True | True | |
| 1 | | True | | True | True | |
| 2 | | True | | True | True | |
| 3 | | True | | True | True | |
| 4 | | True | | True | True | |
| 5 | | True | | True | True | |
| 6 | | True | | True | True | |
| 7 | | True | | True | True | |
| 8 | | True | | True | True | |
| 9 | | True | | True | True | |
| 10 | | True | | True | True | |
| 11 | | True | | True | True | |
| 12 | | True | | True | True | |
| 13 | | True | | True | True | |
| 14 | | True | | True | True | |
| 15 | | True | | True | True | |
| 16 | | True | | True | True | |
| | | 1140 | | 1140 | 11 46 | |

| 17 | True | e True | • | True |
|-------|------------------|---------------------------------|-----------|-------------------|
| 18 | True | e True | | True |
| 19 | True | e True | • | True |
| 20 | True | e True | • | True |
| 21 | True | e True | • | True |
| 22 | True | e True | | True |
| 23 | True | e True | | True |
| 24 | True | e True | | True |
| 25 | True | | | True |
| 26 | True | | | True |
| 27 | True | | | True |
| 28 | True | | | True |
| 29 | True | | | True |
| | ••• | | | |
| 42508 | True | | | True |
| 42509 | True | | | True |
| 42510 | True | | | True |
| 42511 | True | | | True |
| 42512 | True | | | True |
| 42513 | True | | | True |
| 42514 | True | | | True |
| 42514 | True | | | True |
| 42516 | | | | |
| 42517 | Truo Truo | | | True |
| | | | | True |
| 42518 | True | | | True |
| 42519 | True | | | True |
| 42520 | True | | | True |
| 42521 | True | | | True |
| 42522 | True | | | True |
| 42523 | True | | | True |
| 42524 | True | | | True |
| 42525 | True | | | True |
| 42526 | True | | | True |
| 42527 | True | | | True |
| 42528 | True | | | True |
| 42529 | True | | | True |
| 42530 | True | | | True |
| 42531 | True | | | True |
| 42532 | True | | | True |
| 42533 | True | | | True |
| 42534 | True | e True | • | True |
| 42535 | True | e True | • | True |
| 42536 | True | e True | • | True |
| 42537 | True | e True | | True |
| | | | | |
| | percent_bc_gt_75 | <pre>pub_rec_bankruptcies</pre> | tax_liens | tot_hi_cred_lim \ |
| 0 | True | False | False | True |
| 1 | True | False | False | True |
| | | | | |

| 2 | True | False | False | True |
|-------|---------|-------|-------|------|
| 3 | True | False | False | True |
| 4 | True | False | False | True |
| 5 | True | False | False | True |
| 6 | True | False | False | True |
| 7 | True | False | False | True |
| 8 | True | False | False | True |
| 9 | True | False | False | True |
| 10 | True | False | False | True |
| 11 | True | False | False | True |
| 12 | True | False | False | True |
| 13 | True | False | False | True |
| 14 | True | False | False | True |
| 15 | True | False | False | True |
| 16 | True | False | False | True |
| 17 | True | False | False | True |
| 18 | True | False | False | True |
| 19 | True | False | False | True |
| 20 | True | False | False | True |
| 21 | True | False | False | True |
| 22 | True | False | False | True |
| 23 | True | False | False | True |
| 24 | True | False | False | True |
| 25 | True | False | False | True |
| 26 | True | False | False | True |
| 27 | True | False | False | True |
| 28 | True | False | False | True |
| 29 | True | False | False | True |
| | • • • • | | | |
| 42508 | True | True | True | True |
| 42509 | True | True | True | True |
| 42510 | True | True | True | True |
| 42511 | True | True | True | True |
| 42512 | True | True | True | True |
| 42513 | True | True | True | True |
| 42514 | True | True | True | True |
| 42515 | True | True | True | True |
| 42516 | True | True | True | True |
| 42517 | True | True | True | True |
| 42518 | True | True | True | True |
| 42519 | True | True | True | True |
| 42520 | True | True | True | True |
| 42521 | True | True | True | True |
| 42522 | True | True | True | True |
| 42523 | True | True | True | True |
| 42524 | True | True | True | True |
| 42525 | True | True | True | True |
| 42526 | True | True | True | True |
| 72020 | irue | iiue | 11 ue | irue |

| 42527 | True | | True | True | | True |
|----------------------|-------------------|----------------|----------|------------------------|-------|------|
| | | | | | | |
| 42528 | True | | True | True | | True |
| 42529 | True | | True | True | | True |
| 42530 | True | | True | True | | True |
| 42531 | True | | True | True | | True |
| 42532 | True | | True | True | | True |
| 42533 | True | | True | True | | True |
| 42534 | True | | True | True | | True |
| 42535 | True | | True | True | | True |
| 42536 | True | | True | True | | True |
| 42537 | True | | True | True | | True |
| | | | | | | |
| | total_bal_ex_mort | total_bc_limit | total_il | $_{	t high_credit_}$ | limit | |
| 0 | True | True | | | True | |
| 1 | True | True | | | True | |
| 2 | True | True | | | True | |
| 3 | True | True | | | True | |
| 4 | True | True | | | True | |
| 5 | True | True | | | True | |
| 6 | True | True | | | True | |
| 7 | True | True | | | True | |
| 8 | True | True | | | True | |
| 9 | True | True | | | True | |
| 10 | True | True | | | True | |
| 11 | True | True | | | True | |
| 12 | True | True | | | True | |
| 13 | True | True | | | True | |
| 14 | True | True | | | True | |
| 15 | True | True | | | True | |
| 16 | True | True | | | True | |
| 17 | True | True | | | True | |
| 18 | True | True | | | True | |
| 19 | True | True | | | True | |
| 20 | True | True | | | True | |
| 21 | True | True | | | True | |
| 22 | True | True | | | True | |
| 23 | True | True | | | True | |
| 23 24 | True | True | | | True | |
| 2 4 25 | True | True | | | True | |
| | | | | | | |
| 26 | True | True | | | True | |
| 27 | True | True | | | True | |
| 28 | True | True | | | True | |
| 29 | True | True | | | True | |
| 40500 | | | | | | |
| 42508 | True | True | | | True | |
| 42509 | True | True | | | True | |
| 42510 | True | True | | | True | |
| 42511 | True | True | | | True | |

```
42512
                              True
                                                True
                                                                              True
        42513
                              True
                                                                              True
                                               True
        42514
                              True
                                               True
                                                                              True
        42515
                              True
                                               True
                                                                              True
                                                                              True
        42516
                              True
                                               True
        42517
                              True
                                                True
                                                                              True
        42518
                              True
                                               True
                                                                              True
        42519
                              True
                                                True
                                                                              True
        42520
                              True
                                                True
                                                                              True
                                                                              True
        42521
                              True
                                                True
        42522
                                                                              True
                              True
                                               True
        42523
                              True
                                               True
                                                                              True
                                                                              True
        42524
                              True
                                               True
                                                                              True
        42525
                              True
                                               True
        42526
                              True
                                               True
                                                                              True
        42527
                              True
                                               True
                                                                              True
        42528
                              True
                                               True
                                                                              True
        42529
                              True
                                               True
                                                                              True
        42530
                              True
                                               True
                                                                              True
        42531
                              True
                                               True
                                                                              True
        42532
                              True
                                               True
                                                                              True
                                                                              True
        42533
                              True
                                                True
        42534
                              True
                                               True
                                                                              True
        42535
                              True
                                                True
                                                                              True
        42536
                              True
                                                True
                                                                              True
                                                                              True
        42537
                              True
                                                True
        [42538 rows x 115 columns]
In [7]: new_df = df.dropna(axis=0,how='any')
        # comparing sizes of data frames
        print("Old data frame length:", len(df), "\nNew data frame length:",
                len(new_df), "\nNumber of rows with at least 1 NA value: ",
                (len(df)-len(new_df)))
Old data frame length: 42538
New data frame length: 0
Number of rows with at least 1 NA value:
```

Always dropping rows that contain any null values may not be a good strategy as we may miss the randomness is the dataset So, what is the next best option to handle missing value?

1.1.2. Setting threshold for null values in a row

we can set a threshold count and if a rows exceeds the threshold count for null values then we can drop the row. We will drop a row from data_1 only if a row contains 50% of the data as null values

```
In [8]: half_count = len(df) / 2
        df.dropna(thresh=half_count, axis=1,inplace=True) # Drop any column with more than 50%
        df
Out[8]:
                                                                 id member_id
                                                                                 loan_amnt
        0
                                                            1077501
                                                                      1296599.0
                                                                                     5000.0
        1
                                                            1077430
                                                                      1314167.0
                                                                                     2500.0
        2
                                                            1077175
                                                                     1313524.0
                                                                                     2400.0
        3
                                                            1076863
                                                                     1277178.0
                                                                                    10000.0
        4
                                                            1075358
                                                                      1311748.0
                                                                                     3000.0
        5
                                                            1075269
                                                                      1311441.0
                                                                                     5000.0
        6
                                                            1069639
                                                                     1304742.0
                                                                                     7000.0
        7
                                                            1072053
                                                                     1288686.0
                                                                                     3000.0
        8
                                                            1071795
                                                                      1306957.0
                                                                                     5600.0
        9
                                                            1071570
                                                                     1306721.0
                                                                                     5375.0
        10
                                                            1070078
                                                                      1305201.0
                                                                                     6500.0
        11
                                                            1069908
                                                                      1305008.0
                                                                                    12000.0
        12
                                                                                     9000.0
                                                            1064687
                                                                      1298717.0
        13
                                                            1069866
                                                                      1304956.0
                                                                                     3000.0
        14
                                                            1069057
                                                                      1303503.0
                                                                                    10000.0
        15
                                                            1069759
                                                                      1304871.0
                                                                                     1000.0
        16
                                                            1065775
                                                                     1299699.0
                                                                                    10000.0
        17
                                                            1069971
                                                                      1304884.0
                                                                                     3600.0
        18
                                                            1062474
                                                                     1294539.0
                                                                                     6000.0
        19
                                                            1069742
                                                                      1304855.0
                                                                                     9200.0
        20
                                                            1069740
                                                                      1284848.0
                                                                                    20250.0
        21
                                                            1039153
                                                                     1269083.0
                                                                                    21000.0
        22
                                                            1069710
                                                                      1304821.0
                                                                                    10000.0
        23
                                                            1069700
                                                                      1304810.0
                                                                                    10000.0
        24
                                                            1069559
                                                                      1304634.0
                                                                                     6000.0
        25
                                                                      1273773.0
                                                                                    15000.0
                                                            1069697
        26
                                                            1069800
                                                                      1304679.0
                                                                                    15000.0
        27
                                                                      1304764.0
                                                                                     5000.0
                                                            1069657
        28
                                                            1069799
                                                                      1304678.0
                                                                                     4000.0
        29
                                                            1047704
                                                                      1278806.0
                                                                                     8500.0
                                                                . . .
                                                                                        . . .
        42508
                                                              91175
                                                                        91170.0
                                                                                     6000.0
        42509
                                                              91126
                                                                        91067.0
                                                                                     5350.0
        42510
                                                                                     1900.0
                                                              91023
                                                                        70879.0
        42511
                                                              90106
                                                                        90090.0
                                                                                    10000.0
        42512
                                                              89258
                                                                        80039.0
                                                                                     2000.0
        42513
                                                              88637
                                                                        88629.0
                                                                                     6000.0
        42514
                                                              88046
                                                                        88023.0
                                                                                     4400.0
        42515
                                                              85961
                                                                        85923.0
                                                                                     1200.0
        42516
                                                              85818
                                                                        85802.0
                                                                                     5000.0
        42517
                                                              85781
                                                                        85727.0
                                                                                     1400.0
        42518
                                                              85675
                                                                        85667.0
                                                                                     1000.0
        42519
                                                              84670
                                                                        79576.0
                                                                                     5000.0
```

```
84098
                                                                84091.0
                                                                             2500.0
42520
42521
                                                      83979
                                                                83974.0
                                                                             3000.0
42522
                                                                             2600.0
                                                      83489
                                                                83471.0
42523
                                                                83132.0
                                                                             1000.0
                                                      83185
42524
                                                      76629
                                                                76623.0
                                                                             1275.0
42525
                                                      74014
                                                                73890.0
                                                                             6450.0
42526
                                                      81085
                                                                80973.0
                                                                            10500.0
42527
                                                      77792
                                                                77764.0
                                                                             3000.0
42528
                                                      77757
                                                                70626.0
                                                                             3000.0
42529
                                                      74505
                                                                74469.0
                                                                             2000.0
42530
                                                      74323
                                                                74301.0
                                                                             6500.0
42531
                                                      73582
                                                                73096.0
                                                                             3500.0
                                                      72998
                                                                72992.0
42532
                                                                             1000.0
42533
                                                      72176
                                                                70868.0
                                                                             2525.0
42534
                                                      71623
                                                                70735.0
                                                                             6500.0
42535
                                                                70681.0
                                                      70686
                                                                             5000.0
42536
       Total amount funded in policy code 1: 460296150
                                                                    NaN
                                                                                 NaN
42537
                Total amount funded in policy code 2: 0
                                                                    NaN
                                                                                 NaN
       funded amnt
                      funded_amnt_inv
                                               term int_rate
                                                                installment grade
                                                       10.65%
0
             5000.0
                          4975.000000
                                          36 months
                                                                      162.87
                                                                                  В
                                                                                  С
1
             2500.0
                          2500.000000
                                          60 months
                                                       15.27%
                                                                       59.83
2
             2400.0
                          2400.000000
                                          36 months
                                                       15.96%
                                                                       84.33
                                                                                  C
3
                                                                                  C
            10000.0
                         10000.000000
                                          36 months
                                                       13.49%
                                                                      339.31
4
                          3000.000000
                                          60 months
                                                       12.69%
                                                                       67.79
                                                                                  В
             3000.0
5
                          5000.000000
                                          36 months
                                                        7.90%
                                                                                  Α
             5000.0
                                                                      156.46
6
                                                                                  C
             7000.0
                          7000.000000
                                          60 months
                                                       15.96%
                                                                      170.08
7
                                                                                  Е
             3000.0
                          3000.000000
                                          36 months
                                                       18.64%
                                                                      109.43
                                                                                  F
8
             5600.0
                          5600.000000
                                          60 months
                                                       21.28%
                                                                      152.39
9
             5375.0
                          5350.000000
                                          60 months
                                                       12.69%
                                                                                  В
                                                                      121.45
10
                                          60 months
                                                                                  C
             6500.0
                          6500.000000
                                                       14.65%
                                                                      153.45
11
            12000.0
                         12000.000000
                                          36 months
                                                       12.69%
                                                                      402.54
                                                                                  В
12
             9000.0
                          9000.000000
                                          36 months
                                                       13.49%
                                                                      305.38
                                                                                  C
                          3000.000000
                                                                                  В
13
             3000.0
                                          36 months
                                                        9.91%
                                                                       96.68
                                                                                  В
            10000.0
                         10000.000000
                                          36 months
                                                       10.65%
14
                                                                      325.74
15
             1000.0
                          1000.000000
                                          36 months
                                                       16.29%
                                                                       35.31
                                                                                  D
                                                                                  C
16
            10000.0
                         10000.000000
                                          36 months
                                                       15.27%
                                                                      347.98
                                          36 months
                                                        6.03%
17
             3600.0
                          3600.000000
                                                                      109.57
                                                                                  Α
                                          36 months
                                                                                  В
18
             6000.0
                          6000.000000
                                                       11.71%
                                                                      198.46
19
             9200.0
                          9200.000000
                                          36 months
                                                        6.03%
                                                                                  Α
                                                                      280.01
                                                                                  C
20
            20250.0
                         19142.161077
                                          60 months
                                                       15.27%
                                                                      484.63
                                                                                  В
21
            21000.0
                         21000.000000
                                          36 months
                                                       12.42%
                                                                      701.73
22
                                          36 months
                                                                                  В
            10000.0
                         10000.000000
                                                       11.71%
                                                                      330.76
                                                                                  В
23
            10000.0
                         10000.000000
                                          36 months
                                                       11.71%
                                                                      330.76
24
             6000.0
                          6000.000000
                                             months
                                                       11.71%
                                                                      198.46
                                                                                  В
25
            15000.0
                         15000.000000
                                          36 months
                                                        9.91%
                                                                      483.38
                                                                                  В
26
            15000.0
                          8725.000000
                                          36 months
                                                       14.27%
                                                                      514.64
                                                                                  C
27
             5000.0
                          5000.000000
                                          60 months
                                                                                  D
                                                       16.77%
                                                                      123.65
```

| 28 | 4000.0 | 4000.000 | 000 36 | months | 11.71% | 132.31 | В |
|----------------|------------------|----------|------------|---------------|------------|-------------|-------|
| 29 | 8500.0 | | | months | | 281.15 | В |
| | | | 000 30 | montans | | | |
| 42508 | 6000.0 | | 000 36 | months | 13.12% | 202.51 | D |
| 42509 | 5350.0 | | | months | | 180.57 | D |
| 42510 | 1900.0 | | | months | | 61.00 | В |
| 42511 | 10000.0 | | | months | | 345.18 | E |
| 42512 | 2000.0 | | | months | | 61.87 | A |
| 42513 | 6000.0 | | | months | | 195.28 | C |
| 42513 | 4400.0 | | | months | | 141.25 | В |
| 42514 | 1200.0 | | | months | | 38.17 | В |
| 42516 | 5000.0 | | | months | | 164.23 | С |
| | | | | | | | C |
| 42517 | 1400.0 1000.0 | | | months | | 45.78 | |
| 42518 42519 | | | | months | | 34.21 | E |
| | 5000.0 2500.0 | | | months months | | 156.11 | A |
| 42520 | | | | | | 77.69 | A |
| 42521 | 3000.0 | | | months | | 93.23 | A |
| 42522 | 2600.0 | | | months | | 81.94 | A |
| 42523 | 1000.0 | | | months | | 30.94 | A |
| 42524 | 1275.0 | | | months | • • | 42.65 | D |
| 42525 | 6450.0 | | | months | | 211.85 | C |
| 42526 | 10500.0 | | | months | | 344.87 | C |
| 42527 | 3000.0 | | | months | | 95.42 | В |
| 42528 | 3000.0 | | | months | | 95.86 | В |
| 42529 | 2000.0 | | | months | | 64.50 | В |
| 42530 | 6500.0 | | | months | | 208.66 | В |
| 42531 | 3500.0 | | | months | | 113.39 | C |
| 42532 | 1000.0 | | | months | | 32.11 | В |
| 42533 | 2525.0 | | | months | | 80.69 | В |
| 42534 | 6500.0 | | | months | | 204.84 | A |
| 42535 | 5000.0 | | | months | | 156.11 | A |
| 42536 | NaN | | NaN | NaN | | NaN | NaN |
| 42537 | NaN | V . | NaN | NaN | NaN | NaN | NaN |
| | sub_grade | lagt f | ico rang | a high | last fico | range_low \ | |
| 0 | B2 | | rco_rang | 744.0 | 1450_1100_ | 740.0 | |
| 1 | C4 | • • • | | 499.0 | | 0.0 | |
| 2 | C5 | • • • | | 719.0 | | 715.0 | |
| 3 | C1 | • • • | | 604.0 | | 600.0 | |
| 4 | B5 | • • • | | 694.0 | | 690.0 | |
| 5 | A4 | • • • | | 679.0 | | 675.0 | |
| 6 | C5 | • • • | | 654.0 | | 650.0 | |
| 7 | E1 | • • • | | 689.0 | | 685.0 | |
| 8 | F2 | • • • | | 499.0 | | 0.0 | |
| 9 | ь В5 | • • • | | 519.0 | | 515.0 | |
| 9 10 | C3 | • • • | | 734.0 | | 730.0 | |
| 10 | B5 | • • • | | 669.0 | | 665.0 | |
| 12 | ъз С1 | • • • | | 619.0 | | 615.0 | |
| 14 | CI | • • • | | 019.0 | | 010.0 | |

