```
In [1]:
          import pandas as pd
          import numpy as np
          import matplotlib as plt
          %matplotlib inline
          df = pd.read_csv("C:\\Users\\Suresh Jamadagni\\downloads\\train.csv")
In [2]:
          print(df)
               Loan_ID Gender Married Dependents
                                                         Education Self_Employed
         0
              LP001002
                           Male
                                      No
                                                          Graduate
         1
              LP001003
                           Male
                                     Yes
                                                   1
                                                          Graduate
                                                                                No
         2
              LP001005
                           Male
                                     Yes
                                                  0
                                                          Graduate
                                                                               Yes
         3
              LP001006
                           Male
                                     Yes
                                                  0 Not Graduate
                                                                                No
                                                  0
         4
              LP001008
                           Male
                                     No
                                                          Graduate
                                                                                No
                            . . .
                                     . . .
                   . . .
                                                 . . .
                                                                               . . .
         . .
              LP002978
                                                          Graduate
         609
                         Female
                                     No
                                                  0
                                                                                No
                                     Yes
         610
              LP002979
                           Male
                                                 3+
                                                          Graduate
                                                                                No
         611
              LP002983
                           Male
                                     Yes
                                                  1
                                                          Graduate
                                                                                No
         612
              LP002984
                           Male
                                     Yes
                                                   2
                                                          Graduate
                                                                                No
             LP002990 Female
                                                          Graduate
         613
                                      No
                                                  a
                                                                               Yes
              ApplicantIncome
                               CoapplicantIncome LoanAmount Loan_Amount_Term
         0
                          5849
                                               0.0
                                                            NaN
                                                                              360.0
         1
                          4583
                                            1508.0
                                                          128.0
                                                                              360.0
         2
                          3000
                                               0.0
                                                           66.0
                                                                              360.0
         3
                          2583
                                            2358.0
                                                          120.0
                                                                              360.0
         4
                          6000
                                                          141.0
                                               0.0
                                                                              360.0
                                                . . .
                                                            . . .
         609
                          2900
                                               0.0
                                                           71.0
                                                                              360.0
                                                           40.0
         610
                          4106
                                               0.0
                                                                              180.0
         611
                          8072
                                             240.0
                                                          253.0
                                                                              360.0
         612
                          7583
                                               0.0
                                                          187.0
                                                                              360.0
         613
                          4583
                                               0.0
                                                          133.0
                                                                              360.0
              Credit_History Property_Area Loan_Status
                                       Urban
         0
                          1.0
         1
                          1.0
                                       Rural
                                                        N
         2
                          1.0
                                       Urban
                                                        Υ
         3
                          1.0
                                       Urban
         4
                                       Urban
                                                        Υ
                          1.0
                          . . .
                                         . . .
         609
                          1.0
                                       Rural
                                                        Υ
         610
                          1.0
                                       Rural
                                                        Υ
         611
                          1.0
                                       Urban
                                       Urban
         612
                          1.0
                                                        Υ
         613
                          0.0
                                  Semiurban
         [614 rows x 13 columns]
In [3]:
          df.head()
                     #displays first 5 rows
                     Gender Married
                                                            Self_Employed ApplicantIncome
Out[3]:
             Loan_ID
                                      Dependents
                                                  Education
                                                                                           Coapplican
```