| 13 | B1 | | 734.0 | 730.0 |
|-------|------------|-------|------------|-------|
| 14 | B2 | | 654.0 | 650.0 |
| 15 | D1 | | 769.0 | 765.0 |
| 16 | C4 | | 639.0 | 635.0 |
| 17 | A1 | | 794.0 | 790.0 |
| 18 | В3 | | 559.0 | 555.0 |
| 19 | A1 | | 644.0 | 640.0 |
| 20 | C4 | | 684.0 | 680.0 |
| 21 | B4 | | 534.0 | 530.0 |
| 22 | В3 | | 779.0 | 775.0 |
| 23 | В3 | | 684.0 | 680.0 |
| 24 | В3 | | 499.0 | 0.0 |
| 25 | B1 | | 734.0 | 730.0 |
| 26 | C2 | | 509.0 | 505.0 |
| 27 | D2 | | 579.0 | 575.0 |
| 28 | В3 | | 689.0 | 685.0 |
| 29 | В3 | | 599.0 | 595.0 |
| | | | | |
| 42508 | D5 | | 664.0 | 660.0 |
| 42509 | D5 | | 704.0 | 700.0 |
| 42510 | B4 | | 709.0 | 705.0 |
| 42511 | E5 | | 499.0 | 0.0 |
| 42512 | A1 | • • • | 809.0 | 805.0 |
| 42513 | C2 | • • • | 769.0 | 765.0 |
| 42514 | B4 | • • • | 584.0 | 580.0 |
| 42515 | B2 | • • • | 694.0 | 690.0 |
| 42516 | C4 | • • • | 679.0 | 675.0 |
| 42517 | C3 | • • • | 679.0 | 675.0 |
| 42518 | E3 | • • • | 499.0 | 0.0 |
| 42519 | A3 | • • • | 759.0 | 755.0 |
| 42520 | A2 | • • • | 789.0 | 785.0 |
| 42521 | A2 | • • • | 799.0 | 795.0 |
| 42522 | A5 | ••• | 614.0 | 610.0 |
| 42523 | AJ | • • • | 599.0 | 595.0 |
| 42524 | D3 | • • • | 624.0 | 620.0 |
| 42525 | C4 | • • • | 574.0 | 570.0 |
| 42526 | C4 | • • • | 744.0 | 740.0 |
| 42527 | B2 | • • • | 594.0 | 590.0 |
| 42528 | B3 | • • • | 714.0 | 710.0 |
| 42529 | B5 | • • • | 594.0 | 590.0 |
| 42529 | В3 В4 | • • • | 684.0 | 680.0 |
| | C1 | • • • | 819.0 | 815.0 |
| 42531 | | • • • | | |
| 42532 | B4 | • • • | 784.0 | 780.0 |
| 42533 | B3 | • • • | 714.0 | 710.0 |
| 42534 | A5 | • • • | 724.0 | 720.0 |
| 42535 | A3 NaN | • • • | 794.0 | 790.0 |
| 42536 | NaN NaN | • • • | NaN NaN | NaN |
| 42537 | NaN | • • • | NaN | NaN |
| | | | | |

| | collections_12_mths_ex_med | policy code | application type | acc now deling | \ |
|-------|----------------------------|-------------|------------------|----------------|---|
| 0 | False | True | INDIVIDUAL | False | |
| 1 | False | True | INDIVIDUAL | False | |
| 2 | False | True | INDIVIDUAL | False | |
| 3 | False | True | INDIVIDUAL | False | |
| 4 | False | True | INDIVIDUAL | False | |
| 5 | False | True | INDIVIDUAL | False | |
| 6 | False | True | INDIVIDUAL | False | |
| 7 | False | True | INDIVIDUAL | False | |
| 8 | False | True | INDIVIDUAL | False | |
| 9 | False | True | INDIVIDUAL | False | |
| 10 | False | True | INDIVIDUAL | False | |
| 11 | False | True | INDIVIDUAL | False | |
| 12 | False | True | INDIVIDUAL | False | |
| 13 | False | True | INDIVIDUAL | False | |
| 14 | False | True | INDIVIDUAL | False | |
| 15 | False | True | INDIVIDUAL | False | |
| 16 | False | True | INDIVIDUAL | False | |
| 17 | False | True | INDIVIDUAL | False | |
| 18 | False | True | INDIVIDUAL | False | |
| 19 | False | True | INDIVIDUAL | False | |
| 20 | False | True | INDIVIDUAL | False | |
| 21 | False | True | INDIVIDUAL | False | |
| 22 | False | True | INDIVIDUAL | False | |
| 23 | False | True | INDIVIDUAL | False | |
| 24 | False | True | INDIVIDUAL | False | |
| 25 | False | True | INDIVIDUAL | False | |
| 26 | False | True | INDIVIDUAL | False | |
| 27 | False | True | INDIVIDUAL | False | |
| 28 | False | True | INDIVIDUAL | False | |
| 29 | False | True | INDIVIDUAL | False | |
| | ••• | | | | |
| 42508 | NaN | True | INDIVIDUAL | False | |
| 42509 | NaN | True | INDIVIDUAL | False | |
| 42510 | NaN | True | INDIVIDUAL | NaN | |
| 42511 | . NaN | True | INDIVIDUAL | False | |
| 42512 | NaN | True | INDIVIDUAL | False | |
| 42513 | NaN | True | INDIVIDUAL | False | |
| 42514 | . NaN | True | INDIVIDUAL | False | |
| 42515 | NaN | True | INDIVIDUAL | NaN | |
| 42516 | NaN | True | INDIVIDUAL | NaN | |
| 42517 | NaN | True | INDIVIDUAL | NaN | |
| 42518 | NaN | True | INDIVIDUAL | NaN | |
| 42519 | NaN | True | INDIVIDUAL | NaN | |
| 42520 | NaN | True | INDIVIDUAL | NaN | |
| 42521 | NaN | True | INDIVIDUAL | NaN | |
| 42522 | NaN | True | INDIVIDUAL | NaN | |

| 42523 | Na | aN Tr | ue INDIVIDUAL | NaN |
|-------|--------------------------|-------------|---------------------------------|-----------|
| 42524 | Na | aN Tr | ue INDIVIDUAL | NaN |
| 42525 | Na | aN Tr | ue INDIVIDUAL | NaN |
| 42526 | Na | aN Tr | ue INDIVIDUAL | NaN |
| 42527 | | aN Tr | | NaN |
| | | | | |
| 42528 | | | | NaN |
| 42529 | | | | NaN |
| 42530 | Na | aN Tr | ue INDIVIDUAL | NaN |
| 42531 | Na | aN Tr | ue INDIVIDUAL | NaN |
| 42532 | Na | aN Tr | ue INDIVIDUAL | NaN |
| 42533 | Na | aN Tr | ue INDIVIDUAL | NaN |
| 42534 | | | | NaN |
| | | | | |
| 42535 | | | | NaN |
| 42536 | | | aN NaN | NaN |
| 42537 | Na | aN Ne | aN NaN | NaN |
| | | | | |
| | chargeoff_within_12_mths | delinq_amnt | <pre>pub_rec_bankruptcies</pre> | tax_liens |
| 0 | False | 0.0 | 0.0 | False |
| 1 | False | 0.0 | 0.0 | False |
| | | | | |
| 2 | False | 0.0 | 0.0 | False |
| 3 | False | 0.0 | 0.0 | False |
| 4 | False | 0.0 | 0.0 | False |
| 5 | False | 0.0 | 0.0 | False |
| 6 | False | 0.0 | 0.0 | False |
| 7 | False | 0.0 | 0.0 | False |
| 8 | False | 0.0 | 0.0 | False |
| 9 | False | 0.0 | 0.0 | |
| | | | | False |
| 10 | False | 0.0 | 0.0 | False |
| 11 | False | 0.0 | 0.0 | False |
| 12 | False | 0.0 | 0.0 | False |
| 13 | False | 0.0 | 0.0 | False |
| 14 | False | 0.0 | 0.0 | False |
| 15 | False | 0.0 | 0.0 | False |
| 16 | False | 0.0 | 0.0 | False |
| 17 | False | 0.0 | 0.0 | |
| | | | | False |
| 18 | False | 0.0 | 0.0 | False |
| 19 | False | 0.0 | 0.0 | False |
| 20 | False | 0.0 | 0.0 | False |
| 21 | False | 0.0 | 0.0 | False |
| 22 | False | 0.0 | 0.0 | False |
| 23 | False | 0.0 | 0.0 | False |
| 24 | False | 0.0 | 0.0 | False |
| | | | | |
| 25 | False | 0.0 | 0.0 | False |
| 26 | False | 0.0 | 0.0 | False |
| 27 | False | 0.0 | 0.0 | False |
| 28 | False | 0.0 | 0.0 | False |
| 29 | False | 0.0 | 0.0 | False |
| | | | | |

. . .

| 42508 | NaN | 0.0 | NaN | NaN |
|-------|-----|-----|-----|-----|
| 42509 | NaN | 0.0 | NaN | NaN |
| 42510 | NaN | NaN | NaN | NaN |
| 42511 | NaN | 0.0 | NaN | NaN |
| 42512 | NaN | 0.0 | NaN | NaN |
| 42513 | NaN | 0.0 | NaN | NaN |
| 42514 | NaN | 0.0 | NaN | NaN |
| 42515 | NaN | NaN | NaN | NaN |
| 42516 | NaN | NaN | NaN | NaN |
| 42517 | NaN | NaN | NaN | NaN |
| 42518 | NaN | NaN | NaN | NaN |
| 42519 | NaN | NaN | NaN | NaN |
| 42520 | NaN | NaN | NaN | NaN |
| 42521 | NaN | NaN | NaN | NaN |
| 42522 | NaN | NaN | NaN | NaN |
| 42523 | NaN | NaN | NaN | NaN |
| 42524 | NaN | NaN | NaN | NaN |
| 42525 | NaN | NaN | NaN | NaN |
| 42526 | NaN | NaN | NaN | NaN |
| 42527 | NaN | NaN | NaN | NaN |
| 42528 | NaN | NaN | NaN | NaN |
| 42529 | NaN | NaN | NaN | NaN |
| 42530 | NaN | NaN | NaN | NaN |
| 42531 | NaN | NaN | NaN | NaN |
| 42532 | NaN | NaN | NaN | NaN |
| 42533 | NaN | NaN | NaN | NaN |
| 42534 | NaN | NaN | NaN | NaN |
| 42535 | NaN | NaN | NaN | NaN |
| 42536 | NaN | NaN | NaN | NaN |
| 42537 | NaN | NaN | NaN | NaN |
| | | | | |

[42538 rows x 58 columns]

What if I do not want to drop any rows with missing data instead fill it with more meaning value?

1.1.3 Filling Missing data with Value There are different ways we can fill the missing data with some meaningful values like mean or median or most frequently value available in the column. One of the method is to use fillna() by specifying how we want to fill the null or missing values.

| 4 | 1075358 | 1311748.0 |
|-------|---------|-----------|
| 5 | 1075269 | 1311441.0 |
| 6 | 1069639 | 1304742.0 |
| 7 | 1072053 | 1288686.0 |
| 8 | 1071795 | 1306957.0 |
| 9 | 1071570 | 1306721.0 |
| 10 | 1070078 | 1305201.0 |
| 11 | 1069908 | 1305008.0 |
| 12 | 1064687 | 1298717.0 |
| 13 | 1069866 | 1304956.0 |
| 14 | 1069057 | 1303503.0 |
| 15 | 1069759 | 1304871.0 |
| 16 | 1065775 | 1299699.0 |
| 17 | 1069971 | 1304884.0 |
| 18 | 1062474 | 1294539.0 |
| 19 | 1069742 | 1304855.0 |
| 20 | 1069740 | 1284848.0 |
| 21 | 1039153 | 1269083.0 |
| 22 | 1069710 | 1304821.0 |
| 23 | 1069700 | 1304810.0 |
| 24 | 1069559 | 1304634.0 |
| 25 | 1069697 | 1273773.0 |
| 26 | 1069800 | 1304679.0 |
| 27 | 1069657 | 1304764.0 |
| 28 | 1069799 | 1304678.0 |
| 29 | 1047704 | 1278806.0 |
| | | 1270000.0 |
| 42508 | 91175 | 91170.0 |
| 42509 | 91126 | 91067.0 |
| 42510 | 91120 | 70879.0 |
| 42510 | 90106 | 90090.0 |
| 42511 | | 80039.0 |
| | 89258 | |
| 42513 | 88637 | 88629.0 |
| 42514 | 88046 | 88023.0 |
| 42515 | 85961 | 85923.0 |
| 42516 | 85818 | 85802.0 |
| 42517 | 85781 | 85727.0 |
| 42518 | 85675 | 85667.0 |
| 42519 | 84670 | 79576.0 |
| 42520 | 84098 | 84091.0 |
| 42521 | 83979 | 83974.0 |
| 42522 | 83489 | 83471.0 |
| 42523 | 83185 | 83132.0 |
| 42524 | 76629 | 76623.0 |
| 42525 | 74014 | 73890.0 |
| 42526 | 81085 | 80973.0 |
| 42527 | 77792 | 77764.0 |
| 42528 | 77757 | 70626.0 |
| | | |

```
74505
42529
                                                               74469.0
42530
                                                     74323
                                                               74301.0
42531
                                                     73582
                                                               73096.0
42532
                                                     72998
                                                               72992.0
42533
                                                     72176
                                                               70868.0
42534
                                                     71623
                                                               70735.0
42535
                                                     70686
                                                               70681.0
42536
       Total amount funded in policy code 1: 460296150
                                                                   NaN
                Total amount funded in policy code 2: 0
42537
                                                                   NaN
           loan_amnt
                       funded_amnt
                                     funded_amnt_inv
                                                              term int_rate
                                                                              \
0
        5000.000000
                            5000.0
                                         4975.000000
                                                         36 months
                                                                      10.65%
1
        2500.000000
                                                         60 months
                                                                      15.27%
                            2500.0
                                         2500.000000
2
        2400.000000
                            2400.0
                                         2400.000000
                                                         36 months
                                                                      15.96%
3
       10000.000000
                           10000.0
                                        10000.000000
                                                           months
                                                                      13.49%
4
                                                         60 months
                                                                      12.69%
        3000.000000
                            3000.0
                                         3000.000000
5
        5000.000000
                            5000.0
                                         5000.000000
                                                         36 months
                                                                       7.90%
6
                                                                      15.96%
        7000.000000
                            7000.0
                                         7000.000000
                                                         60 months
7
        3000.000000
                                                         36 months
                                                                      18.64%
                            3000.0
                                         3000.000000
8
        5600.000000
                                         5600.000000
                                                         60 months
                                                                      21.28%
                            5600.0
9
        5375.000000
                            5375.0
                                         5350.000000
                                                         60 months
                                                                      12.69%
10
        6500.000000
                            6500.0
                                         6500.000000
                                                         60 months
                                                                      14.65%
11
       12000.000000
                           12000.0
                                        12000.000000
                                                         36 months
                                                                      12.69%
                                                                      13.49%
12
        9000.000000
                            9000.0
                                         9000.000000
                                                         36 months
13
                                         3000.000000
                                                         36 months
                                                                       9.91%
        3000.000000
                            3000.0
14
       10000.000000
                                        10000.000000
                                                         36 months
                                                                      10.65%
                           10000.0
15
        1000.000000
                            1000.0
                                         1000.000000
                                                         36 months
                                                                      16.29%
16
       10000.000000
                           10000.0
                                        10000.000000
                                                           months
                                                                      15.27%
17
        3600.000000
                            3600.0
                                         3600.000000
                                                           months
                                                                       6.03%
18
        6000.000000
                            6000.0
                                         6000.000000
                                                           months
                                                                      11.71%
                                                           months
19
        9200.000000
                            9200.0
                                         9200.000000
                                                         36
                                                                       6.03%
20
       20250.000000
                           20250.0
                                        19142.161077
                                                         60 months
                                                                      15.27%
21
       21000.000000
                           21000.0
                                        21000.000000
                                                         36 months
                                                                      12.42%
22
       10000.000000
                           10000.0
                                        10000.000000
                                                         36 months
                                                                      11.71%
23
       10000.000000
                                        10000.000000
                                                         36 months
                                                                      11.71%
                           10000.0
24
        6000.000000
                            6000.0
                                         6000.000000
                                                         36 months
                                                                      11.71%
25
       15000.000000
                           15000.0
                                        15000.000000
                                                         36 months
                                                                       9.91%
26
       15000.000000
                           15000.0
                                                         36 months
                                                                      14.27%
                                         8725.000000
27
        5000.000000
                            5000.0
                                         5000.000000
                                                         60 months
                                                                      16.77%
28
        4000.000000
                            4000.0
                                         4000.000000
                                                         36 months
                                                                      11.71%
29
        8500.000000
                            8500.0
                                         8500.000000
                                                         36 months
                                                                      11.71%
. . .
                               . . .
                                                                         . . .
                                                                      13.12%
42508
        6000.000000
                            6000.0
                                         1200.000000
                                                         36 months
42509
        5350.000000
                            5350.0
                                          625.000000
                                                         36 months
                                                                      13.12%
42510
        1900.000000
                            1900.0
                                          900.000000
                                                           months
                                                                       9.64%
42511
       10000.000000
                           10000.0
                                          350.000000
                                                         36 months
                                                                      14.70%
42512
        2000.000000
                            2000.0
                                         1275.000000
                                                         36 months
                                                                       7.12%
42513
        6000.000000
                            6000.0
                                          650.000000
                                                         36 months
                                                                      10.59%
```

| 42514 | 4400.000000 | | 4400.0 | | 000000 | | months | | 64% |
|---|---|--|--|---|--------|-------------------|--|--|-----|
| 42515 | 1200.000000 | | 1200.0 | | 000000 | | months | | 01% |
| 42516 | 5000.000000 | | 5000.0 | | 000000 | | months | | 22% |
| 42517 | 1400.000000 | | 1400.0 | | 000000 | | months | | 91% |
| 42518 | 1000.000000 | | 1000.0 | | 000000 | | months | | 07% |
| 42519 | 5000.000000 | | 5000.0 | | 000000 | | months | | 75% |
| 42520 | 2500.000000 | | 2500.0 | | 000000 | | months | | 43% |
| 42521 | 3000.000000 | | 3000.0 | | 000000 | | months | | 43% |
| 42522 | 2600.000000 | | 2600.0 | | 000000 | | months | | 38% |
| 42523 | 1000.000000 | | 1000.0 | 625. | 000000 | 36 | months | 7. | 12% |
| 42524 | 1275.000000 | | 1275.0 | 0. | 000000 | 36 | months | 12. | 49% |
| 42525 | 6450.000000 | | 6450.0 | 0. | 000000 | 36 | months | 11. | 22% |
| 42526 | 10500.000000 | 1 | .0500.0 | 275. | 000000 | 36 | months | 11. | 22% |
| 42527 | 3000.000000 | | 3000.0 | 125. | 000000 | 36 | months | 9. | 01% |
| 42528 | 3000.000000 | | 3000.0 | 0. | 000000 | 36 | months | 9. | 33% |
| 42529 | 2000.000000 | | 2000.0 | 225. | 000000 | 36 | months | 9. | 96% |
| 42530 | 6500.000000 | | 6500.0 | 0. | 000000 | 36 | months | 9. | 64% |
| 42531 | 3500.000000 | | 3500.0 | 225. | 000000 | 36 | months | 10. | 28% |
| 42532 | 1000.000000 | | 1000.0 | 0. | 000000 | 36 | months | 9. | 64% |
| 42533 | 2525.000000 | | 2525.0 | 225. | 000000 | 36 | months | 9. | 33% |
| 42534 | 6500.000000 | | 6500.0 | 0. | 000000 | 36 | months | 8. | 38% |
| 42535 | 5000.000000 | | 5000.0 | 0. | 000000 | 36 | months | 7. | 75% |
| 42536 | 11089.722581 | | NaN | | NaN | | NaN | | NaN |
| 42537 | 11089.722581 | | NaN | | NaN | | NaN | | NaN |
| | | | | | | | | | |
| | installment gr | | ush manada | | lagt | | | l- | \ |
| | THE GATTMONG PT | ade s | sub_grade | | Table_ | fico | _range_h | ngn | \ |
| 0 | 162.87 | ade s B | B2 | | Tabe_ | fico _. | _ | 11gn 14.0 | \ |
| 0 1 | | | _ | | 1450_ | fico _. | 74 | - | ` |
| | 162.87 | В | B2 | | 1450_ | fico _. | 74 49 | 14.0 | \ |
| 1 | 162.87 59.83 | B C | B2 C4 | | 1450_ | fico _. | 74 49 71 | 14.0 99.0 | \ |
| 1 2 | 162.87 59.83 84.33 | B C C | B2 C4 C5 | • | 1450 | fico _. | 74 49 71 60 | 14.0 99.0 19.0 | ` |
| 1 2 3 | 162.87 59.83 84.33 339.31 | B C C | B2 C4 C5 C1 | | 1450_ | fico _. | 74 49 71 60 69 | 14.0 99.0 19.0 04.0 | ` |
| 1 2 3 4 | 162.87 59.83 84.33 339.31 67.79 | B C C C | B2 C4 C5 C1 B5 | | 1450_ | fico _. | 74 49 71 60 69 | 14.0 99.0 19.0 04.0 | |
| 1 2 3 4 5 | 162.87 59.83 84.33 339.31 67.79 156.46 | B C C B A | B2 C4 C5 C1 B5 A4 | | 1450_ | fico _. | 74 49 71 60 69 67 | 14.0 99.0 19.0 04.0 94.0 | |
| 1 2 3 4 5 | 162.87 59.83 84.33 339.31 67.79 156.46 170.08 | B C C B A | B2 C4 C5 C1 B5 A4 C5 | | 1450_ | fico _. | 74 49 71 60 69 67 65 | 14.0 99.0 19.0 04.0 94.0 79.0 | |
| 1 2 3 4 5 6 7 | 162.87 59.83 84.33 339.31 67.79 156.46 170.08 109.43 | B C C B A C | B2 C4 C5 C1 B5 A4 C5 E1 | | 1450_ | fico _. | 74 49 71 60 69 67 68 49 | 14.0 99.0 19.0 04.0 94.0 79.0 54.0 | |
| 1 2 3 4 5 6 7 8 | 162.87 59.83 84.33 339.31 67.79 156.46 170.08 109.43 152.39 | B C C B A C E | B2 C4 C5 C1 B5 A4 C5 E1 | | 1450_ | fico _. | 74 49 71 60 69 67 65 68 49 | 14.0 99.0 19.0 04.0 79.0 54.0 39.0 | |
| 1 2 3 4 5 6 7 8 | 162.87 59.83 84.33 339.31 67.79 156.46 170.08 109.43 152.39 121.45 | B C C C B A C E F | B2 C4 C5 C1 B5 A4 C5 E1 F2 B5 | | 1450_ | fico | 74 49 71 60 69 67 65 68 49 51 | 14.0 99.0 19.0 94.0 94.0 79.0 54.0 99.0 | |
| 1 2 3 4 5 6 7 8 9 10 | 162.87 59.83 84.33 339.31 67.79 156.46 170.08 109.43 152.39 121.45 153.45 | B C C B A C E F B | B2 C4 C5 C1 B5 A4 C5 E1 F2 B5 C3 | | 1450_ | fico | 74 49 71 60 69 67 68 49 51 73 | 14.0 99.0 19.0 94.0 94.0 79.0 54.0 39.0 99.0 | |
| 1 2 3 4 5 6 7 8 9 10 | 162.87 59.83 84.33 339.31 67.79 156.46 170.08 109.43 152.39 121.45 153.45 402.54 | B C C B A C E F B | B2 C4 C5 C1 B5 A4 C5 E1 F2 B5 C3 B5 | | 1450_ | fico | 74 49 71 60 69 67 65 68 49 51 73 66 | 14.0 99.0 19.0 94.0 94.0 79.0 54.0 39.0 19.0 34.0 | |
| 1 2 3 4 5 6 7 8 9 10 11 12 | 162.87 59.83 84.33 339.31 67.79 156.46 170.08 109.43 152.39 121.45 153.45 402.54 305.38 | B C C B A C E F B C | B2 C4 C5 C1 B5 A4 C5 E1 F2 B5 C3 B5 C1 | | 1450_ | fico | 74 49 71 60 69 67 65 68 49 51 73 66 61 73 | 14.0 99.0 19.0 94.0 94.0 79.0 54.0 89.0 19.0 34.0 | |
| 1 2 3 4 5 6 7 8 9 10 11 12 13 | 162.87 59.83 84.33 339.31 67.79 156.46 170.08 109.43 152.39 121.45 153.45 402.54 305.38 96.68 | B C C B A C E B C B | B2 C4 C5 C1 B5 A4 C5 E1 F2 B5 C3 B5 C1 B1 | | | fico | 74 49 71 60 69 67 68 49 51 73 66 61 73 | 14.0 99.0 19.0 94.0 94.0 79.0 54.0 89.0 19.0 84.0 89.0 | |
| 1 2 3 4 5 6 7 8 9 10 11 12 13 14 | 162.87 59.83 84.33 339.31 67.79 156.46 170.08 109.43 152.39 121.45 153.45 402.54 305.38 96.68 325.74 | B C C B A C E F B C B B | B2 C4 C5 C1 B5 A4 C5 E1 F2 B5 C3 B5 C1 B1 B2 | | | fico | 74 49 71 60 69 67 68 49 51 73 66 61 73 | 14.0 99.0 19.0 94.0 94.0 79.0 54.0 39.0 19.0 34.0 39.0 19.0 34.0 34.0 34.0 | |
| 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 | 162.87 59.83 84.33 339.31 67.79 156.46 170.08 109.43 152.39 121.45 153.45 402.54 305.38 96.68 325.74 35.31 | B C C B A C E F B C B B D | B2 C4 C5 C1 B5 A4 C5 E1 F2 B5 C3 B5 C1 B1 B2 D1 | | | fico | 74 49 71 60 69 67 65 68 49 51 73 66 61 73 65 | 14.0 99.0 19.0 94.0 94.0 79.0 54.0 39.0 99.0 19.0 34.0 39.0 39.0 39.0 39.0 39.0 | |
| 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 | 162.87 59.83 84.33 339.31 67.79 156.46 170.08 109.43 152.39 121.45 153.45 402.54 305.38 96.68 325.74 35.31 347.98 | B C C B A C B B C C B | B2 C4 C5 C1 B5 A4 C5 E1 F2 B5 C3 B5 C1 B1 B2 D1 | | | fico | 74 49 71 60 69 67 68 49 51 73 66 61 73 65 79 | 14.0 99.0 19.0 94.0 94.0 79.0 54.0 39.0 19.0 34.0 59.0 19.0 34.0 59.0 39.0 39.0 | |
| 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 | 162.87 59.83 84.33 339.31 67.79 156.46 170.08 109.43 152.39 121.45 153.45 402.54 305.38 96.68 325.74 35.31 347.98 109.57 | B C C B A C B B C A | B2 C4 C5 C1 B5 A4 C5 E1 F2 B5 C3 B5 C1 B1 B2 D1 C4 A1 | | | fico | 74 49 71 60 69 67 68 49 51 73 66 61 73 65 76 | 14.0 99.0 19.0 94.0 94.0 79.0 54.0 39.0 19.0 34.0 39.0 19.0 34.0 39.0 34.0 39.0 34.0 39.0 | |
| 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 | 162.87 59.83 84.33 339.31 67.79 156.46 170.08 109.43 152.39 121.45 153.45 402.54 305.38 96.68 325.74 35.31 347.98 109.57 198.46 | B C C B A C E F B C B B C A B | B2 C4 C5 C1 B5 A4 C5 E1 F2 B5 C3 B5 C1 B1 B2 D1 C4 A1 B3 | | | fico | 74 49 71 60 69 67 65 68 49 51 73 66 61 73 65 76 63 79 55 | 14.0 99.0 19.0 94.0 94.0 99.0 54.0 39.0 99.0 19.0 34.0 34.0 39.0 34.0 39.0 34.0 39.0 | |
| 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 | 162.87 59.83 84.33 339.31 67.79 156.46 170.08 109.43 152.39 121.45 153.45 402.54 305.38 96.68 325.74 35.31 347.98 109.57 198.46 280.01 | B C C B A C E F B C B A A B A | B2 C4 C5 C1 B5 A4 C5 E1 F2 B5 C3 B5 C1 B1 B2 D1 C4 A1 B3 A1 | | | fico | 74 49 71 60 69 67 68 49 51 73 66 61 73 65 76 63 79 55 64 | 14.0 99.0 19.0 94.0 94.0 79.0 54.0 39.0 99.0 19.0 34.0 59.0 30.0 30.0 30.0 30.0 30.0 30.0 30.0 30.0 30.0 30.0 30.0 30.0 30.0 | |

| 22 | 330.76 | В | В3 | | | 779.0 | |
|----------|---------------|-------|------------|----------|-------|-------|---|
| 23 | 330.76 | В | В3 | | | 684.0 | |
| 24 | 198.46 | В | В3 | | | 499.0 | |
| 25 | 483.38 | В | B1 | | | 734.0 | |
| 26 | 514.64 | C | C2 | • • • | | 509.0 | |
| 27 | 123.65 | D | D2 | • • • | | 579.0 | |
| 28 | 132.31 | В | B3 | • • • | | 689.0 | |
| 20 29 | 281.15 | | | • • • | | 599.0 | |
| | | В | В3 | • • • | | | |
| 42508 | 202.51 | D | D5 | • • • | | 664.0 | |
| 42509 | | | | • • • | | 704.0 | |
| | | D | D5 | • • • | | | |
| 42510 | | В | B4 | • • • | | 709.0 | |
| 42511 | | E | E5 | • • • | | 499.0 | |
| 42512 | | A | A1 | • • • | | 809.0 | |
| 42513 | | C | C2 | • • • | | 769.0 | |
| 42514 | | В | B4 | • • • | | 584.0 | |
| 42515 | | В | B2 | • • • | | 694.0 | |
| 42516 | 164.23 | C | C4 | | | 679.0 | |
| 42517 | 45.78 | C | C3 | | | 679.0 | |
| 42518 | 34.21 | E | E3 | | | 499.0 | |
| 42519 | 156.11 | Α | A3 | | | 759.0 | |
| 42520 | 77.69 | Α | A2 | | | 789.0 | |
| 42521 | 93.23 | Α | A2 | | | 799.0 | |
| 42522 | 81.94 | Α | A 5 | | | 614.0 | |
| 42523 | 30.94 | Α | A1 | | | 599.0 | |
| 42524 | | D | D3 | | | 624.0 | |
| 42525 | | С | C4 | | | 574.0 | |
| 42526 | | C | C4 | | | 744.0 | |
| 42527 | | В | B2 | | | 594.0 | |
| 42528 | | В | В3 | | | 714.0 | |
| 42529 | | В | B5 | ••• | | 594.0 | |
| 42530 | | В | B4 | • • • | | 684.0 | |
| 42531 | 113.39 | C | C1 | ••• | | 819.0 | |
| 42532 | | В | B4 | • • • | | 784.0 | |
| | | | | • • • | | | |
| 42533 | | В | B3 | • • • | | 714.0 | |
| 42534 | | A | A5 | • • • | | 724.0 | |
| 42535 | | A | A3 | • • • | | 794.0 | |
| 42536 | | NaN | NaN | • • • | | NaN | |
| 42537 | NaN | NaN | NaN | • • • | | NaN | |
| | | _ | | | _ | | |
| | last_fico_ran | _ | collection | s_12_mth | | - • | \ |
| 0 | | 740.0 | | | False | True | |
| 1 | | 0.0 | | | False | True | |
| 2 | | 715.0 | | | False | True | |
| 3 | | 600.0 | | | False | True | |
| 4 | | 690.0 | | | False | True | |
| 5 | | 675.0 | | | False | True | |
| 6 | | 650.0 | | | False | True | |
| | | | | | | | |

| 8 0.0 False 9 515.0 False 10 730.0 False | True True True True True |
|--|--------------------------|
| 9 515.0 False 10 730.0 False | True True True |
| 10 730.0 False | True True |
| | True |
| 11 665.0 False | |
| | True |
| | True |
| | True |
| 15 765.0 False | True |
| 16 635.0 False | True |
| 17 790.0 False | True |
| 18 555.0 False | True |
| 19 640.0 False | True |
| 20 680.0 False | True |
| 21 530.0 False | True |
| 22 775.0 False | True |
| 23 680.0 False | True |
| 24 0.0 False | True |
| 25 730.0 False | True |
| 26 505.0 False | True |
| 27 575.0 False | True |
| 28 685.0 False | True |
| 29 595.0 False | True |
| | |
| 42508 660.0 NaN | True |
| 42509 700.0 NaN | True |
| 42510 705.0 NaN | True |
| 42511 0.0 NaN | True |
| 42512 805.0 NaN | True |
| 42513 765.0 NaN | True |
| 42514 580.0 NaN | True |
| 42515 690.0 NaN | True |
| 42516 675.0 NaN | True |
| | True |
| 42531 815.0 NaN | True |

| 42532 | 780 | 0.0 | NaN | Tr | rue | |
|-------|------------------|----------------|-----------------|-----------|-------------|---|
| 42533 | 710 | 0.0 | NaN | Tr | rue | |
| 42534 | 720 | 0.0 | NaN | Tr | rue | |
| 42535 | 790 | 0.0 | NaN | Tr | ue | |
| 42536 | ľ | NaN | NaN | N | aN | |
| 42537 | ľ | NaN | NaN | N | aN | |
| | | | | | | |
| | application_type | acc_now_delinq | chargeoff_withi | n_12_mths | delinq_amnt | \ |
| 0 | INDIVIDUAL | False | | False | 0.0 | |
| 1 | INDIVIDUAL | False | | False | 0.0 | |
| 2 | INDIVIDUAL | False | | False | 0.0 | |
| 3 | INDIVIDUAL | False | | False | 0.0 | |
| 4 | INDIVIDUAL | False | | False | 0.0 | |
| 5 | INDIVIDUAL | False | | False | 0.0 | |
| 6 | INDIVIDUAL | False | | False | 0.0 | |
| 7 | INDIVIDUAL | False | | False | 0.0 | |
| 8 | INDIVIDUAL | False | | False | 0.0 | |
| 9 | INDIVIDUAL | False | | False | 0.0 | |
| 10 | INDIVIDUAL | False | | False | 0.0 | |
| 11 | INDIVIDUAL | False | | False | 0.0 | |
| 12 | INDIVIDUAL | False | | False | 0.0 | |
| 13 | INDIVIDUAL | False | | False | 0.0 | |
| 14 | INDIVIDUAL | False | | False | 0.0 | |
| 15 | INDIVIDUAL | False | | False | 0.0 | |
| 16 | INDIVIDUAL | False | | False | 0.0 | |
| 17 | INDIVIDUAL | False | | False | 0.0 | |
| 18 | INDIVIDUAL | False | | False | 0.0 | |
| 19 | INDIVIDUAL | False | | False | 0.0 | |
| 20 | INDIVIDUAL | False | | False | 0.0 | |
| 21 | INDIVIDUAL | False | | False | 0.0 | |
| 22 | INDIVIDUAL | False | | False | 0.0 | |
| 23 | INDIVIDUAL | False | | False | 0.0 | |
| 24 | INDIVIDUAL | False | | False | 0.0 | |
| 25 | INDIVIDUAL | False | | False | 0.0 | |
| 26 | INDIVIDUAL | False | | False | 0.0 | |
| 27 | INDIVIDUAL | False | | False | 0.0 | |
| 28 | INDIVIDUAL | False | | False | 0.0 | |
| 29 | INDIVIDUAL | False | | False | 0.0 | |
| • • • | • • • | • • • | | | • • • | |
| 42508 | | False | | NaN | 0.0 | |
| 42509 | | False | | NaN | 0.0 | |
| 42510 | | NaN | | NaN | NaN | |
| 42511 | | False | | NaN | 0.0 | |
| 42512 | | False | | NaN | 0.0 | |
| 42513 | | False | | NaN | 0.0 | |
| 42514 | | False | | NaN | 0.0 | |
| 42515 | | NaN | | NaN | NaN | |
| 42516 | INDIVIDUAL | NaN | | NaN | NaN | |

| 42517 | INDIVIDUAL | NaN | NaN | NaN |
|-------|------------|-----|-----|-----|
| 42518 | INDIVIDUAL | NaN | NaN | NaN |
| 42519 | INDIVIDUAL | NaN | NaN | NaN |
| 42520 | INDIVIDUAL | NaN | NaN | NaN |
| 42521 | INDIVIDUAL | NaN | NaN | NaN |
| 42522 | INDIVIDUAL | NaN | NaN | NaN |
| 42523 | INDIVIDUAL | NaN | NaN | NaN |
| 42524 | INDIVIDUAL | NaN | NaN | NaN |
| 42525 | INDIVIDUAL | NaN | NaN | NaN |
| 42526 | INDIVIDUAL | NaN | NaN | NaN |
| 42527 | INDIVIDUAL | NaN | NaN | NaN |
| 42528 | INDIVIDUAL | NaN | NaN | NaN |
| 42529 | INDIVIDUAL | NaN | NaN | NaN |
| 42530 | INDIVIDUAL | NaN | NaN | NaN |
| 42531 | INDIVIDUAL | NaN | NaN | NaN |
| 42532 | INDIVIDUAL | NaN | NaN | NaN |
| 42533 | INDIVIDUAL | NaN | NaN | NaN |
| 42534 | INDIVIDUAL | NaN | NaN | NaN |
| 42535 | INDIVIDUAL | NaN | NaN | NaN |
| 42536 | NaN | NaN | NaN | NaN |
| 42537 | NaN | NaN | NaN | NaN |
| | | | | |

| | <pre>pub_rec_bankruptcies</pre> | tax_liens |
|----|---------------------------------|-----------|
| 0 | 0.0 | False |
| 1 | 0.0 | False |
| 2 | 0.0 | False |
| 3 | 0.0 | False |
| 4 | 0.0 | False |
| 5 | 0.0 | False |
| 6 | 0.0 | False |
| 7 | 0.0 | False |
| 8 | 0.0 | False |
| 9 | 0.0 | False |
| 10 | 0.0 | False |
| 11 | 0.0 | False |
| 12 | 0.0 | False |
| 13 | 0.0 | False |
| 14 | 0.0 | False |
| 15 | 0.0 | False |
| 16 | 0.0 | False |
| 17 | 0.0 | False |
| 18 | 0.0 | False |
| 19 | 0.0 | False |
| 20 | 0.0 | False |
| 21 | 0.0 | False |
| 22 | 0.0 | False |
| 23 | 0.0 | False |
| 24 | 0.0 | False |

```
0.0
          25
                                                   False
          26
                                        0.0
                                                   False
          27
                                        0.0
                                                   False
          28
                                        0.0
                                                   False
          29
                                        0.0
                                                   False
          . . .
                                        . . .
          42508
                                        NaN
                                                     NaN
          42509
                                        {\tt NaN}
                                                     NaN
          42510
                                        {\tt NaN}
                                                     NaN
          42511
                                        {\tt NaN}
                                                     {\tt NaN}
          42512
                                        {\tt NaN}
                                                     NaN
          42513
                                                     NaN
                                        NaN
          42514
                                                     NaN
                                        {\tt NaN}
          42515
                                                     NaN
                                        NaN
          42516
                                                     NaN
                                        {\tt NaN}
          42517
                                        {\tt NaN}
                                                     NaN
          42518
                                        {\tt NaN}
                                                     {\tt NaN}
          42519
                                                     NaN
                                        {\tt NaN}
          42520
                                        {\tt NaN}
                                                     NaN
          42521
                                        NaN
                                                     NaN
          42522
                                        {\tt NaN}
                                                     NaN
          42523
                                        NaN
                                                     NaN
          42524
                                        {\tt NaN}
                                                     NaN
          42525
                                        {\tt NaN}
                                                     NaN
          42526
                                        {\tt NaN}
                                                     NaN
          42527
                                                     NaN
                                        NaN
          42528
                                                     NaN
                                        NaN
                                                     NaN
          42529
                                        {\tt NaN}
          42530
                                                     NaN
                                        NaN
          42531
                                        {\tt NaN}
                                                     NaN
          42532
                                        {\tt NaN}
                                                     NaN
          42533
                                        {\tt NaN}
                                                     {\tt NaN}
          42534
                                                     NaN
                                        NaN
          42535
                                        {\tt NaN}
                                                     NaN
          42536
                                                     NaN
                                        NaN
          42537
                                        NaN
                                                     {\tt NaN}
          [42538 rows x 58 columns]
In [10]: #fill with median
           df['funded_amnt'].fillna(df['funded_amnt'].median(), inplace = True)
Out[10]:
                                                                                id member_id \
                                                                                     1296599.0
           0
                                                                         1077501
           1
                                                                         1077430
                                                                                     1314167.0
           2
                                                                         1077175
                                                                                     1313524.0
           3
                                                                         1076863
                                                                                     1277178.0
```