Graduate

Graduate

Graduate

1

No

No

Yes

5849

4583

3000

0 LP001002

2 LP001005

LP001003

Male

Male

Male

No

Yes

Yes

df.head()	3	B LF	P001006	Male	Yes	0	Not Graduate		No	2583	
to_drop = ['Dependents'] #df.drop(to_drop) print(df) Loan_ID Gender Married Dependents	4	1 LF	P001008	Male	No	0	Graduate		No	6000	
to_drop = ['Dependents'] #df.drop(to_drop) print(df) Loan_ID Gender Married Dependents											
#df.drop(to_drop) print(df) Loan_ID Gender Married Dependents 6 LP001002 Male No 0 Graduate No 1 LP001003 Male Yes 1 Graduate No 2 LP001005 Male Yes 0 Graduate Yes 3 LP001006 Male Yes 0 Not Graduate No 4 LP001008 Male No 0 Graduate No 4 LP001008 Male Yes 0 Not Graduate No 5 LP001006 Male Yes 0 Not Graduate No 6 LP002978 Female No 0 Graduate No 6 Gra											
### Polication		#df	.drop(to_		s']						
LP001003 Male Yes 1 Graduate No			_					cation Se	elf_Employed	\	
LP001005	0)	LP001002		No)	0 Gra	aduate	No		
LP001006 Male Yes 0 Not Graduate No			LP001003		Yes	;	1 Gra	aduate	No		
### LP001008 Male No 0 Graduate No			LP001005						Yes		
							0 Not Gra	aduate			
Company Comp	4	ŀ	LP001008	Male	No)	0 Gra	aduate	No		
610 LP002979 Male Yes 3+ Graduate No 611 LP002983 Male Yes 1 Graduate No 612 LP002984 Male Yes 2 Graduate No 613 LP002990 Female No 0 Graduate Yes ApplicantIncome CoapplicantIncome LoanAmount Loan_Amount_Term 0 S849 0.0 NaN 360.0 1 4583 1508.0 128.0 360.0 2 3000 0.0 66.0 360.0 3 2583 2358.0 120.0 360.0 4 6000 0.0 141.0 360.0	•	•					•	• • •	• • •		
611 LP002983 Male Yes 1 Graduate No 612 LP002984 Male Yes 2 Graduate No 613 LP002990 Female No 0 Graduate Yes ApplicantIncome CoapplicantIncome LoanAmount Loan_Amount_Term 0 5849 0.0 NaN 360.0 1 4583 1508.0 128.0 360.0 2 3000 0.0 66.0 360.0 3 2583 2358.0 120.0 360.0 4 6000 0.0 141.0 360.0 609 2900 0.0 71.0 360.0 610 4106 0.0 40.0 180.0 611 8072 240.0 253.0 360.0 612 7583 0.0 187.0 360.0 613 4583 0.0 133.0 360.0 613 4583 0.0 133.0 360.0 Credit_History Property_Area Loan_Status 0 1.0 Urban Y 1 1.0 Rural N 2 1.0 Urban Y 4 1.0 Urban Y 4 1.0 Urban Y 5 1.0 Urban Y 6 1.0 Urban Y 7 1.0 Urban Y 8 1.0 Urban Y 9 1.0 Urban Y 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	6	609			No)	0 Gra	aduate	No		
ApplicantIncome	6	10			Yes	3	+ Gra	aduate	No		
ApplicantIncome CoapplicantIncome LoanAmount Loan_Amount_Term \ 0	6	511	LP002983		Yes	;	1 Gra	aduate	No		
ApplicantIncome CoapplicantIncome LoanAmount Loan_Amount_Term \ 0	6	512	LP002984	Male	Yes	;	2 Gra	aduate	No		
0 5849 0.0 NaN 360.0 1 4583 1508.0 128.0 360.0 2 3000 0.0 66.0 360.0 3 2583 2358.0 120.0 360.0 4 6000 0.0 141.0 360.0 609 2900 0.0 71.0 360.0 610 4106 0.0 40.0 180.0 611 8072 240.0 253.0 360.0 612 7583 0.0 187.0 360.0 613 4583 0.0 133.0 360.0 Credit_History Property_Area Loan_Status 0 1.0 Urban Y 1 1.0 Rural N 2 1.0 Urban Y 4 1.0 Urban Y 4 1.0 Urban Y 609 1.0 Rural Y 610 1.0 Rural Y 611 1.0 Urban Y 611 1.0 Urban Y 612 1.0 Urban Y 613 0.0 Semiurban N [614 rows x 13 columns]	6	513	LP002990	Female	No)	0 Gra	aduate	Yes		