| 4 | 1075358 | 1311748.0 |
|---|---|--|
| 5 | 1075269 | 1311441.0 |
| 6 | 1069639 | 1304742.0 |
| 7 | 1072053 | 1288686.0 |
| 8 | 1071795 | 1306957.0 |
| 9 | 1071570 | 1306937.0 |
| 10 | | 1305721.0 |
| 11 | 1070078 | |
| | 1069908 | 1305008.0 |
| 12 | 1064687 | 1298717.0 |
| 13 | 1069866 | 1304956.0 |
| 14 | 1069057 | 1303503.0 |
| 15 | 1069759 | 1304871.0 |
| 16 | 1065775 | 1299699.0 |
| 17 | 1069971 | 1304884.0 |
| 18 | 1062474 | 1294539.0 |
| 19 | 1069742 | 1304855.0 |
| 20 | 1069740 | 1284848.0 |
| 21 | 1039153 | 1269083.0 |
| 22 | 1069710 | 1304821.0 |
| 23 | 1069700 | 1304810.0 |
| 24 | 1069559 | 1304634.0 |
| 25 | 1069697 | 1273773.0 |
| 26 | 1069800 | 1304679.0 |
| 27 | 1069657 | 1304764.0 |
| 28 | 1069799 | 1304678.0 |
| 30 | 1017701 | 1070000 0 |
| 29 | 1047704 | 1278806.0 |
| | 1047704 | 12/8806.0 |
| | | 91170.0 |
| | | |
| 42508 | 91175 | 91170.0 |
| 42508 42509 | 91175 91126 | 91170.0 91067.0 |
| 42508 42509 42510 | 91175 91126 91023 | 91170.0 91067.0 70879.0 |
| 42508 42509 42510 42511 | 91175 91126 91023 90106 | 91170.0 91067.0 70879.0 90090.0 |
| 42508 42509 42510 42511 42512 | 91175 91126 91023 90106 89258 | 91170.0 91067.0 70879.0 90090.0 80039.0 |
| 42508 42509 42510 42511 42512 42513 | 91175 91126 91023 90106 89258 88637 | 91170.0 91067.0 70879.0 90090.0 80039.0 88629.0 |
| 42508 42509 42510 42511 42512 42513 42514 | 91175 91126 91023 90106 89258 88637 88046 | 91170.0 91067.0 70879.0 90090.0 80039.0 88629.0 88023.0 |
| 42508 42509 42510 42511 42512 42513 42514 42515 | 91175 91126 91023 90106 89258 88637 88046 85961 | 91170.0 91067.0 70879.0 90090.0 80039.0 88629.0 88023.0 85923.0 |
| 42508 42509 42510 42511 42512 42513 42514 42515 42516 | 91175 91126 91023 90106 89258 88637 88046 85961 | 91170.0 91067.0 70879.0 90090.0 80039.0 88629.0 88023.0 85923.0 |
| 42508 42509 42510 42511 42512 42513 42514 42515 42516 42517 | 91175 91126 91023 90106 89258 88637 88046 85961 85818 | 91170.0 91067.0 70879.0 90090.0 80039.0 88629.0 88023.0 85923.0 85802.0 |
| 42508 42509 42510 42511 42512 42513 42514 42515 42516 42517 42518 | 91175 91126 91023 90106 89258 88637 88046 85961 85818 85781 | 91170.0 91067.0 70879.0 90090.0 80039.0 88629.0 85923.0 85923.0 85727.0 |
| 42508 42509 42510 42511 42512 42513 42514 42515 42516 42517 42518 42519 | 91175 91126 91023 90106 89258 88637 88046 85961 85818 85781 85675 84670 | 91170.0 91067.0 70879.0 90090.0 80039.0 88629.0 85923.0 85923.0 85802.0 85727.0 85667.0 79576.0 |
| 42508 42509 42510 42511 42512 42513 42514 42515 42516 42517 42518 42519 42520 | 91175 91126 91023 90106 89258 88637 88046 85961 85818 85781 85675 84670 84098 | 91170.0 91067.0 70879.0 90090.0 80039.0 88629.0 85923.0 85923.0 85727.0 85667.0 79576.0 84091.0 |
| 42508 42509 42510 42511 42512 42513 42514 42515 42516 42517 42518 42519 42520 42521 | 91175 91126 91023 90106 89258 88637 88046 85961 85818 85781 85675 84670 84098 83979 | 91170.0 91067.0 70879.0 90090.0 80039.0 88629.0 85923.0 85923.0 85727.0 85667.0 79576.0 84091.0 |
| 42508 42509 42510 42511 42512 42513 42514 42515 42516 42517 42518 42519 42520 42521 | 91175 91126 91023 90106 89258 88637 88046 85961 85818 85781 85675 84670 84098 83979 83489 | 91170.0 91067.0 70879.0 90090.0 80039.0 88629.0 85923.0 85923.0 85727.0 85667.0 79576.0 84091.0 83974.0 |
| 42508 42509 42510 42511 42512 42513 42514 42515 42516 42517 42518 42519 42520 42521 42523 | 91175 91126 91023 90106 89258 88637 88046 85961 85818 85781 85675 84670 84098 83979 83489 83185 76629 | 91170.0 91067.0 70879.0 90090.0 80039.0 88629.0 85923.0 85923.0 85727.0 85667.0 79576.0 84091.0 83974.0 83471.0 83132.0 76623.0 |
| 42508 42509 42510 42511 42512 42513 42514 42515 42516 42517 42518 42519 42520 42521 42522 42523 42524 42525 | 91175 91126 91023 90106 89258 88637 88046 85961 85818 85781 85675 84670 84098 83979 83489 83185 76629 74014 | 91170.0 91067.0 70879.0 90090.0 80039.0 88629.0 85923.0 85923.0 85727.0 85667.0 79576.0 84091.0 83974.0 83471.0 |
| 42508 42509 42510 42511 42512 42513 42514 42515 42516 42517 42518 42519 42520 42521 42522 42523 42524 42525 42526 | 91175 91126 91023 90106 89258 88637 88046 85961 85818 85781 85675 84670 84098 83979 83489 83185 76629 74014 81085 | 91170.0 91067.0 70879.0 90090.0 80039.0 88629.0 85923.0 85923.0 85727.0 85667.0 79576.0 84091.0 83471.0 83471.0 83132.0 76623.0 73890.0 80973.0 |
| 42508 42509 42510 42511 42512 42513 42514 42515 42516 42517 42518 42519 42520 42521 42522 42523 42524 42525 | 91175 91126 91023 90106 89258 88637 88046 85961 85818 85781 85675 84670 84098 83979 83489 83185 76629 74014 | 91170.0 91067.0 70879.0 90090.0 80039.0 88629.0 85923.0 85923.0 85727.0 85667.0 79576.0 84091.0 83974.0 83132.0 76623.0 73890.0 |

```
42529
                                                     74505
                                                               74469.0
42530
                                                     74323
                                                               74301.0
42531
                                                     73582
                                                               73096.0
42532
                                                     72998
                                                               72992.0
42533
                                                     72176
                                                               70868.0
42534
                                                     71623
                                                               70735.0
42535
                                                     70686
                                                               70681.0
42536
       Total amount funded in policy code 1: 460296150
                                                                   NaN
                Total amount funded in policy code 2: 0
42537
                                                                   NaN
                       funded_amnt
           loan_amnt
                                     funded_amnt_inv
                                                              term int_rate
                                                                              \
0
         5000.000000
                            5000.0
                                         4975.000000
                                                         36 months
                                                                      10.65%
1
                                                         60 months
                                                                      15.27%
         2500.000000
                            2500.0
                                         2500.000000
2
         2400.000000
                            2400.0
                                         2400.000000
                                                         36 months
                                                                      15.96%
3
       10000.000000
                           10000.0
                                        10000.000000
                                                         36 months
                                                                      13.49%
4
                                                         60 months
                                                                      12.69%
         3000.000000
                            3000.0
                                         3000.000000
5
         5000.000000
                            5000.0
                                         5000.000000
                                                         36 months
                                                                       7.90%
                                                                      15.96%
6
         7000.000000
                            7000.0
                                         7000.000000
                                                         60 months
7
                                                         36 months
         3000.000000
                            3000.0
                                         3000.000000
                                                                      18.64%
8
                            5600.0
                                                         60 months
                                                                      21.28%
         5600.000000
                                         5600.000000
9
         5375.000000
                            5375.0
                                         5350.000000
                                                         60 months
                                                                      12.69%
10
         6500.000000
                            6500.0
                                         6500.000000
                                                         60 months
                                                                      14.65%
11
       12000.000000
                           12000.0
                                        12000.000000
                                                         36 months
                                                                      12.69%
12
                                                                      13.49%
         9000.000000
                            9000.0
                                         9000.000000
                                                         36 months
13
                                                         36 months
                                                                       9.91%
         3000.000000
                            3000.0
                                         3000.000000
14
                           10000.0
                                        10000.000000
                                                         36 months
                                                                      10.65%
       10000.000000
15
         1000.000000
                            1000.0
                                         1000.000000
                                                         36 months
                                                                      16.29%
16
       10000.000000
                           10000.0
                                        10000.000000
                                                         36 months
                                                                      15.27%
17
         3600.000000
                            3600.0
                                         3600.000000
                                                         36 months
                                                                       6.03%
18
         6000.000000
                            6000.0
                                         6000.000000
                                                         36 months
                                                                      11.71%
19
                                                         36 months
         9200.000000
                            9200.0
                                         9200.000000
                                                                       6.03%
20
       20250.000000
                           20250.0
                                        19142.161077
                                                         60 months
                                                                      15.27%
21
       21000.000000
                           21000.0
                                        21000.000000
                                                         36 months
                                                                      12.42%
22
       10000.000000
                           10000.0
                                        10000.000000
                                                         36 months
                                                                      11.71%
23
                           10000.0
                                        10000.000000
                                                         36 months
                                                                      11.71%
       10000.000000
24
         6000.000000
                            6000.0
                                         6000.000000
                                                         36 months
                                                                      11.71%
25
       15000.000000
                           15000.0
                                        15000.000000
                                                         36 months
                                                                       9.91%
26
                           15000.0
                                                         36 months
                                                                      14.27%
       15000.000000
                                         8725.000000
27
         5000.000000
                            5000.0
                                         5000.000000
                                                         60 months
                                                                      16.77%
28
         4000.000000
                            4000.0
                                         4000.000000
                                                         36 months
                                                                      11.71%
29
         8500.000000
                                         8500.000000
                            8500.0
                                                         36 months
                                                                      11.71%
. . .
                                . . .
                                                                         . . .
42508
                                                                      13.12%
         6000.000000
                            6000.0
                                         1200.000000
                                                         36 months
42509
         5350.000000
                            5350.0
                                          625.000000
                                                         36 months
                                                                      13.12%
42510
                                          900.000000
                                                         36 months
                                                                       9.64%
         1900.000000
                            1900.0
42511
       10000.000000
                           10000.0
                                          350.000000
                                                         36 months
                                                                      14.70%
42512
         2000.000000
                            2000.0
                                         1275.000000
                                                         36 months
                                                                       7.12%
42513
         6000.000000
                            6000.0
                                          650.000000
                                                         36 months
                                                                      10.59%
```

| 42514 | 4400.000000 | 4400. | | | .000000 | | months | 9.64% |
|---|--|--|--|-----|---------|-------|---|--|
| 42515 | 1200.000000 | 1200. | | | .000000 | | months | 9.01% |
| 42516 | 5000.000000 | 5000. | | | .000000 | | months | 11.22% |
| 42517 | 1400.000000 | 1400. | | | .000000 | | months | 10.91% |
| 42518 | 1000.000000 | 1000. | | | .000000 | | months | 14.07% |
| 42519 | 5000.000000 | 5000. | 0 | | .000000 | | months | 7.75% |
| 42520 | 2500.000000 | 2500. | 0 | 225 | .000000 | 36 | months | 7.43% |
| 42521 | 3000.000000 | 3000. | 0 | 250 | .000000 | 36 | months | 7.43% |
| 42522 | 2600.000000 | 2600. | 0 | | .000000 | 36 | months | 8.38% |
| 42523 | 1000.000000 | 1000. | 0 | 625 | .000000 | 36 | months | 7.12% |
| 42524 | 1275.000000 | 1275. | 0 | 0. | .000000 | 36 | months | 12.49% |
| 42525 | 6450.000000 | 6450. | 0 | 0. | .000000 | 36 | months | 11.22% |
| 42526 | 10500.000000 | 10500. | 0 | 275 | .000000 | 36 | months | 11.22% |
| 42527 | 3000.000000 | 3000. | 0 | 125 | .000000 | 36 | months | 9.01% |
| 42528 | 3000.000000 | 3000. | 0 | 0. | .000000 | 36 | months | 9.33% |
| 42529 | 2000.000000 | 2000. | 0 | 225 | .000000 | 36 | months | 9.96% |
| 42530 | 6500.000000 | 6500. | 0 | 0. | .000000 | 36 | months | 9.64% |
| 42531 | 3500.000000 | 3500. | 0 | 225 | .000000 | 36 | months | 10.28% |
| 42532 | 1000.000000 | 1000. | 0 | 0. | .000000 | 36 | months | 9.64% |
| 42533 | 2525.000000 | 2525. | 0 | 225 | .000000 | 36 | months | 9.33% |
| 42534 | 6500.000000 | 6500. | 0 | 0. | .000000 | 36 | months | 8.38% |
| 42535 | 5000.000000 | 5000. | 0 | 0. | .000000 | 36 | months | 7.75% |
| 42536 | 11089.722581 | 9600. | | | NaN | | NaN | NaN |
| 42537 | 11089.722581 | 9600. | | | NaN | | NaN | NaN |
| | | | | | | | | |
| | | | | | | | | |
| | | | | | | fico | | |
| 0 | installment gra | | | | | _fico | _range_h | |
| 0 1 | installment gra | ade sub_gr | ade B2 | | | _fico | _range_h 74 | nigh \ 14.0 |
| 1 | installment grant 162.87 59.83 | ade sub_gr B | rade B2 C4 | | | _fico | _range_h 74 49 | nigh \ 14.0 99.0 |
| 1 2 | installment gr: 162.87 59.83 84.33 | ade sub_gr B C | rade B2 C4 C5 | | | _fico | _range_h 74 49 71 | nigh \ 14.0 99.0 |
| 1 2 3 | installment gra 162.87 59.83 84.33 339.31 | ade sub_gr B C C | rade B2 C4 C5 C1 | | | _fico | _range_t 74 49 71 60 | nigh \ 14.0 99.0 19.0 |
| 1 2 3 4 | installment gra 162.87 59.83 84.33 339.31 67.79 | ade sub_gr B C C C B | rade B2 C4 C5 C1 B5 | | | _fico | _range_t 74 49 71 60 | nigh \ 14.0 99.0 19.0 94.0 |
| 1 2 3 4 5 | installment gr: 162.87 59.83 84.33 339.31 67.79 156.46 | ade sub_gr B C C C B A | rade B2 C4 C5 C1 B5 | | | _fico | _range_t 74 49 71 60 69 | nigh \ 14.0 99.0 19.0 04.0 94.0 |
| 1 2 3 4 5 | installment gra 162.87 59.83 84.33 339.31 67.79 156.46 170.08 | ade sub_gr B C C C B A | cade B2 C4 C5 C1 B5 A4 C5 | | | _fico | _range_t 74 49 71 60 69 67 | nigh \ 14.0 99.0 19.0 94.0 94.0 79.0 |
| 1 2 3 4 5 6 7 | installment gr: 162.87 59.83 84.33 339.31 67.79 156.46 170.08 109.43 | ade sub_gr B C C C B A C | rade B2 C4 C5 C1 B5 A4 C5 | | | _fico | _range_t 74 49 71 60 69 67 68 | nigh \ 14.0 99.0 19.0 94.0 94.0 79.0 54.0 |
| 1 2 3 4 5 6 7 8 | installment gr: 162.87 59.83 84.33 339.31 67.79 156.46 170.08 109.43 152.39 | ade sub_gr B C C B A C E | rade B2 C4 C5 C1 B5 A4 C5 E1 | | | _fico | _range_t 74 49 71 60 69 67 68 49 | nigh \ 14.0 99.0 19.0 94.0 94.0 79.0 54.0 39.0 |
| 1 2 3 4 5 6 7 8 | installment gr: 162.87 59.83 84.33 339.31 67.79 156.46 170.08 109.43 152.39 121.45 | ade sub_gr B C C B A C E F | rade B2 C4 C5 C1 B5 A4 C5 E1 F2 | | | _fico | _range_f 74 49 71 60 69 67 65 49 | nigh \ 14.0 99.0 19.0 04.0 94.0 79.0 54.0 39.0 |
| 1 2 3 4 5 6 7 8 9 10 | installment gr: 162.87 59.83 84.33 339.31 67.79 156.46 170.08 109.43 152.39 121.45 153.45 | ade sub_gr B C C B A C E F B C | rade B2 C4 C5 C1 B5 A4 C5 E1 F2 B5 C3 | | | _fico | _range_t 74 49 71 60 69 67 68 49 51 | nigh \ 14.0 99.0 19.0 94.0 94.0 79.0 54.0 99.0 19.0 |
| 1 2 3 4 5 6 7 8 9 10 | installment gr: 162.87 59.83 84.33 339.31 67.79 156.46 170.08 109.43 152.39 121.45 153.45 402.54 | ade sub_gr B C C B A C E F B C | rade B2 C4 C5 C1 B5 A4 C5 E1 F2 B5 C3 B5 | | | _fico | _range_t 74 49 71 60 69 67 68 49 51 73 | nigh \ 14.0 99.0 19.0 94.0 94.0 79.0 54.0 39.0 19.0 34.0 |
| 1 2 3 4 5 6 7 8 9 10 11 12 | installment gr: 162.87 59.83 84.33 339.31 67.79 156.46 170.08 109.43 152.39 121.45 153.45 402.54 305.38 | ade sub_gr B C C B A C E F B C B C | rade B2 C4 C5 C1 B5 A4 C5 E1 F2 B5 C3 B5 C1 | | | _fico | _range_t 74 49 71 60 69 67 68 49 51 73 66 | nigh \ 14.0 99.0 19.0 04.0 94.0 94.0 39.0 39.0 19.0 34.0 |
| 1 2 3 4 5 6 7 8 9 10 11 12 13 | installment gr: 162.87 59.83 84.33 339.31 67.79 156.46 170.08 109.43 152.39 121.45 153.45 402.54 305.38 96.68 | ade sub_gr B C C B A C E F B C B C B | rade B2 C4 C5 C1 B5 A4 C5 E1 F2 B5 C3 B5 C1 B1 | | | _fico | _range_t 74 49 71 60 69 67 68 49 51 73 66 61 73 | nigh \ 14.0 99.0 19.0 94.0 94.0 99.0 54.0 99.0 19.0 19.0 19.0 |
| 1 2 3 4 5 6 7 8 9 10 11 12 13 14 | installment gr: 162.87 59.83 84.33 339.31 67.79 156.46 170.08 109.43 152.39 121.45 153.45 402.54 305.38 96.68 325.74 | ade sub_gr B C C B A C E F B C B C B B C B B | Cade B2 C4 C5 C1 B5 A4 C5 E1 F2 B5 C3 B5 C1 B1 B2 | | | _fico | _range_t 74 49 71 60 69 67 68 49 51 73 66 61 73 | nigh \ 14.0 99.0 19.0 94.0 94.0 99.0 54.0 39.0 19.0 34.0 39.0 |
| 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 | installment gr: 162.87 59.83 84.33 339.31 67.79 156.46 170.08 109.43 152.39 121.45 153.45 402.54 305.38 96.68 325.74 35.31 | ade sub_gr B C C B A C E F B C B C B D | rade B2 C4 C5 C1 B5 A4 C5 E1 F2 B5 C3 B5 C1 B1 B2 D1 | | | _fico | _range_t 74 49 71 60 69 67 65 68 49 51 73 66 61 73 | nigh \ 14.0 99.0 19.0 04.0 94.0 94.0 39.0 39.0 19.0 34.0 39.0 34.0 39.0 |
| 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 | installment gr: 162.87 59.83 84.33 339.31 67.79 156.46 170.08 109.43 152.39 121.45 153.45 402.54 305.38 96.68 325.74 35.31 347.98 | ade sub_gr B C C B A C E F B C B C B C C C C C C C C C C C C C C | Cade B2 C4 C5 C1 B5 A4 C5 E1 F2 B5 C3 B5 C1 B1 B2 D1 C4 | | | _fico | _range_f 74 49 71 60 69 67 65 68 49 51 73 66 61 73 65 65 | nigh \ 14.0 99.0 19.0 94.0 94.0 94.0 39.0 19.0 34.0 39.0 19.0 34.0 39.0 39.0 |
| 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 | installment gr: 162.87 59.83 84.33 339.31 67.79 156.46 170.08 109.43 152.39 121.45 153.45 402.54 305.38 96.68 325.74 35.31 347.98 109.57 | ade sub_gr B C C B A C E F B C B C B C A | rade B2 C4 C5 C1 B5 A4 C5 E1 F2 B5 C3 B5 C1 B1 B2 D1 C4 A1 | | | _fico | _range_t 74 49 71 60 69 67 68 49 51 73 66 61 73 68 79 | nigh \ 14.0 99.0 19.0 94.0 94.0 94.0 39.0 19.0 34.0 39.0 19.0 34.0 39.0 39.0 |
| 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 | installment gr: 162.87 59.83 84.33 339.31 67.79 156.46 170.08 109.43 152.39 121.45 153.45 402.54 305.38 96.68 325.74 35.31 347.98 109.57 198.46 | ade sub_gr B C C B A C E F B C B C B C B B C B B C B B B C B B B C B B B C B B B B C B B B B B B C B | rade B2 C4 C5 C1 B5 A4 C5 E1 F2 B5 C3 B5 C1 B1 B2 D1 C4 A1 B3 | | | _fico | _range_t 74 49 71 60 69 67 68 49 51 73 66 61 73 68 76 63 78 | nigh \ 14.0 99.0 19.0 94.0 94.0 94.0 39.0 99.0 19.0 34.0 39.0 34.0 39.0 34.0 39.0 |
| 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 | installment gr: 162.87 59.83 84.33 339.31 67.79 156.46 170.08 109.43 152.39 121.45 153.45 402.54 305.38 96.68 325.74 35.31 347.98 109.57 198.46 280.01 | ade sub_gr B C C B A C E F B C B C B A C B A C B A C B A C B A A A | rade B2 C4 C5 C1 B5 A4 C5 E1 F2 B5 C3 B5 C1 B1 B2 D1 C4 A1 B3 A1 | | | _fico | _range_f 74 49 71 60 69 67 65 68 49 51 73 66 67 65 65 66 67 65 66 67 66 67 66 67 66 67 66 67 66 67 66 67 66 67 66 67 66 67 66 67 67 | nigh \ 14.0 99.0 19.0 94.0 94.0 94.0 39.0 19.0 34.0 39.0 34.0 39.0 34.0 39.0 34.0 39.0 |
| 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 | installment gr: 162.87 59.83 84.33 339.31 67.79 156.46 170.08 109.43 152.39 121.45 153.45 402.54 305.38 96.68 325.74 35.31 347.98 109.57 198.46 | ade sub_gr B C C B A C E F B C B C B C B B C B B C B B B C B B B C B B B C B B B B C B B B B B B C B | rade B2 C4 C5 C1 B5 A4 C5 E1 F2 B5 C3 B5 C1 B1 B2 D1 C4 A1 B3 | | | _fico | _range_t 74 49 71 60 69 67 68 49 51 73 66 61 73 68 79 58 64 | nigh \ 14.0 99.0 19.0 94.0 94.0 94.0 39.0 99.0 19.0 34.0 39.0 34.0 39.0 34.0 39.0 |

| 22 | 330.76 | В | В3 | | 779.0 | |
|-------|---------------|-------|-------------|-----------------|-----------|---|
| 23 | 330.76 | В | В3 | | 684.0 | |
| 24 | 198.46 | В | В3 | • • • | 499.0 | |
| 25 | 483.38 | В | B1 | | 734.0 | |
| 26 | 514.64 | С | C2 | | 509.0 | |
| 27 | 123.65 | D | D2 | | 579.0 | |
| 28 | 132.31 | В | В3 | | 689.0 | |
| 29 | 281.15 | В | B3 | | 599.0 | |
| | | | | ••• | | |
| 42508 | | D | D5 | ••• | 664.0 | |
| 42509 | | D | D5 | • • • | 704.0 | |
| 42510 | | В | B4 | • • • | 709.0 | |
| 42510 | | E | E5 | ••• | 499.0 | |
| | | | | • • • | | |
| 42512 | | A | A1 | • • • | 809.0 | |
| 42513 | | C | C2 | • • • | 769.0 | |
| 42514 | | В | B4 | • • • | 584.0 | |
| 42515 | | В | B2 | • • • | 694.0 | |
| 42516 | | С | C4 | • • • | 679.0 | |
| 42517 | | C | C3 | • • • | 679.0 | |
| 42518 | | E | E3 | • • • | 499.0 | |
| 42519 | 156.11 | Α | A3 | • • • | 759.0 | |
| 42520 | 77.69 | Α | A2 | | 789.0 | |
| 42521 | 93.23 | Α | A2 | | 799.0 | |
| 42522 | 81.94 | Α | A 5 | • • • | 614.0 | |
| 42523 | 30.94 | Α | A1 | • • • | 599.0 | |
| 42524 | 42.65 | D | D3 | | 624.0 | |
| 42525 | 211.85 | C | C4 | | 574.0 | |
| 42526 | 344.87 | С | C4 | | 744.0 | |
| 42527 | 95.42 | В | B2 | • • • | 594.0 | |
| 42528 | | В | В3 | • • • | 714.0 | |
| 42529 | | В | В5 | | 594.0 | |
| 42530 | | В | B4 | | 684.0 | |
| 42531 | | С | C1 | • • • | 819.0 | |
| 42532 | | В | B4 | • • • | 784.0 | |
| 42533 | | В | В3 | ••• | 714.0 | |
| 42534 | | A | A5 | • • • | 724.0 | |
| 42535 | | A | A3 | • • • | 794.0 | |
| 42536 | | NaN | NaN | • • • | NaN | |
| | | | | • • • | | |
| 42537 | NaN | NaN | NaN | • • • | NaN | |
| | lost fice mem |] | aallaa+iama | 10 m+ha or mod | no]: codo | \ |
| ^ | last_lico_ran | _ | corrections | _12_mths_ex_med | - | \ |
| 0 | | 740.0 | | False | True | |
| 1 | | 0.0 | | False | True | |
| 2 | | 715.0 | | False | True | |
| 3 | | 600.0 | | False | True | |
| 4 | | 690.0 | | False | True | |
| 5 | | 675.0 | | False | True | |
| 6 | | 650.0 | | False | True | |

| 7 | 685.0 | False | True |
|-------|-------|-------|------|
| 8 | 0.0 | False | True |
| 9 | 515.0 | False | True |
| 10 | 730.0 | False | True |
| 11 | 665.0 | False | True |
| 12 | 615.0 | False | True |
| 13 | 730.0 | False | True |
| 14 | 650.0 | False | True |
| 15 | 765.0 | False | True |
| 16 | 635.0 | False | True |
| 17 | 790.0 | False | True |
| 18 | 555.0 | False | True |
| 19 | 640.0 | False | True |
| 20 | 680.0 | False | True |
| 21 | 530.0 | False | True |
| 22 | 775.0 | False | True |
| 23 | 680.0 | False | True |
| 24 | 0.0 | False | True |
| 25 | 730.0 | False | True |
| 26 | 505.0 | False | True |
| 27 | 575.0 | False | True |
| 28 | 685.0 | False | True |
| 29 | 595.0 | False | True |
| | ••• | • • • | |
| 42508 | 660.0 | NaN | True |
| 42509 | 700.0 | NaN | True |
| 42510 | 705.0 | NaN | True |
| 42511 | 0.0 | NaN | True |
| 42512 | 805.0 | NaN | True |
| 42513 | 765.0 | NaN | True |
| 42514 | 580.0 | NaN | True |
| 42515 | 690.0 | NaN | True |
| 42516 | 675.0 | NaN | True |
| 42517 | 675.0 | NaN | True |
| 42518 | 0.0 | NaN | True |
| 42519 | 755.0 | NaN | True |
| 42520 | 785.0 | NaN | True |
| 42521 | 795.0 | NaN | True |
| 42522 | 610.0 | NaN | True |
| 42523 | 595.0 | NaN | True |
| 42524 | 620.0 | NaN | True |
| 42525 | 570.0 | NaN | True |
| 42526 | 740.0 | NaN | True |
| 42527 | 590.0 | NaN | True |
| 42528 | 710.0 | NaN | True |
| 42529 | 590.0 | NaN | True |
| 42530 | 680.0 | NaN | True |
| 42531 | 815.0 | NaN | True |
| | | | |

| 42532 | 780 |) () | NaN | Тъ | rue | |
|-------|------------------|----------------|-----------------|-----------|-------------|---|
| 42533 | 710 | | NaN | | rue | |
| 42534 | 720 | | NaN | | rue | |
| 42535 | 790 | | NaN | | rue | |
| 42536 | | VaN | NaN | | JaN | |
| 42537 | | van Van | NaN | | van Van | |
| 12001 | · | vaiv | waw | | vaiv | |
| | application_type | acc_now_deling | chargeoff_withi | n_12_mths | delinq_amnt | \ |
| 0 | INDIVIDUAL | False | | False | 0.0 | |
| 1 | INDIVIDUAL | False | | False | 0.0 | |
| 2 | INDIVIDUAL | False | | False | 0.0 | |
| 3 | INDIVIDUAL | False | | False | 0.0 | |
| 4 | INDIVIDUAL | False | | False | 0.0 | |
| 5 | INDIVIDUAL | False | | False | 0.0 | |
| 6 | INDIVIDUAL | False | | False | 0.0 | |
| 7 | INDIVIDUAL | False | | False | 0.0 | |
| 8 | INDIVIDUAL | False | | False | 0.0 | |
| 9 | INDIVIDUAL | False | | False | 0.0 | |
| 10 | INDIVIDUAL | False | | False | 0.0 | |
| 11 | INDIVIDUAL | False | | False | 0.0 | |
| 12 | INDIVIDUAL | False | | False | 0.0 | |
| 13 | INDIVIDUAL | False | | False | 0.0 | |
| 14 | INDIVIDUAL | False | | False | 0.0 | |
| 15 | INDIVIDUAL | False | | False | 0.0 | |
| 16 | INDIVIDUAL | False | | False | 0.0 | |
| 17 | INDIVIDUAL | False | | False | 0.0 | |
| 18 | INDIVIDUAL | False | | False | 0.0 | |
| 19 | INDIVIDUAL | False | | False | 0.0 | |
| 20 | INDIVIDUAL | False | | False | 0.0 | |
| 21 | INDIVIDUAL | False | | False | 0.0 | |
| 22 | INDIVIDUAL | False | | False | 0.0 | |
| 23 | INDIVIDUAL | False | | False | 0.0 | |
| 24 | INDIVIDUAL | False | | False | 0.0 | |
| 25 | INDIVIDUAL | False | | False | 0.0 | |
| 26 | INDIVIDUAL | False | | False | 0.0 | |
| 27 | INDIVIDUAL | False | | False | 0.0 | |
| 28 | INDIVIDUAL | False | | False | 0.0 | |
| 29 | INDIVIDUAL | False | | False | 0.0 | |
| • • • | • • • | ••• | | • • • | • • • | |
| 42508 | INDIVIDUAL | False | | NaN | 0.0 | |
| 42509 | INDIVIDUAL | False | | NaN | 0.0 | |
| 42510 | INDIVIDUAL | NaN | | NaN | NaN | |
| 42511 | INDIVIDUAL | False | | NaN | 0.0 | |
| 42512 | INDIVIDUAL | False | | NaN | 0.0 | |
| 42513 | INDIVIDUAL | False | | NaN | 0.0 | |
| 42514 | INDIVIDUAL | False | | NaN | 0.0 | |
| 42515 | INDIVIDUAL | NaN | | NaN | NaN | |
| 42516 | INDIVIDUAL | NaN | | NaN | NaN | |

| 42517 | INDIVIDUAL | NaN | NaN | NaN |
|-------|------------|-----|-----|-----|
| 42518 | INDIVIDUAL | NaN | NaN | NaN |
| 42519 | INDIVIDUAL | NaN | NaN | NaN |
| 42520 | INDIVIDUAL | NaN | NaN | NaN |
| 42521 | INDIVIDUAL | NaN | NaN | NaN |
| 42522 | INDIVIDUAL | NaN | NaN | NaN |
| 42523 | INDIVIDUAL | NaN | NaN | NaN |
| 42524 | INDIVIDUAL | NaN | NaN | NaN |
| 42525 | INDIVIDUAL | NaN | NaN | NaN |
| 42526 | INDIVIDUAL | NaN | NaN | NaN |
| 42527 | INDIVIDUAL | NaN | NaN | NaN |
| 42528 | INDIVIDUAL | NaN | NaN | NaN |
| 42529 | INDIVIDUAL | NaN | NaN | NaN |
| 42530 | INDIVIDUAL | NaN | NaN | NaN |
| 42531 | INDIVIDUAL | NaN | NaN | NaN |
| 42532 | INDIVIDUAL | NaN | NaN | NaN |
| 42533 | INDIVIDUAL | NaN | NaN | NaN |
| 42534 | INDIVIDUAL | NaN | NaN | NaN |
| 42535 | INDIVIDUAL | NaN | NaN | NaN |
| 42536 | NaN | NaN | NaN | NaN |
| 42537 | NaN | NaN | NaN | NaN |

| 1 | 0.0 | False |
|----|-----|-------|
| 2 | 0.0 | False |
| 3 | 0.0 | False |
| 4 | 0.0 | False |
| 5 | 0.0 | False |
| 6 | 0.0 | False |
| 7 | 0.0 | False |
| 8 | 0.0 | False |
| 9 | 0.0 | False |
| 10 | 0.0 | False |
| 11 | 0.0 | False |
| 12 | 0.0 | False |
| 13 | 0.0 | False |
| 14 | 0.0 | False |
| 15 | 0.0 | False |
| 16 | 0.0 | False |
| 17 | 0.0 | False |
| 18 | 0.0 | False |
| 19 | 0.0 | False |
| 20 | 0.0 | False |
| 21 | 0.0 | False |
| 22 | 0.0 | False |
| 23 | 0.0 | False |
| 24 | 0.0 | False |
| | | |

| 25 | 0.0 | False |
|-------|-----|-------|
| 26 | 0.0 | False |
| 27 | 0.0 | False |
| 28 | 0.0 | False |
| 29 | 0.0 | False |
| • • • | | |
| 42508 | NaN | NaN |
| 42509 | NaN | NaN |
| 42510 | NaN | NaN |
| 42511 | NaN | NaN |
| 42512 | NaN | NaN |
| 42513 | NaN | NaN |
| 42514 | NaN | NaN |
| 42515 | NaN | NaN |
| 42516 | NaN | NaN |
| 42517 | NaN | NaN |
| 42518 | NaN | NaN |
| 42519 | NaN | NaN |
| 42520 | NaN | NaN |
| 42521 | NaN | NaN |
| 42522 | NaN | NaN |
| 42523 | NaN | NaN |
| 42524 | NaN | NaN |
| 42525 | NaN | NaN |
| 42526 | NaN | NaN |
| 42527 | NaN | NaN |
| 42528 | NaN | NaN |
| 42529 | NaN | NaN |
| 42530 | NaN | NaN |
| 42531 | NaN | NaN |
| 42532 | NaN | NaN |
| 42533 | NaN | NaN |
| 42534 | NaN | NaN |
| 42535 | NaN | NaN |
| 42536 | NaN | NaN |
| 42537 | NaN | NaN |
| | | |

[42538 rows x 58 columns]