0 5849 0.0 NaN 360.0 1 4583 1508.0 128.0 360.0 2 3000 0.0 66.0 360.0 3 2583 2358.0 120.0 360.0 4 6000 0.0 141.0 360.0 609 2900 0.0 71.0 360.0 610 4106 0.0 40.0 180.0 611 8072 240.0 253.0 360.0 612 7583 0.0 187.0 360.0 613 4583 0.0 133.0 360.0 Credit_History Property_Area Loan_Status 0 1.0 Urban Y 1 1.0 Rural N 2 1.0 Urban Y 4 1.0 Urban Y 4 1.0 Urban Y 609 1.0 Rural Y 610 1.0 Rural Y 611 1.0 Urban Y 611 1.0 Urban Y 612 1.0 Urban Y 613 0.0 Semiurban N [614 rows x 13 columns]			Annlican	†Tncome	Coannli	cantIncome	LoanAmoi	ınt Loar	n Amount Term	\	
1 4583 1508.0 128.0 360.0 2 3000 0.0 66.0 360.0 3 2583 2358.0 120.0 360.0 4 6000 0.0 141.0 360.0 609 2900 0.0 71.0 360.0 610 4106 0.0 40.0 180.0 611 8072 240.0 253.0 360.0 612 7583 0.0 187.0 360.0 613 4583 0.0 133.0 360.0 Credit_History Property_Area Loan_Status 0 1.0 Urban Y 1 1.0 Rural N 2 1.0 Urban Y 4 1.0 Urban Y 4 1.0 Urban Y 610 1.0 Rural Y 610 1.0 Rural Y 611 1.0 Urban Y 612 1.0 Urban Y 613 0.0 Semiurban N [614 rows x 13 columns]	a)	пррітсин		соирріі					\	
2 3000 0.0 66.0 360.0 3 2583 2358.0 120.0 360.0 4 6000 0.0 141.0 360.0 609 2900 0.0 71.0 360.0 610 4106 0.0 40.0 180.0 611 8072 240.0 253.0 360.0 612 7583 0.0 187.0 360.0 613 4583 0.0 187.0 360.0 613 4583 0.0 133.0 360.0 Credit_History Property_Area Loan_Status 0 1.0 Urban Y 1 1.0 Rural N 2 1.0 Urban Y 3 1.0 Urban Y 4 1.0 Urban Y 609 1.0 Rural Y 610 1.0 Rural Y 611 1.0 Urban Y 612 1.0 Urban Y 613 0.0 Semiurban N [614 rows x 13 columns]											
3											
4 6000 0.0 141.0 360.0											
609											
610											
611 8072 240.0 253.0 360.0 612 7583 0.0 187.0 360.0 613 4583 0.0 133.0 360.0 Credit_History Property_Area Loan_Status 0 1.0 Urban Y 1 1.0 Rural N 2 1.0 Urban Y 4 1.0 Urban Y											
612											
Credit_History Property_Area Loan_Status 0											
0 1.0 Urban Y 1 1.0 Rural N 2 1.0 Urban Y 3 1.0 Urban Y 4 1.0 Urban Y 609 1.0 Rural Y 610 1.0 Rural Y 611 1.0 Urban Y 612 1.0 Urban Y 613 0.0 Semiurban N [614 rows x 13 columns]											
0 1.0 Urban Y 1 1.0 Rural N 2 1.0 Urban Y 3 1.0 Urban Y 4 1.0 Urban Y 609 1.0 Rural Y 610 1.0 Rural Y 611 1.0 Urban Y 612 1.0 Urban Y 613 0.0 Semiurban N [614 rows x 13 columns]			6 1:1 11				.				
1	_		Credit_H	-	. , _						
2 1.0 Urban Y 3 1.0 Urban Y 4 1.0 Urban Y 609 1.0 Rural Y 610 1.0 Rural Y 611 1.0 Urban Y 612 1.0 Urban Y 613 0.0 Semiurban N [614 rows x 13 columns]											
1.0 Urban Y 1.0 Rural Y 1.0 Urban N											
4 1.0 Urban Y 609 1.0 Rural Y 610 1.0 Rural Y 611 1.0 Urban Y 612 1.0 Urban Y 613 0.0 Semiurban N [614 rows x 13 columns]											
<pre></pre>											
609					Ĺ						
610					п						
611											
612 1.0 Urban Y 613 0.0 Semiurban N [614 rows x 13 columns] df.tail() #displays last 5 rows											
613 0.0 Semiurban N [614 rows x 13 columns] df.tail() #displays last 5 rows											
[614 rows x 13 columns] df.tail() #displays last 5 rows											
			rows x 1								
unitali() #utsptays tust 5 rows											
Loan_ID Gender Married Dependents Education Self_Employed ApplicantIncome Coap		df.	tail()	#display	s last	5 rows					
			Loan_ID	Gender N	/larried	Dependents	Education	Self_Emp	oloyed Applica	ntlncome	Coappli