1.1.4 Imputing missing data with mean, median or most frequently used value for the column we will create the Imputer object and set a strategy for handling the missing values. Different strategy available are mean, median and most_frequent we then fit the imputter objects on the column where we want to handle the missing values After fitting the data we transform the data from the dataframe

```
Out[11]:
             Age Country
                            Salary
                   France
         0
             NaN
                               NaN
         1 45.0
                    Spain
                           90000.0
         2
            NaN Germany
                               NaN
         3 32.0
                      USA
                          75000.0
In [12]: imp = Imputer(missing_values='NaN', strategy='mean', axis=0)
         imp.fit(countryData.iloc[:,[0,2]])
         countryData.iloc[:,[0,2]]=imp.transform(countryData.iloc[:,[0,2]])
         countryData
Out[12]:
             Age Country
                            Salary
         0 38.5
                   France 82500.0
                    Spain 90000.0
         1 45.0
         2 38.5 Germany
                           82500.0
         3 32.0
                      USA
                           75000.0
  1.1.5 Using replace method
In [13]: countryData = pd.DataFrame({"Country":["France", "Spain", "Germany", "USA"], "Age":[np
         countryData
Out[13]:
             Age Country
                            Salary
            {\tt NaN}
                   France
                               NaN
         1 45.0
                           90000.0
                    Spain
            NaN Germany
                               NaN
         3 32.0
                      USA
                          75000.0
In [14]: countryData.replace({'Age':np.NAN},40 )
Out[14]:
                            Salary
             Age Country
         0 40.0
                   France
                               NaN
         1 45.0
                    Spain
                           90000.0
         2 40.0 Germany
                               NaN
         3 32.0
                      USA
                          75000.0
  1.2 Removing duplicate entries from the dataset
In [15]: df.drop_duplicates(keep='first')
Out[15]:
                                                              id member_id \
         0
                                                        1077501 1296599.0
         1
                                                        1077430 1314167.0
         2
                                                        1077175 1313524.0
         3
                                                        1076863 1277178.0
         4
                                                        1075358 1311748.0
         5
                                                        1075269 1311441.0
         6
                                                        1069639 1304742.0
         7
                                                        1072053 1288686.0
```

| 8 | 1071795 | 1306957.0 |
|-------|---------|-----------|
| 9 | 1071730 | 1306721.0 |
| 10 | 1071370 | 1305721.0 |
| | | |
| 11 | 1069908 | 1305008.0 |
| 12 | 1064687 | 1298717.0 |
| 13 | 1069866 | 1304956.0 |
| 14 | 1069057 | 1303503.0 |
| 15 | 1069759 | 1304871.0 |
| 16 | 1065775 | 1299699.0 |
| 17 | 1069971 | 1304884.0 |
| 18 | 1062474 | 1294539.0 |
| 19 | 1069742 | 1304855.0 |
| 20 | 1069740 | 1284848.0 |
| 21 | 1039153 | 1269083.0 |
| 22 | 1069710 | 1304821.0 |
| 23 | 1069700 | 1304810.0 |
| 24 | 1069559 | 1304634.0 |
| 25 | 1069697 | 1273773.0 |
| 26 | 1069800 | 1304679.0 |
| 27 | 1069657 | 1304764.0 |
| 28 | 1069799 | 1304678.0 |
| 29 | 1047704 | 1278806.0 |
| ••• | | |
| 42508 | 91175 | 91170.0 |
| 42509 | 91126 | 91067.0 |
| 42510 | 91023 | 70879.0 |
| 42511 | 90106 | 90090.0 |
| 42512 | 89258 | 80039.0 |
| 42513 | 88637 | 88629.0 |
| 42514 | 88046 | 88023.0 |
| 42515 | 85961 | 85923.0 |
| 42516 | 85818 | 85802.0 |
| 42517 | 85781 | 85727.0 |
| 42518 | 85675 | 85667.0 |
| 42519 | 84670 | 79576.0 |
| 42520 | 84098 | 84091.0 |
| 42521 | 83979 | 83974.0 |
| 42522 | 83489 | 83471.0 |
| 42523 | 83185 | 83132.0 |
| 42524 | | |
| | 76629 | 76623.0 |
| 42525 | 74014 | 73890.0 |
| 42526 | 81085 | 80973.0 |
| 42527 | 77792 | 77764.0 |
| 42528 | 77757 | 70626.0 |
| 42529 | 74505 | 74469.0 |
| 42530 | 74323 | 74301.0 |
| 42531 | 73582 | 73096.0 |
| 42532 | 72998 | 72992.0 |

```
42533
                                                     72176
                                                               70868.0
42534
                                                     71623
                                                               70735.0
42535
                                                     70686
                                                               70681.0
       Total amount funded in policy code 1: 460296150
42536
                                                                   NaN
                Total amount funded in policy code 2: 0
42537
                                                                   NaN
           loan amnt
                       funded amnt
                                     funded amnt inv
                                                              term int_rate
0
         5000.000000
                            5000.0
                                         4975.000000
                                                        36 months
                                                                      10.65%
1
         2500.000000
                            2500.0
                                         2500.000000
                                                        60 months
                                                                      15.27%
2
         2400.000000
                            2400.0
                                         2400.000000
                                                        36 months
                                                                      15.96%
3
                                                        36 months
                                                                      13.49%
       10000.000000
                           10000.0
                                        10000.000000
4
                                         3000.000000
         3000.000000
                            3000.0
                                                         60 months
                                                                      12.69%
5
                                                                       7.90%
         5000.000000
                            5000.0
                                         5000.000000
                                                         36 months
6
         7000.000000
                            7000.0
                                         7000.000000
                                                        60 months
                                                                      15.96%
7
         3000.000000
                            3000.0
                                         3000.000000
                                                         36 months
                                                                      18.64%
8
                                                                      21.28%
         5600.000000
                            5600.0
                                         5600.000000
                                                         60 months
9
                                                        60 months
                                                                      12.69%
         5375.000000
                            5375.0
                                         5350.000000
                                                                      14.65%
10
         6500.000000
                            6500.0
                                         6500.000000
                                                        60 months
11
                                                         36 months
       12000.000000
                           12000.0
                                        12000.000000
                                                                      12.69%
12
                                                         36 months
                                                                      13.49%
         9000.000000
                            9000.0
                                         9000.000000
13
         3000.000000
                            3000.0
                                         3000.000000
                                                         36 months
                                                                       9.91%
14
       10000.000000
                           10000.0
                                        10000.000000
                                                        36 months
                                                                      10.65%
15
         1000.000000
                            1000.0
                                         1000.000000
                                                         36 months
                                                                      16.29%
16
                                                        36 months
       10000.000000
                           10000.0
                                        10000.000000
                                                                      15.27%
17
                                                        36 months
                                                                       6.03%
         3600.000000
                            3600.0
                                         3600.000000
18
                                                         36 months
         6000.000000
                            6000.0
                                         6000.000000
                                                                      11.71%
19
                            9200.0
         9200.000000
                                         9200.000000
                                                         36 months
                                                                       6.03%
20
       20250.000000
                           20250.0
                                        19142.161077
                                                         60 months
                                                                      15.27%
21
       21000.000000
                           21000.0
                                        21000.000000
                                                        36 months
                                                                      12.42%
22
                                                         36 months
                                                                      11.71%
       10000.000000
                           10000.0
                                        10000.000000
23
                                                         36 months
       10000.000000
                           10000.0
                                        10000.000000
                                                                      11.71%
24
                                                        36 months
                                                                      11.71%
         6000.000000
                            6000.0
                                         6000.000000
25
                           15000.0
                                        15000.000000
                                                         36 months
                                                                       9.91%
       15000.000000
26
                                                                      14.27%
       15000.000000
                           15000.0
                                         8725.000000
                                                        36 months
27
                                                                      16.77%
                                                         60 months
         5000.000000
                            5000.0
                                         5000.000000
28
         4000.000000
                            4000.0
                                         4000.000000
                                                        36 months
                                                                      11.71%
29
         8500.000000
                            8500.0
                                         8500.000000
                                                         36 months
                                                                      11.71%
                                                                         . . .
. . .
                                . . .
42508
         6000.000000
                            6000.0
                                         1200.000000
                                                        36 months
                                                                      13.12%
42509
                                                        36 months
                                                                      13.12%
         5350.000000
                            5350.0
                                          625.000000
42510
         1900.000000
                            1900.0
                                          900.000000
                                                        36 months
                                                                       9.64%
42511
        10000.000000
                           10000.0
                                          350.000000
                                                         36 months
                                                                      14.70%
42512
                                                         36 months
                                                                       7.12%
         2000.000000
                            2000.0
                                         1275.000000
42513
         6000.000000
                            6000.0
                                          650.000000
                                                         36 months
                                                                      10.59%
42514
                                                         36 months
                                                                       9.64%
         4400.000000
                            4400.0
                                         1400.000000
42515
         1200.000000
                            1200.0
                                          500.000000
                                                         36 months
                                                                       9.01%
42516
         5000.000000
                            5000.0
                                          375.000000
                                                         36 months
                                                                      11.22%
42517
         1400.000000
                            1400.0
                                          475.000000
                                                         36 months
                                                                      10.91%
```

| 42518 | 1000.000000 | 1000.0 | 625.000000 36 months 14.07% |
|-------|-----------------|---------------|------------------------------------|
| 42519 | 5000.000000 | 5000.0 | 300.000000 36 months 7.75% |
| 42520 | 2500.000000 | 2500.0 | 225.000000 36 months 7.43% |
| 42521 | 3000.000000 | 3000.0 | 250.000000 36 months 7.43% |
| 42522 | 2600.000000 | 2600.0 | 575.000000 36 months 8.38% |
| 42523 | 1000.000000 | 1000.0 | 625.000000 36 months 7.12% |
| 42524 | 1275.000000 | 1275.0 | 0.000000 36 months 12.49% |
| 42525 | 6450.000000 | 6450.0 | 0.000000 36 months 11.22% |
| 42526 | 10500.000000 | 10500.0 | 275.000000 36 months 11.22% |
| 42527 | 3000.000000 | 3000.0 | 125.000000 36 months 9.01% |
| 42528 | 3000.000000 | 3000.0 | 0.000000 36 months 9.33% |
| 42529 | 2000.000000 | 2000.0 | 225.000000 36 months 9.96% |
| 42530 | 6500.000000 | 6500.0 | 0.000000 36 months 9.64% |
| 42531 | 3500.000000 | 3500.0 | 225.000000 36 months 10.28% |
| 42532 | 1000.000000 | 1000.0 | 0.000000 36 months 9.64% |
| 42533 | 2525.000000 | 2525.0 | 225.000000 36 months 9.33% |
| 42534 | 6500.000000 | 6500.0 | 0.000000 36 months 8.38% |
| 42535 | 5000.000000 | 5000.0 | 0.000000 36 months 7.75% |
| 42536 | 11089.722581 | 9600.0 | NaN NaN NaN |
| 42537 | 11089.722581 | 9600.0 | NaN NaN NaN |
| | | | |
| | installment gra | ade sub_grade | <pre> last_fico_range_high \</pre> |
| 0 | 162.87 | B B2 | 744.0 |
| 1 | 59.83 | C C4 | 499.0 |
| 2 | 84.33 | C C5 | 719.0 |
| 3 | 339.31 | C C1 | 604.0 |
| 4 | 67.79 | B B5 | 694.0 |
| 5 | 156.46 | A A4 | 679.0 |
| 6 | 170.08 | C C5 | 654.0 |
| 7 | 109.43 | E E1 | 689.0 |
| 8 | 152.39 | F F2 | 499.0 |
| 9 | 121.45 | В В5 | 519.0 |
| 10 | 153.45 | C C3 | 734.0 |
| 11 | 402.54 | В В5 | 669.0 |
| 12 | 305.38 | C C1 | 619.0 |
| 13 | 96.68 | B B1 | 734.0 |
| 14 | 325.74 | В В2 | 654.0 |
| 15 | 35.31 | D D1 | 769.0 |
| 16 | 347.98 | C C4 | 639.0 |
| 17 | 109.57 | A A1 | 794.0 |
| 18 | 198.46 | В ВЗ | |
| 19 | 280.01 | A A1 | |
| 20 | 484.63 | C C4 | |
| 21 | 701.73 | B B4 | |
| 22 | 330.76 | В ВЗ | |
| 23 | 330.76 | В ВЗ | |
| 24 | 198.46 | В ВЗ | |
| 25 | 483.38 | B B1 | |
| | | | |

| 26 | 514.64 | C | C2 | | | 509.0 | |
|-------|---------------|-------|------------|-----------|-------|-------|---|
| 27 | 123.65 | D | D2 | | | 579.0 | |
| 28 | 132.31 | В | В3 | | | 689.0 | |
| 29 | 281.15 | В | В3 | | | 599.0 | |
| | | | | | | | |
| 42508 | 202.51 | D | D5 | | | 664.0 | |
| 42509 | 180.57 | D | D5 | | | 704.0 | |
| 42510 | 61.00 | В | B4 | | | 709.0 | |
| 42511 | 345.18 | E | E5 | | | 499.0 | |
| 42512 | 61.87 | Α | A1 | | | 809.0 | |
| 42513 | 195.28 | C | C2 | | | 769.0 | |
| 42514 | 141.25 | В | B4 | | | 584.0 | |
| 42515 | 38.17 | В | B2 | | | 694.0 | |
| 42516 | 164.23 | C | C4 | | | 679.0 | |
| 42517 | 45.78 | C | C3 | | | 679.0 | |
| 42518 | 34.21 | E | E3 | | | 499.0 | |
| 42519 | 156.11 | Α | A3 | | | 759.0 | |
| 42520 | 77.69 | Α | A2 | | | 789.0 | |
| 42521 | 93.23 | Α | A2 | | | 799.0 | |
| 42522 | | Α | A 5 | | | 614.0 | |
| 42523 | | Α | A1 | | | 599.0 | |
| 42524 | | D | D3 | | | 624.0 | |
| 42525 | | С | C4 | | | 574.0 | |
| 42526 | | C | C4 | | | 744.0 | |
| 42527 | | В | В2 | | | 594.0 | |
| 42528 | | В | В3 | | | 714.0 | |
| 42529 | | В | В5 | | | 594.0 | |
| 42530 | | В | В4 | | | 684.0 | |
| 42531 | | С | C1 | | | 819.0 | |
| 42532 | | В | В4 | | | 784.0 | |
| 42533 | | В | В3 | | | 714.0 | |
| 42534 | | Α | A 5 | | | 724.0 | |
| 42535 | | Α | A3 | | | 794.0 | |
| 42536 | | NaN | NaN | | | NaN | |
| 42537 | | NaN | NaN | | | NaN | |
| | log+ #4.5 | .ma] | aallaa+:- | a 10±1 | | mali | ` |
| 0 | last_fico_ran | _ | collection | S_12_mtns | | - • | \ |
| 0 | | 740.0 | | | False | True | |
| 1 | | 0.0 | | | False | True | |
| 2 | | 715.0 | | | False | True | |
| 3 | | 600.0 | | | False | True | |
| 4 | | 690.0 | | | False | True | |
| 5 | | 675.0 | | | False | True | |
| 6 | | 650.0 | | | False | True | |
| 7 | | 685.0 | | | False | True | |
| 8 | | 0.0 | | | False | True | |
| 9 | | 515.0 | | | False | True | |
| 10 | | 730.0 | | | False | True | |

| 11 | 665.0 | False | True |
|----------------|----------------|------------|--------------|
| 12 | 615.0 | False | True |
| 13 | 730.0 | False | True |
| 14 | 650.0 | False | True |
| 15 | 765.0 | False | True |
| 16 | 635.0 | False | True |
| 17 | 790.0 | False | True |
| 18 | 555.0 | False | True |
| 19 | 640.0 | False | True |
| 20 | 680.0 | False | True |
| 21 | 530.0 | False | True |
| 22 | 775.0 | False | True |
| 23 | 680.0 | False | True |
| 24 | 0.0 | False | True |
| 25 | 730.0 | False | True |
| 26 | 505.0 | False | True |
| 27 | 575.0 | False | True |
| 28 | 685.0 | False | True |
| 29 | 595.0 | False | True |
| | • • • • | | |
| 42508 | 660.0 | NaN | True |
| 42509 | 700.0 | NaN | True |
| 42510 | 705.0 | NaN | True |
| 42511 | 0.0 | NaN | True |
| 42512 | 805.0 | NaN | True |
| 42513 | 765.0 | NaN | True |
| 42514 | 580.0 | NaN | True |
| 42515 | 690.0 | NaN | True |
| 42516 | 675.0 | NaN | True |
| 42517 | 675.0 | NaN | True |
| 42518 | 0.0 | NaN | True |
| 42519 | 755.0 | NaN | True |
| 42520 | 785.0 | NaN | True |
| 42521 | 795.0 | NaN | True |
| 42522 | 610.0 | NaN | True |
| 42523 | 595.0 | NaN | True |
| 42524 | 620.0 | NaN | True |
| 42525 | 570.0 | NaN | True |
| 42526 | 740.0 | NaN | True |
| 42527 | 590.0 | NaN | True |
| 42528 | 710.0 | NaN | True |
| 42529 | 590.0 | | |
| 42529 42530 | 680.0 | NaN NaN | True True |
| 42530 | 815.0 | NaN | True |
| 42531 | 780.0 | | |
| | | NaN | True |
| 42533 42534 | 710.0 720.0 | NaN NaN | True |
| 42534 42535 | 720.0 | NaN NaN | True |
| 42000 | 130.0 | NaN | True |

| 42536 | r | vaiv | nan i | Nan | |
|-------|-------------|----------------|-------------------------------------|-------|---|
| 42537 | N | VaN | NaN 1 | NaN | |
| | - | | | | |
| | | | | | |
| | | acc_now_delinq | <pre>chargeoff_within_12_mths</pre> | - | \ |
| 0 | INDIVIDUAL | False | False | 0.0 | |
| 1 | INDIVIDUAL | False | False | 0.0 | |
| 2 | INDIVIDUAL | False | False | 0.0 | |
| | | | | | |
| 3 | INDIVIDUAL | False | False | 0.0 | |
| 4 | INDIVIDUAL | False | False | 0.0 | |
| 5 | INDIVIDUAL | False | False | 0.0 | |
| 6 | INDIVIDUAL | False | False | 0.0 | |
| 7 | INDIVIDUAL | False | False | 0.0 | |
| 8 | INDIVIDUAL | False | False | 0.0 | |
| 9 | INDIVIDUAL | False | False | 0.0 | |
| | | | | | |
| 10 | INDIVIDUAL | False | False | 0.0 | |
| 11 | INDIVIDUAL | False | False | 0.0 | |
| 12 | INDIVIDUAL | False | False | 0.0 | |
| 13 | INDIVIDUAL | False | False | 0.0 | |
| 14 | INDIVIDUAL | False | False | 0.0 | |
| 15 | INDIVIDUAL | False | False | 0.0 | |
| 16 | INDIVIDUAL | False | False | 0.0 | |
| | | | | | |
| 17 | INDIVIDUAL | False | False | 0.0 | |
| 18 | INDIVIDUAL | False | False | 0.0 | |
| 19 | INDIVIDUAL | False | False | 0.0 | |
| 20 | INDIVIDUAL | False | False | 0.0 | |
| 21 | INDIVIDUAL | False | False | 0.0 | |
| 22 | INDIVIDUAL | False | False | 0.0 | |
| | INDIVIDUAL | False | False | | |
| 23 | | | | 0.0 | |
| 24 | INDIVIDUAL | False | False | 0.0 | |
| 25 | INDIVIDUAL | False | False | 0.0 | |
| 26 | INDIVIDUAL | False | False | 0.0 | |
| 27 | INDIVIDUAL | False | False | 0.0 | |
| 28 | INDIVIDUAL | False | False | 0.0 | |
| 29 | INDIVIDUAL | False | False | 0.0 | |
| | | | | | |
| 40500 | | T-1 | NT - NT | | |
| 42508 | INDIVIDUAL | False | NaN | 0.0 | |
| 42509 | INDIVIDUAL | False | NaN | 0.0 | |
| 42510 | INDIVIDUAL | NaN | NaN | NaN | |
| 42511 | INDIVIDUAL | False | NaN | 0.0 | |
| 42512 | INDIVIDUAL | False | NaN | 0.0 | |
| 42513 | INDIVIDUAL | False | NaN | 0.0 | |
| 42514 | INDIVIDUAL | False | NaN | 0.0 | |
| | | | | | |
| 42515 | INDIVIDUAL | NaN | NaN | NaN | |
| 42516 | INDIVIDUAL | NaN | NaN | NaN | |
| 42517 | INDIVIDUAL | NaN | NaN | NaN | |
| 42518 | INDIVIDUAL | NaN | NaN | NaN | |
| 42519 | INDIVIDUAL | NaN | NaN | NaN | |
| 42520 | INDIVIDUAL | NaN | NaN | NaN | |
| 72020 | TINDTATDOUP | INGIN | Nan | ivalv | |

 ${\tt NaN}$

 ${\tt NaN}$

42536

 ${\tt NaN}$

| 42521 | INDIVIDUAL | NaN | NaN | NaN |
|-------|------------|-----|-----|-----|
| 42522 | INDIVIDUAL | NaN | NaN | NaN |
| 42523 | INDIVIDUAL | NaN | NaN | NaN |
| 42524 | INDIVIDUAL | NaN | NaN | NaN |
| 42525 | INDIVIDUAL | NaN | NaN | NaN |
| 42526 | INDIVIDUAL | NaN | NaN | NaN |
| 42527 | INDIVIDUAL | NaN | NaN | NaN |
| 42528 | INDIVIDUAL | NaN | NaN | NaN |
| 42529 | INDIVIDUAL | NaN | NaN | NaN |
| 42530 | INDIVIDUAL | NaN | NaN | NaN |
| 42531 | INDIVIDUAL | NaN | NaN | NaN |
| 42532 | INDIVIDUAL | NaN | NaN | NaN |
| 42533 | INDIVIDUAL | NaN | NaN | NaN |
| 42534 | INDIVIDUAL | NaN | NaN | NaN |
| 42535 | INDIVIDUAL | NaN | NaN | NaN |
| 42536 | NaN | NaN | NaN | NaN |
| 42537 | NaN | NaN | NaN | NaN |
| | | | | |

| | <pre>pub_rec_bankruptcies</pre> | tax_liens |
|----|---------------------------------|-----------|
| 0 | 0.0 | False |
| 1 | 0.0 | False |
| 2 | 0.0 | False |
| 3 | 0.0 | False |
| 4 | 0.0 | False |
| 5 | 0.0 | False |
| 6 | 0.0 | False |
| 7 | 0.0 | False |
| 8 | 0.0 | False |
| 9 | 0.0 | False |
| 10 | 0.0 | False |
| 11 | 0.0 | False |
| 12 | 0.0 | False |
| 13 | 0.0 | False |
| 14 | 0.0 | False |
| 15 | 0.0 | False |
| 16 | 0.0 | False |
| 17 | 0.0 | False |
| 18 | 0.0 | False |
| 19 | 0.0 | False |
| 20 | 0.0 | False |
| 21 | 0.0 | False |
| 22 | 0.0 | False |
| 23 | 0.0 | False |
| 24 | 0.0 | False |
| 25 | 0.0 | False |
| 26 | 0.0 | False |
| 27 | 0.0 | False |
| 28 | 0.0 | False |

| 29 | 0.0 | False |
|-------|-----|-------|
| | | |
| 42508 | NaN | NaN |
| 42509 | NaN | NaN |
| 42510 | NaN | NaN |
| 42511 | NaN | NaN |
| 42512 | NaN | NaN |
| 42513 | NaN | NaN |
| 42514 | NaN | NaN |
| 42515 | NaN | NaN |
| 42516 | NaN | NaN |
| 42517 | NaN | NaN |
| 42518 | NaN | NaN |
| 42519 | NaN | NaN |
| 42520 | NaN | NaN |
| 42521 | NaN | NaN |
| 42522 | NaN | NaN |
| 42523 | NaN | NaN |
| 42524 | NaN | NaN |
| 42525 | NaN | NaN |
| 42526 | NaN | NaN |
| 42527 | NaN | NaN |
| 42528 | NaN | NaN |
| 42529 | NaN | NaN |
| 42530 | NaN | NaN |
| 42531 | NaN | NaN |
| 42532 | NaN | NaN |
| 42533 | NaN | NaN |
| 42534 | NaN | NaN |
| 42535 | NaN | NaN |
| 42536 | NaN | NaN |
| 42537 | NaN | NaN |
| | | |