609 LP002978 Female

Male

610 LP002979

No

Yes

0

3+

Graduate

Graduate

No

No

2900

4106

Loan_ID Gender Married Dependents Education Self_Employed ApplicantIncome Coapplican

	Loan_ID	Gender	Married	Dependents	Education	Self_Employed	ApplicantIncome	Coapplic
611	LP002983	Male	Yes	1	Graduate	No	8072	
612	LP002984	Male	Yes	2	Graduate	No	7583	
613	LP002990	Female	No	0	Graduate	Yes	4583	

In [5]: print(df.tail(10))

pri	nt(df.tail	(10))					
	Loan_ID		Married	Dependents	Educati	on Self_Employed	\
604	LP002959	Female	Yes	1	Gradua)
605	LP002960	Male	Yes	0	Not Gradua		
606	LP002961	Male	Yes	1	Gradua	te No)
607	LP002964	Male	Yes	2	Not Gradua	te No)
608	LP002974	Male	Yes	0	Gradua)
609	LP002978	Female	No	0	Gradua		
610	LP002979	Male	Yes	3+	Gradua	te No)
611	LP002983	Male	Yes	1	Gradua)
612	LP002984	Male	Yes	2	Gradua)
613	LP002990	Female	No	0	Gradua	te Yes	
	Applicant	Income	Coapplic	antIncome	LoanAmount	Loan_Amount_Ter	
604		12000		0.0	496.0	360.	
605		2400		3800.0	NaN	180.	
606		3400		2500.0	173.0	360.	0
607		3987		1411.0	157.0	360.	0
608		3232		1950.0	108.0	360.	
609		2900		0.0	71.0	360.	
610		4106		0.0	40.0	180.	
611		8072		240.0	253.0	360.	
612		7583		0.0	187.0	360.	
613		4583		0.0	133.0	360.	0
	Credit_Hi	-		rea Loan_St			
604		1.0	Semiur		Υ		
605		1.0		ban	N		
606		1.0	Semiur	ban	Υ		
607		1.0	Ru	ral	Υ		
608		1.0	Ru	ral	Υ		
609		1.0	Ru	ral	Υ		
610		1.0		ral	Υ		
611		1.0	Ur	ban	Υ		
612		1.0		ban	Υ		
613		0.0	Semiur	1	N		

In [6]: df.describe() #Provides summary measures for variables of type : int or float