[42538 rows x 58 columns]

1.3. Removing irrelevenat observations

| Out[16]: | id | member_id | loan_amnt | funded_amnt | funded_amnt_inv | \ |
|----------|---------|-----------|-----------|-------------|-----------------|---|
| 0 | 1077501 | 1296599.0 | 5000.0 | 5000.0 | 4975.000000 | |
| 1 | 1077430 | 1314167.0 | 2500.0 | 2500.0 | 2500.000000 | |
| 2 | 1077175 | 1313524.0 | 2400.0 | 2400.0 | 2400.000000 | |
| 3 | 1076863 | 1277178.0 | 10000.0 | 10000.0 | 10000.000000 | |
| 4 | 1075358 | 1311748.0 | 3000.0 | 3000.0 | 3000.000000 | |
| 5 | 1075269 | 1311441.0 | 5000.0 | 5000.0 | 5000.000000 | |
| 6 | 1069639 | 1304742.0 | 7000.0 | 7000.0 | 7000.000000 | |

| 7 | 1072053 | 1288686.0 | 3000.0 | 3000.0 | 3000.000000 |
|-------|---------|-----------|---------|---------|--------------|
| 8 | 1071795 | 1306957.0 | 5600.0 | 5600.0 | 5600.000000 |
| 9 | 1071570 | 1306721.0 | 5375.0 | 5375.0 | 5350.000000 |
| 10 | 1070078 | 1305201.0 | 6500.0 | 6500.0 | 6500.000000 |
| 11 | 1069908 | 1305008.0 | 12000.0 | 12000.0 | 12000.000000 |
| 12 | 1064687 | 1298717.0 | 9000.0 | 9000.0 | 9000.000000 |
| 13 | 1069866 | 1304956.0 | 3000.0 | 3000.0 | 3000.000000 |
| 14 | 1069057 | 1303503.0 | 10000.0 | 10000.0 | 10000.000000 |
| 15 | 1069759 | 1304871.0 | 1000.0 | 1000.0 | 1000.000000 |
| 16 | 1065775 | 1299699.0 | 10000.0 | 10000.0 | 10000.000000 |
| 17 | 1069971 | 1304884.0 | 3600.0 | 3600.0 | 3600.000000 |
| 18 | 1062474 | 1294539.0 | 6000.0 | 6000.0 | 6000.000000 |
| 19 | 1069742 | 1304855.0 | 9200.0 | 9200.0 | 9200.000000 |
| 20 | 1069740 | 1284848.0 | 20250.0 | 20250.0 | 19142.161077 |
| 21 | 1039153 | 1269083.0 | 21000.0 | 21000.0 | 21000.000000 |
| 22 | 1069710 | 1304821.0 | 10000.0 | 10000.0 | 10000.000000 |
| 23 | 1069700 | 1304810.0 | 10000.0 | 10000.0 | 10000.000000 |
| 24 | 1069559 | 1304634.0 | 6000.0 | 6000.0 | 6000.000000 |
| 25 | 1069697 | 1273773.0 | 15000.0 | 15000.0 | 15000.000000 |
| 26 | 1069800 | 1304679.0 | 15000.0 | 15000.0 | 8725.000000 |
| 27 | 1069657 | 1304764.0 | 5000.0 | 5000.0 | 5000.000000 |
| 28 | 1069799 | 1304678.0 | 4000.0 | 4000.0 | 4000.000000 |
| 29 | 1047704 | 1278806.0 | 8500.0 | 8500.0 | 8500.000000 |
| | | | | | |
| 42506 | 94406 | 94385.0 | 6725.0 | 6725.0 | 825.000000 |
| 42507 | 93055 | 92947.0 | 2000.0 | 2000.0 | 1025.000000 |
| 42508 | 91175 | 91170.0 | 6000.0 | 6000.0 | 1200.000000 |
| 42509 | 91126 | 91067.0 | 5350.0 | 5350.0 | 625.000000 |
| 42510 | 91023 | 70879.0 | 1900.0 | 1900.0 | 900.000000 |
| 42511 | 90106 | 90090.0 | 10000.0 | 10000.0 | 350.000000 |
| 42512 | 89258 | 80039.0 | 2000.0 | 2000.0 | 1275.000000 |
| 42513 | 88637 | 88629.0 | 6000.0 | 6000.0 | 650.000000 |
| 42514 | 88046 | 88023.0 | 4400.0 | 4400.0 | 1400.000000 |
| 42515 | 85961 | 85923.0 | 1200.0 | 1200.0 | 500.000000 |
| 42516 | 85818 | 85802.0 | 5000.0 | 5000.0 | 375.000000 |
| 42517 | 85781 | 85727.0 | 1400.0 | 1400.0 | 475.000000 |
| 42518 | 85675 | 85667.0 | 1000.0 | 1000.0 | 625.000000 |
| 42519 | 84670 | 79576.0 | 5000.0 | 5000.0 | 300.000000 |
| 42520 | 84098 | 84091.0 | 2500.0 | 2500.0 | 225.000000 |
| 42521 | 83979 | 83974.0 | 3000.0 | 3000.0 | 250.000000 |
| 42522 | 83489 | 83471.0 | 2600.0 | 2600.0 | 575.000000 |
| 42523 | 83185 | 83132.0 | 1000.0 | 1000.0 | 625.000000 |
| 42524 | 76629 | 76623.0 | 1275.0 | 1275.0 | 0.000000 |
| 42525 | 74014 | 73890.0 | 6450.0 | 6450.0 | 0.000000 |
| 42526 | 81085 | 80973.0 | 10500.0 | 10500.0 | 275.000000 |
| 42527 | 77792 | 77764.0 | 3000.0 | 3000.0 | 125.000000 |
| 42528 | 77757 | 70626.0 | 3000.0 | 3000.0 | 0.000000 |
| 42529 | 74505 | 74469.0 | 2000.0 | 2000.0 | 225.000000 |
| | | | | | |

| 42530 | 74323 | 74301.0 | 6500.0 | 6500.0 | | 0.000000 | |
|-------|-----------|----------|-------------|-----------|-------|------------|---|
| 42531 | 73582 | 73096.0 | 3500.0 | 3500.0 | | 225.000000 | |
| 42532 | 72998 | 72992.0 | 1000.0 | 1000.0 | | 0.000000 | |
| 42533 | 72176 | 70868.0 | 2525.0 | 2525.0 | | 225.000000 | |
| 42534 | 71623 | 70735.0 | 6500.0 | 6500.0 | | 0.000000 | |
| 42535 | 70686 | 70681.0 | 5000.0 | 5000.0 | | 0.000000 | |
| 12000 | , 0000 | 1000110 | 3333.3 | 0000.0 | | 0.00000 | |
| | term | int_rate | installment | grade sub | grade | | \ |
| 0 | 36 months | 10.65% | 162.87 | В | B2 | | • |
| 1 | 60 months | 15.27% | 59.83 | С | C4 | | |
| 2 | 36 months | 15.96% | 84.33 | C | C5 | | |
| 3 | 36 months | 13.49% | 339.31 | C | C1 | | |
| 4 | 60 months | 12.69% | 67.79 | В | B5 | | |
| 5 | 36 months | 7.90% | 156.46 | A | A4 | | |
| 6 | 60 months | 15.96% | 170.08 | C | C5 | | |
| 7 | 36 months | 18.64% | 109.43 | Ē | E1 | | |
| 8 | 60 months | 21.28% | 152.39 | F | F2 | | |
| 9 | 60 months | 12.69% | 121.45 | В | B5 | | |
| 10 | 60 months | 14.65% | 153.45 | C | C3 | ••• | |
| 11 | 36 months | 12.69% | 402.54 | В | B5 | • • • | |
| 12 | 36 months | 13.49% | 305.38 | C | C1 | • • • | |
| 13 | 36 months | 9.91% | 96.68 | В | B1 | • • • | |
| 14 | 36 months | 10.65% | 325.74 | В | B2 | • • • | |
| 15 | 36 months | 16.29% | 35.31 | D | D1 | • • • | |
| 16 | 36 months | | 347.98 | C | C4 | • • • | |
| | | 15.27% | | | | • • • | |
| 17 | 36 months | 6.03% | 109.57 | A | A1 | • • • | |
| 18 | 36 months | 11.71% | 198.46 | В | B3 | • • • | |
| 19 | 36 months | 6.03% | 280.01 | A | A1 | • • • | |
| 20 | 60 months | 15.27% | 484.63 | C | C4 | | |
| 21 | 36 months | 12.42% | 701.73 | В | B4 | • • • | |
| 22 | 36 months | 11.71% | 330.76 | В | B3 | • • • | |
| 23 | 36 months | 11.71% | 330.76 | В | B3 | • • • | |
| 24 | 36 months | 11.71% | 198.46 | В - | В3 | • • • | |
| 25 | 36 months | | 483.38 | В | B1 | • • • | |
| 26 | 36 months | | 514.64 | | C2 | | |
| 27 | 60 months | | 123.65 | D - | D2 | | |
| 28 | 36 months | 11.71% | 132.31 | В | В3 | • • • | |
| 29 | 36 months | 11.71% | 281.15 | В | ВЗ | | |
| | • • • | • • • | • • • | | • • • | | |
| 42506 | 36 months | 13.12% | 226.98 | D | D5 | | |
| 42507 | 36 months | 12.80% | 67.20 | D | D4 | | |
| 42508 | 36 months | 13.12% | 202.51 | D | D5 | • • • | |
| 42509 | 36 months | 13.12% | 180.57 | D | D5 | • • • | |
| 42510 | 36 months | 9.64% | 61.00 | В | В4 | • • • | |
| 42511 | 36 months | 14.70% | 345.18 | E | E5 | | |
| 42512 | 36 months | 7.12% | 61.87 | Α | A1 | | |
| 42513 | 36 months | 10.59% | 195.28 | C | C2 | | |
| 42514 | 36 months | 9.64% | 141.25 | В | B4 | | |
| | | | | | | | |

| 42515 | 36 | months | 9.01% | 38.17 | В | B2 | |
|-------|------|----------------|-----------|---------------|---------|------------|----------|
| 42516 | 36 | months | 11.22% | 164.23 | C | C4 | |
| 42517 | 36 | months | 10.91% | 45.78 | C | C3 | |
| 42518 | 36 | months | 14.07% | 34.21 | E | E3 | |
| 42519 | 36 | months | 7.75% | 156.11 | Α | A3 | |
| 42520 | 36 | months | 7.43% | 77.69 | Α | A2 | |
| 42521 | 36 | months | 7.43% | 93.23 | Α | A2 | |
| 42522 | 36 | months | 8.38% | 81.94 | Α | A5 | |
| 42523 | 36 | months | 7.12% | 30.94 | Α | A1 | |
| 42524 | 36 | months | 12.49% | 42.65 | D | D3 | |
| 42525 | 36 | months | 11.22% | 211.85 | C | C4 | |
| 42526 | 36 | months | 11.22% | 344.87 | C | C4 | |
| 42527 | 36 | months | 9.01% | 95.42 | В | B2 | |
| 42528 | 36 | months | 9.33% | 95.86 | В | В3 | |
| 42529 | 36 | ${\tt months}$ | 9.96% | 64.50 | В | B5 | |
| 42530 | 36 | ${\tt months}$ | 9.64% | 208.66 | В | B4 | |
| 42531 | 36 | ${\tt months}$ | 10.28% | 113.39 | C | C1 | |
| 42532 | 36 | months | 9.64% | 32.11 | В | B4 | |
| 42533 | 36 | months | 9.33% | 80.69 | В | В3 | |
| 42534 | 36 | ${\tt months}$ | 8.38% | 204.84 | Α | A5 | |
| 42535 | 36 | months | 7.75% | 156.11 | Α | A3 | |
| | | | | | | | |
| | last | _fico_ra | ange_high | last_fico_ran | nge_low | collection | ons_12_m |
| 0 | | | 744.0 | | 740.0 | | |
| 1 | | | 499.0 | | 0.0 | | |

| | <pre>last_fico_range_high</pre> | <pre>last_fico_range_low</pre> | collections_12_mths_ex_med ' |
|----|---------------------------------|--------------------------------|------------------------------|
| 0 | 744.0 | 740.0 | False |
| 1 | 499.0 | 0.0 | False |
| 2 | 719.0 | 715.0 | False |
| 3 | 604.0 | 600.0 | False |
| 4 | 694.0 | 690.0 | False |
| 5 | 679.0 | 675.0 | False |
| 6 | 654.0 | 650.0 | False |
| 7 | 689.0 | 685.0 | False |
| 8 | 499.0 | 0.0 | False |
| 9 | 519.0 | 515.0 | False |
| 10 | 734.0 | 730.0 | False |
| 11 | 669.0 | 665.0 | False |
| 12 | 619.0 | 615.0 | False |
| 13 | 734.0 | 730.0 | False |
| 14 | 654.0 | 650.0 | False |
| 15 | 769.0 | 765.0 | False |
| 16 | 639.0 | 635.0 | False |
| 17 | 794.0 | 790.0 | False |
| 18 | 559.0 | 555.0 | False |
| 19 | 644.0 | 640.0 | False |
| 20 | 684.0 | 680.0 | False |
| 21 | 534.0 | 530.0 | False |
| 22 | 779.0 | 775.0 | False |
| 23 | 684.0 | 680.0 | False |
| 24 | 499.0 | 0.0 | False |

| 25 | 734.0 | 730.0 | False | |
|-------|---|----------------|--------------------------|---|
| 26 | 509.0 | 505.0 | False | |
| 27 | 579.0 | 575.0 | False | |
| 28 | 689.0 | 685.0 | False | |
| 29 | 599.0 | 595.0 | False | |
| | | | raise | |
| 42506 | 634.0 | 630.0 | NaN | |
| 42507 | 514.0 | 510.0 | NaN | |
| 42508 | 664.0 | 660.0 | NaN | |
| 42509 | 704.0 | 700.0 | NaN | |
| 42510 | 709.0 | 705.0 | NaN | |
| 42511 | 499.0 | 0.0 | NaN | |
| 42512 | 809.0 | 805.0 | NaN | |
| 42513 | 769.0 | 765.0 | NaN | |
| 42514 | 584.0 | 580.0 | NaN | |
| 42515 | 694.0 | 690.0 | NaN | |
| 42516 | 679.0 | 675.0 | NaN | |
| 42517 | 679.0 | 675.0 | NaN | |
| 42518 | 499.0 | 0.0 | NaN | |
| 42519 | 759.0 | 755.0 | NaN NaN | |
| 42520 | 789.0 | 785.0 | NaN NaN | |
| 42521 | 799.0 | 795.0 | NaN | |
| 42521 | 614.0 | 610.0 | NaN | |
| 42523 | 599.0 | 595.0 | NaN | |
| 42523 | 624.0 | | NaN | |
| | | 620.0 | | |
| 42525 | 574.0 | 570.0 | NaN NaN | |
| 42526 | 744.0 | 740.0 | NaN | |
| 42527 | 594.0 | 590.0 | NaN | |
| 42528 | 714.0 | 710.0 | NaN | |
| 42529 | 594.0 | 590.0 | NaN | |
| 42530 | 684.0 | 680.0 | NaN | |
| 42531 | 819.0 | 815.0 | NaN | |
| 42532 | 784.0 | 780.0 | NaN | |
| 42533 | 714.0 | 710.0 | NaN | |
| 42534 | 724.0 | 720.0 | NaN | |
| 42535 | 794.0 | 790.0 | NaN | |
| | <pre>policy_code application_type</pre> | acc now deling | chargeoff within 12 mths | \ |
| 0 | True INDIVIDUAL | _ | False | |
| 1 | True INDIVIDUAL | | False | |
| 2 | True INDIVIDUAL | | False | |
| 3 | True INDIVIDUAL | | False | |
| 4 | True INDIVIDUAL | | False | |
| 5 | True INDIVIDUAL | | False | |
| 6 | True INDIVIDUAL | | False | |
| 7 | True INDIVIDUAL | | False | |
| 8 | True INDIVIDUAL | | False | |
| 9 | True INDIVIDUAL | | False | |
| J | TI GE TWOT A TOOKE | raise | rarse | |

| 10 | True | INDIVIDUAL | False | False |
|-------|------|------------|-------|-------|
| 11 | True | INDIVIDUAL | False | False |
| 12 | True | INDIVIDUAL | False | False |
| 13 | True | INDIVIDUAL | False | False |
| 14 | True | INDIVIDUAL | False | False |
| 15 | True | INDIVIDUAL | False | False |
| 16 | True | INDIVIDUAL | False | False |
| 17 | True | INDIVIDUAL | False | False |
| 18 | True | INDIVIDUAL | False | False |
| 19 | True | INDIVIDUAL | False | False |
| 20 | True | INDIVIDUAL | False | False |
| 21 | True | INDIVIDUAL | False | False |
| 22 | True | INDIVIDUAL | False | False |
| 23 | True | INDIVIDUAL | False | False |
| 24 | True | INDIVIDUAL | False | False |
| 25 | True | INDIVIDUAL | False | False |
| 26 | True | INDIVIDUAL | False | False |
| 27 | True | INDIVIDUAL | False | False |
| 28 | True | INDIVIDUAL | False | False |
| 29 | True | INDIVIDUAL | False | False |
| | | | | |
| 42506 | True | INDIVIDUAL | False | NaN |
| 42507 | True | INDIVIDUAL | False | NaN |
| 42508 | True | INDIVIDUAL | False | NaN |
| 42509 | True | INDIVIDUAL | False | NaN |
| 42510 | True | INDIVIDUAL | NaN | NaN |
| 42511 | True | INDIVIDUAL | False | NaN |
| 42512 | True | INDIVIDUAL | False | NaN |
| 42513 | True | INDIVIDUAL | False | NaN |
| 42514 | True | INDIVIDUAL | False | NaN |
| 42515 | True | INDIVIDUAL | NaN | NaN |
| 42516 | True | INDIVIDUAL | NaN | NaN |
| 42517 | True | INDIVIDUAL | NaN | NaN |
| 42518 | True | INDIVIDUAL | NaN | NaN |
| 42519 | True | INDIVIDUAL | NaN | NaN |
| 42520 | True | INDIVIDUAL | NaN | NaN |
| 42521 | True | INDIVIDUAL | NaN | NaN |
| 42522 | True | INDIVIDUAL | NaN | NaN |
| 42523 | True | INDIVIDUAL | NaN | NaN |
| 42524 | True | INDIVIDUAL | NaN | NaN |
| 42525 | True | INDIVIDUAL | NaN | NaN |
| 42526 | True | INDIVIDUAL | NaN | NaN |
| 42527 | True | INDIVIDUAL | NaN | NaN |
| 42528 | True | INDIVIDUAL | NaN | NaN |
| 42529 | True | INDIVIDUAL | NaN | NaN |
| 42530 | True | INDIVIDUAL | NaN | NaN |
| 42531 | True | INDIVIDUAL | NaN | NaN |
| 42532 | True | INDIVIDUAL | NaN | NaN |
| | | | | |