Out[6]:		ApplicantIncome	CoapplicantIncome	LoanAmount	Loan_Amount_Term	Credit_History
	count	614.000000	614.000000	592.000000	600.00000	564.000000
	mean	5403.459283	1621.245798	146.412162	342.00000	0.842199
	std	6109.041673	2926.248369	85.587325	65.12041	0.364878
	min	150.000000	0.000000	9.000000	12.00000	0.000000
	25%	2877.500000	0.000000	100.000000	360.00000	1.000000
	50%	3812.500000	1188.500000	128.000000	360.00000	1.000000

```
75%
                     5795.000000
                                      2297.250000
                                                    168.000000
                                                                       360.00000
                                                                                      1.000000
           max
                    81000.000000
                                     41667.000000
                                                    700.000000
                                                                       480.00000
                                                                                      1.000000
In [11]:
          print(df['Property_Area'].value_counts())
                                                           #for categorial data(works like group_
          Semiurban
                       233
          Urban
                       202
          Rural
                       179
          Name: Property_Area, dtype: int64
In [12]:
          print(df['Gender'].value_counts())
                                                   #for categorial data(works like group_by)
          Male
                    489
          Female
                    112
          Name: Gender, dtype: int64
In [13]:
          df.dtypes #displays type of each variable
          Loan_ID
                                 object
Out[13]:
          Gender
                                 object
          Married
                                 object
          Dependents
                                 object
          Education
                                 object
          Self_Employed
                                 object
          ApplicantIncome
                                  int64
          CoapplicantIncome
                                float64
          LoanAmount
                                float64
          Loan_Amount_Term
                                float64
          Credit_History
                                float64
          Property_Area
                                 object
          Loan_Status
                                 object
          dtype: object
In [14]:
          df.columns #displays variable names (Column name) as a list
          Index(['Loan_ID', 'Gender', 'Married', 'Dependents', 'Education',
Out[14]:
                 'Self_Employed', 'ApplicantIncome', 'CoapplicantIncome', 'LoanAmount',
                 'Loan_Amount_Term', 'Credit_History', 'Property_Area', 'Loan_Status'],
                dtype='object')
In [15]:
          df.columns[0] #displays 1st column name
          'Loan ID'
Out[15]:
In [16]:
          print(df)
                                                         Education Self_Employed
                Loan_ID
                         Gender Married Dependents
          0
               LP001002
                           Male
                                      No
                                                   0
                                                          Graduate
                                                                               No
                                                  1
                                                          Graduate
          1
               LP001003
                           Male
                                     Yes
                                                                               No
          2
               LP001005
                           Male
                                     Yes
                                                   a
                                                          Graduate
                                                                              Yes
          3
               LP001006
                                                  0 Not Graduate
                           Male
                                     Yes
                                                                               No
          4
               LP001008
                           Male
                                     No
                                                   0
                                                          Graduate
                                                                               No
                                     . . .
                    . . .
                             . . .
                                                 . . .
                                                                              . . .
          609
               LP002978
                                                  0
                                                          Graduate
                         Female
                                     No
                                                                               No
          610
               LP002979
                           Male
                                     Yes
                                                 3+
                                                          Graduate
                                                                               No
               LP002983
                           Male
                                                  1
                                                          Graduate
          611
                                     Yes
                                                                               No
```