| 42533 | True | e INDIVIDUAL | NaN |
|-------|-------------|---------------------------------|-----------|
| 42534 | True | e INDIVIDUAL | NaN |
| 42535 | True | e INDIVIDUAL | NaN |
| | | | |
| | delinq_amnt | <pre>pub_rec_bankruptcies</pre> | tax_liens |
| 0 | 0.0 | 0.0 | False |
| 1 | 0.0 | 0.0 | False |
| 2 | 0.0 | 0.0 | False |
| 3 | 0.0 | 0.0 | False |
| 4 | 0.0 | 0.0 | False |
| 5 | 0.0 | 0.0 | False |
| 6 | 0.0 | 0.0 | False |
| 7 | 0.0 | 0.0 | False |
| 8 | 0.0 | 0.0 | False |
| 9 | 0.0 | 0.0 | False |
| 10 | 0.0 | 0.0 | False |
| 11 | 0.0 | 0.0 | False |
| 12 | 0.0 | 0.0 | False |
| 13 | 0.0 | 0.0 | False |
| 14 | 0.0 | 0.0 | False |
| 15 | 0.0 | 0.0 | False |
| 16 | 0.0 | 0.0 | False |
| 17 | 0.0 | 0.0 | False |
| 18 | 0.0 | 0.0 | False |
| 19 | 0.0 | 0.0 | False |
| 20 | 0.0 | 0.0 | False |
| 21 | 0.0 | 0.0 | False |
| 22 | 0.0 | 0.0 | False |
| 23 | 0.0 | 0.0 | False |
| 24 | 0.0 | 0.0 | False |
| 25 | 0.0 | 0.0 | False |
| 26 | 0.0 | 0.0 | False |
| 27 | 0.0 | 0.0 | False |
| 28 | 0.0 | 0.0 | False |
| 29 | 0.0 | 0.0 | False |
| | | | |
| 42506 | 0.0 | NaN | NaN |
| 42507 | 0.0 | NaN | NaN |
| 42508 | 0.0 | NaN | NaN |
| 42509 | 0.0 | NaN | NaN |
| 42510 | NaN | NaN | NaN |
| 42511 | 0.0 | NaN | NaN |
| 42512 | 0.0 | NaN | NaN |
| 42513 | 0.0 | NaN | NaN |
| 42514 | 0.0 | NaN | NaN |
| 42515 | NaN | NaN | NaN |
| 42516 | NaN | NaN | NaN |
| 42517 | NaN | NaN | NaN |
| | | | |

NaN NaN NaN

| 42518 | NaN | NaN | NaN |
|-------|-----|-----|-----|
| 42519 | NaN | NaN | NaN |
| 42520 | NaN | NaN | NaN |
| 42521 | NaN | NaN | NaN |
| 42522 | NaN | NaN | NaN |
| 42523 | NaN | NaN | NaN |
| 42524 | NaN | NaN | NaN |
| 42525 | NaN | NaN | NaN |
| 42526 | NaN | NaN | NaN |
| 42527 | NaN | NaN | NaN |
| 42528 | NaN | NaN | NaN |
| 42529 | NaN | NaN | NaN |
| 42530 | NaN | NaN | NaN |
| 42531 | NaN | NaN | NaN |
| 42532 | NaN | NaN | NaN |
| 42533 | NaN | NaN | NaN |
| 42534 | NaN | NaN | NaN |
| 42535 | NaN | NaN | NaN |
| | | | |

[42536 rows x 58 columns]

- 2. Data Transformation Feature scaling is the method to limit the range of variables so that they can be compared on common grounds. It is performed on continuous variables. Lets plot the distribution of all the continuous variables in the data set.
- 2.1 Min-Max Normalization or Rescaling xnorm = (X-Xmin)/ (Xmax-Xmin)

```
In [17]: import matplotlib.pyplot as plt
        %matplotlib inline
         # Create data samples x1, x2, x3
        df = pd.DataFrame({
             # positive skew
             'x1': np.random.chisquare(8, 1000),
             # negative skew
             'x2': np.random.beta(8, 2, 1000) * 40,
             # no skew
             'x3': np.random.normal(50, 3, 1000)
        })
        df
Out[17]:
                     x1
                                x2
                                           хЗ
               4.176608 36.578501
                                   53.278469
        0
         1
             11.762533 30.574695 53.018049
         2
               6.502315 36.777874 48.015045
        3
               4.373445 35.828148 50.063159
        4
               4.232660 29.378933 50.233194
        5
               2.130249 35.112108 53.277256
        6
               1.767400 33.833223 49.938346
```

```
7
                             53.343309
      6.356473
                 39.116542
8
     10.524769
                 33.924826
                             51.555350
                 34.847636
9
      7.423300
                             53.520430
      5.700439
                 37.542186
                             51.893499
10
11
      9.870685
                 31.214974
                             50.897999
12
      6.140859
                 30.056021
                             45.247602
13
      3.396192
                 25.785211
                             54.832185
14
      3.349601
                 37.372898
                             51.721646
15
     10.962452
                 28.185094
                             46.317004
16
     10.728609
                 37.629154
                             49.989800
17
      7.254143
                 34.248109
                             53.432171
18
     10.079955
                 32.692325
                             55.423879
19
      7.045191
                 32.750820
                             50.372525
20
      6.398252
                 28.728606
                             47.396270
21
      9.553095
                 28.322186
                             46.954989
22
      5.250622
                 25.688546
                             49.292995
23
      5.567015
                 24.767568
                             50.293283
24
      4.798372
                 34.037735
                             45.470087
25
      2.619429
                 29.994384
                             51.155127
26
      8.811075
                 29.652466
                             50.123312
27
      5.274134
                 29.566529
                             54.941685
                             52.504430
28
      5.422298
                 29.937114
29
     12.474314
                 29.378271
                             52.811595
. .
            . . .
                        . . .
                                    . . .
     20.306841
970
                 33.390367
                             49.175230
971
                             46.325561
      3.156177
                 35.735982
972
     16.536062
                 29.070146
                             52.057782
973
      7.091280
                 32.257698
                             53.663401
974
      1.421279
                 29.969979
                             49.569683
975
     10.389134
                 28.556535
                             47.611947
976
     15.453047
                 35.135510
                             49.778348
977
      2.423340
                 32.950679
                             49.373322
978
      3.701311
                 34.233299
                             51.975502
979
     10.293042
                 26.695088
                             50.696987
980
                 33.822910
                             46.826542
      6.464377
981
      3.656951
                 27.433069
                             51.979545
982
      4.213380
                 31.769439
                             51.703752
983
      8.758165
                 28.098044
                             47.025437
984
      7.005155
                 39.730607
                             48.436875
985
      7.746603
                 38.559149
                             52.544724
986
      7.197039
                 29.312129
                             44.713459
987
     15.169891
                 35.443128
                             44.165024
988
                 37.305274
      4.259811
                             49.496617
989
      6.753583
                 35.842331
                             50.951803
990
     17.070655
                 32.385656
                             53.691918
991
     10.276571
                 30.897110
                             55.165659
992
      6.964506
                 30.248598
                             49.785705
993
                 21.451023
      6.377272
                             51.444754
```

```
994 4.237889 33.841525 54.888862

995 9.352061 24.807878 46.197119

996 8.104146 37.909359 46.066406

997 9.820446 32.563369 53.232693

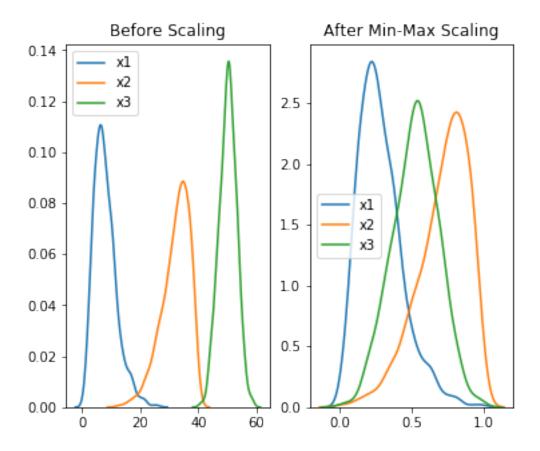
998 6.022620 32.667803 52.162870

999 12.266281 27.496064 46.881607

[1000 rows x 3 columns]
```

We infer that each one of them are in different range For data with attributes of varying scales, we can rescale attributes to possess the same scale. We rescale attributes into the range 0 to 1 and call it normalization. We use the MinMaxScaler class from scikit-learn.

```
In [18]: from sklearn.preprocessing import MinMaxScaler
         import seaborn as sns
         # Use MinMaxScaler
         scaler = MinMaxScaler()
         scaled_df = scaler.fit_transform(df)
         scaled_df = pd.DataFrame(scaled_df, columns=['x1', 'x2', 'x3'])
         # Plot and visualize
         fig, (ax1, ax2) = plt.subplots(ncols=2, figsize=(6, 5))
         ax1.set_title('Before Scaling')
         sns.kdeplot(df['x1'], ax=ax1)
         sns.kdeplot(df['x2'], ax=ax1)
         sns.kdeplot(df['x3'], ax=ax1)
         ax2.set_title('After Min-Max Scaling')
         sns.kdeplot(scaled_df['x1'], ax=ax2)
         sns.kdeplot(scaled_df['x2'], ax=ax2)
         sns.kdeplot(scaled_df['x3'], ax=ax2)
         plt.show()
```



2.2 Z-score Normalization or standardizing or Mean Removal Standardize features by removing the mean and scaling to unit variance

The standard score of a sample x is calculated as:

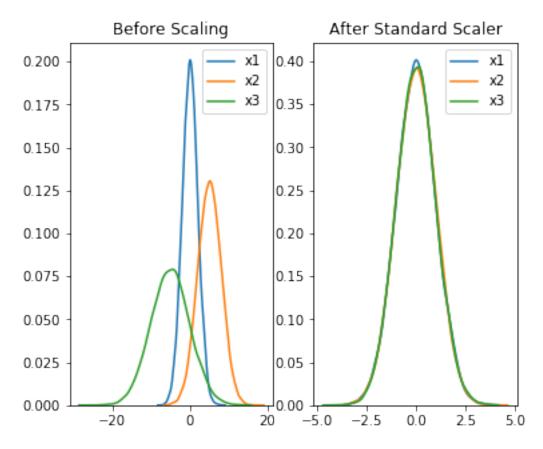
z = (x - u) / s where u is the mean of the training samples or zero if with_mean=False, and s is the standard deviation of the 4 training samples or one if with_std=False.

Centering and scaling happen independently on each feature by computing the relevant statistics on the samples in the training set. Mean and standard deviation are then stored to be used on later data using the transform method.

```
In [19]: from sklearn.preprocessing import StandardScaler
    # Create data samples x1, x2, x3
    np.random.seed(1)
    df = pd.DataFrame({
        'x1': np.random.normal(0, 2, 10000),
        'x2': np.random.normal(5, 3, 10000),
        'x3': np.random.normal(-5, 5, 10000)
})

# Use StandardScaler
scaler = StandardScaler()
scaled_df = scaler.fit_transform(df)
scaled_df = pd.DataFrame(scaled_df, columns=['x1', 'x2', 'x3'])
```

```
# Plot and visualize
fig, (ax1, ax2) = plt.subplots(ncols=2, figsize=(6, 5))
ax1.set_title('Before Scaling')
sns.kdeplot(df['x1'], ax=ax1)
sns.kdeplot(df['x2'], ax=ax1)
sns.kdeplot(df['x3'], ax=ax1)
ax2.set_title('After Standard Scaler')
sns.kdeplot(scaled_df['x1'], ax=ax2)
sns.kdeplot(scaled_df['x2'], ax=ax2)
sns.kdeplot(scaled_df['x3'], ax=ax2)
sns.kdeplot(scaled_df['x3'], ax=ax2)
```



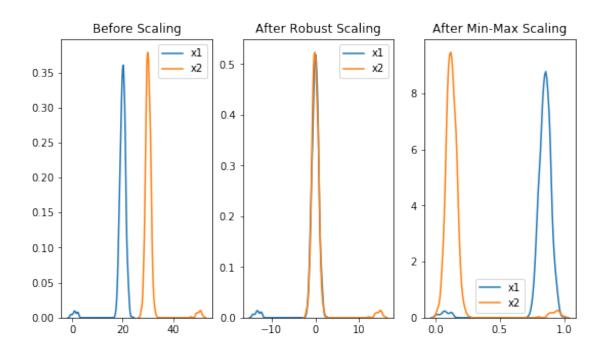
2.3 Robust Scaler Scale features using statistics that are robust to outliers.

This Scaler removes the median and scales the data according to the quantile range (defaults to IQR: Interquartile Range). The IQR is the range between the 1st quartile (25th quantile) and the 3rd quartile (75th quantile).

The RobustScaler uses a similar method to the Min-Max scaler. However, it uses the interquartile range instead of the min-max, which makes it robust to outliers. It follows the following formula for each feature:

(Xi-Q1)/(Q3-Q1)

```
In [20]: from sklearn.preprocessing import RobustScaler
         # Create data samples x1, x2
         x = pd.DataFrame({
             # Distribution with lower outliers
             'x1': np.concatenate([np.random.normal(20, 1, 1000), np.random.normal(1, 1, 25)])
             # Distribution with higher outliers
             'x2': np.concatenate([np.random.normal(30, 1, 1000), np.random.normal(50, 1, 25)]
         })
         # Use RobustScaler
         scaler = RobustScaler()
         robust_scaled_df = scaler.fit_transform(x)
         robust_scaled_df = pd.DataFrame(robust_scaled_df, columns=['x1', 'x2'])
         # Use MinMaxScaler
         scaler = MinMaxScaler()
         minmax_scaled_df = scaler.fit_transform(x)
         minmax_scaled_df = pd.DataFrame(minmax_scaled_df, columns=['x1', 'x2'])
         # Plot and visualize
         fig, (ax1, ax2, ax3) = plt.subplots(ncols=3, figsize=(9, 5))
         ax1.set_title('Before Scaling')
         sns.kdeplot(x['x1'], ax=ax1)
         sns.kdeplot(x['x2'], ax=ax1)
         ax2.set_title('After Robust Scaling')
         sns.kdeplot(robust_scaled_df['x1'], ax=ax2)
         sns.kdeplot(robust_scaled_df['x2'], ax=ax2)
         ax3.set_title('After Min-Max Scaling')
         sns.kdeplot(minmax_scaled_df['x1'], ax=ax3)
         sns.kdeplot(minmax_scaled_df['x2'], ax=ax3)
         plt.show()
```



3. Converting Numeric features to Binary feataures

```
In [26]: from sklearn.preprocessing import Binarizer
         \# Create data samples x1, x2
         x = pd.DataFrame({
             # Distribution with lower outliers
             'x1': np.concatenate([np.random.normal(20, 1, 1000), np.random.normal(1, 1, 25)])
             # Distribution with higher outliers
             'x2': np.concatenate([np.random.normal(30, 1, 1000), np.random.normal(50, 1, 25)]
         })
         # Use Binarizer
         scaler = Binarizer(threshold=5)
         binarized_df = scaler.transform(x)
         binarized_df = pd.DataFrame(binarized_df, columns=['x1', 'x2'])
         binarized_df
Out [26]:
                x1
                     x2
         0
               1.0 1.0
         1
               1.0 1.0
         2
               1.0 1.0
         3
               1.0 1.0
         4
               1.0 1.0
         5
               1.0 1.0
         6
               1.0 1.0
         7
               1.0 1.0
               1.0 1.0
         8
```

```
9
      1.0 1.0
10
      1.0
            1.0
      1.0
            1.0
11
12
      1.0
            1.0
      1.0
            1.0
13
14
      1.0
            1.0
      1.0
15
            1.0
      1.0
            1.0
16
17
      1.0
            1.0
18
      1.0
            1.0
19
      1.0
            1.0
20
      1.0
            1.0
21
      1.0
            1.0
22
      1.0
            1.0
23
      1.0
            1.0
      1.0
24
            1.0
25
      1.0
            1.0
26
      1.0
            1.0
27
      1.0
            1.0
28
      1.0
            1.0
29
      1.0
            1.0
. . .
       . . .
            . . .
      1.0
995
            1.0
996
      1.0
            1.0
997
      1.0
            1.0
998
      1.0
            1.0
999
      1.0
            1.0
1000
      0.0
            1.0
1001
      0.0
            1.0
1002
      0.0
            1.0
1003
      0.0
            1.0
1004
      0.0
            1.0
1005
      0.0
            1.0
1006
      0.0
            1.0
1007
      0.0
            1.0
1008
      0.0
            1.0
1009
      0.0
            1.0
1010
      0.0
            1.0
1011
      0.0
            1.0
1012
      0.0
            1.0
1013
      0.0
            1.0
1014
      0.0
            1.0
1015
      0.0
            1.0
1016
      0.0
            1.0
1017
      0.0
            1.0
1018
      0.0
            1.0
1019
      0.0
            1.0
1020
      0.0
            1.0
```

```
1023 0.0 1.0
         1024 0.0 1.0
         [1025 rows x 2 columns]
  4. Handling Ordinal Categorical Variables
In [60]: df_cat = pd.DataFrame(data =
                              [['green','M',10.1,'class1'],
                               ['blue','L',20.1,'class2'],
                               ['white','M',30.1,'class1']])
         df_cat.columns = ['color', 'size', 'price', 'classlabel']
         df_cat
Out [60]:
            color size price classlabel
                     Μ
                         10.1
                                  class1
         0 green
             blue
                     L
                         20.1
                                  class2
         2 white
                         30.1
                                  class1
In [62]: #Using Label Encoder
         from sklearn.preprocessing import LabelEncoder
         label_encode = LabelEncoder()
         label_encode.fit_transform(['S','L','M'])
         df_cat['size'] =label_encode.transform(df_cat['size'].values)
         df_cat
Out [62]:
            color size price classlabel
         0 green
                          10.1
                                   class1
            blue
                      0
                          20.1
                                   class2
         2 white
                          30.1
                                   class1
                      1
In [63]: #Using Label Encoder
         from sklearn.preprocessing import LabelEncoder
         label_encode = LabelEncoder()
         df_cat['classlabel'] =label_encode.fit_transform(df_cat['classlabel'].values)
         df_cat
Out [63]:
            color size price classlabel
         0 green
                      1
                         10.1
                                         0
           blue
                      0
                          20.1
                                         1
         2 white
                      1
                          30.1
                                         0
  5. Handling Nomial Categorical Variables
In [74]: #Using OneHotEncoder
         from sklearn.preprocessing import OneHotEncoder
         enc = LabelEncoder()
```

1021 0.0 1.0 1022 0.0 1.0

```
new_cat_features = enc.fit_transform(df_cat['color'].values)
        print( new_cat_features) # [1 0 2]
        new_cat_features = new_cat_features.reshape(-1, 1) # Needs to be the correct shape
         ohe = OneHotEncoder(sparse=False) #Easier to read
         print(ohe.fit_transform(new_cat_features))
[1 0 2]
[[0. 1. 0.]
 [1. 0. 0.]
 [0. 0. 1.]]
In [76]: # Using get dummies
        df_cat1 = pd.get_dummies(df_cat[['color','size','price','classlabel']])
Out [76]:
           size price classlabel color_blue color_green color_white
               1
                  10.1
                                 0
                                             0
         1
               0
                  20.1
                                 1
                                             1
                                                          0
                                                                       0
         2
               1
                  30.1
                                 0
                                             0
                                                                       1
                                                          0
```