ApplicantIncome CoapplicantIncome LoanAmount Loan_Amount_Term Credit_History

```
2
          612
               LP002984
                             Male
                                                             Graduate
                                       Yes
                                                                                   No
          613
               LP002990
                           Female
                                                     a
                                                             Graduate
                                        No
                                                                                  Yes
               ApplicantIncome
                                 CoapplicantIncome LoanAmount
                                                                    Loan_Amount_Term
          0
                            5849
                                                                                 360.0
                                                  0.0
                                                               NaN
                                                             128.0
                            4583
                                               1508.0
          1
                                                                                 360.0
          2
                            3000
                                                              66.0
                                                  0.0
                                                                                 360.0
          3
                            2583
                                               2358.0
                                                             120.0
                                                                                 360.0
          4
                            6000
                                                             141.0
                                                  0.0
                                                                                 360.0
                             . . .
                                                  . . .
                                                                                   . . .
          609
                            2900
                                                  0.0
                                                              71.0
                                                                                 360.0
                                                  0.0
                                                              40.0
          610
                            4106
                                                                                 180.0
          611
                            8072
                                                240.0
                                                             253.0
                                                                                 360.0
          612
                            7583
                                                  0.0
                                                             187.0
                                                                                 360.0
          613
                            4583
                                                  0.0
                                                             133.0
                                                                                 360.0
               Credit_History Property_Area Loan_Status
                            1.0
                                         Urban
          0
          1
                            1.0
                                         Rural
                                                           Ν
          2
                            1.0
                                         Urban
                                                           Υ
          3
                            1.0
                                         Urban
          4
                            1.0
                                         Urban
                                                           Υ
          609
                            1.0
                                         Rural
                                                           Υ
          610
                                                           Υ
                            1.0
                                         Rural
                                                           Υ
          611
                            1.0
                                         Urban
          612
                                         Urban
                                                           Υ
                            1.0
          613
                            0.0
                                     Semiurban
          [614 rows x 13 columns]
In [17]:
           print (df.replace('NaN',0))
                                                            Education Self_Employed
                 Loan_ID Gender Married Dependents
                                                             Graduate
          0
               LP001002
                             Male
                                       No
                                                     0
                                                                                   No
          1
               LP001003
                             Male
                                       Yes
                                                     1
                                                             Graduate
                                                                                   No
          2
               LP001005
                             Male
                                       Yes
                                                     0
                                                             Graduate
                                                                                  Yes
          3
               LP001006
                             Male
                                       Yes
                                                     0
                                                        Not Graduate
                                                                                   No
          4
               LP001008
                             Male
                                       No
                                                     0
                                                             Graduate
                                                                                   No
                              . . .
                                       . . .
                                                   . . .
          . .
                     . . .
                                                                                   . . .
               LP002978
                           Female
                                                             Graduate
          609
                                       No
                                                    0
                                                                                   No
          610
               LP002979
                             Male
                                       Yes
                                                    3+
                                                             Graduate
                                                                                   No
               LP002983
          611
                             Male
                                       Yes
                                                     1
                                                             Graduate
                                                                                   No
                                       Yes
                                                     2
          612
               LP002984
                             Male
                                                             Graduate
                                                                                   No
               LP002990
                                                     0
          613
                           Female
                                        No
                                                             Graduate
                                                                                  Yes
               ApplicantIncome CoapplicantIncome LoanAmount
                                                                    Loan_Amount_Term
          0
                            5849
                                                               NaN
                                                                                 360.0
                                                  0.0
                            4583
                                               1508.0
                                                             128.0
                                                                                 360.0
          1
          2
                            3000
                                                  0.0
                                                              66.0
                                                                                 360.0
          3
                            2583
                                               2358.0
                                                             120.0
                                                                                 360.0
          4
                            6000
                                                  0.0
                                                             141.0
                                                                                 360.0
                             . . .
                                                  . . .
                                                               . . .
                                                                                   . . .
          . .
          609
                            2900
                                                  0.0
                                                              71.0
                                                                                 360.0
          610
                            4106
                                                  0.0
                                                              40.0
                                                                                 180.0
          611
                            8072
                                                240.0
                                                             253.0
                                                                                 360.0
          612
                            7583
                                                  0.0
                                                             187.0
                                                                                 360.0
          613
                            4583
                                                  0.0
                                                             133.0
                                                                                 360.0
                Credit_History Property_Area Loan_Status
          0
                                         Urban
                                                           Υ
                            1.0
          1
                            1.0
                                         Rural
                                                           Ν
          2
                                                           Υ
                            1.0
                                         Urban
          3
                            1.0
                                         Urban
                                                           Υ
```

```
4
                          1.0
                                       Urban
                                                        Υ
                           . . .
                          1.0
                                       Rural
          610
                          1.0
                                       Rural
                                                        Υ
          611
                           1.0
                                       Urban
                                                        Υ
          612
                           1.0
                                       Urban
                                                        Υ
          613
                           0.0
                                   Semiurban
          [614 rows x 13 columns]
In [19]:
          print ("NaN replaced with '0':")
          print( df.fillna(method='pad'))
          NaN replaced with '0':
                Loan_ID Gender Married Dependents
                                                         Education Self_Employed
          0
               LP001002
                           Male
                                                          Graduate
                                     No
                                              0
                                                                               No
                                                          Graduate
          1
               LP001003
                           Male
                                     Yes
                                                   1
                                                                               No
          2
               LP001005
                           Male
                                     Yes
                                                  0
                                                          Graduate
                                                                              Yes
          3
               LP001006
                           Male
                                     Yes
                                                  0 Not Graduate
                                                                               No
               LP001008
                           Male
                                                  0
                                                          Graduate
          4
                                     No
                                                 . . .
                    . . .
                            . . .
                                     . . .
                                                          Graduate
          609
               LP002978 Female
                                     No
                                                  0
                                                                               Nο
                                                          Graduate
               LP002979
                                     Yes
                                                  3+
          610
                           Male
                                                                               No
               LP002983
                                                  1
                                                          Graduate
          611
                           Male
                                     Yes
                                                                               No
          612
               LP002984
                           Male
                                     Yes
                                                   2
                                                          Graduate
          613
              LP002990 Female
                                                          Graduate
                                                                              Yes
               ApplicantIncome CoapplicantIncome LoanAmount Loan Amount Term
          0
                           5849
                                                                              360.0
                                                0.0
                                                            NaN
          1
                           4583
                                            1508.0
                                                          128.0
                                                                             360.0
                           3000
          2
                                                           66.0
                                                0.0
                                                                             360.0
          3
                           2583
                                            2358.0
                                                          120.0
                                                                             360.0
          4
                           6000
                                                0.0
                                                          141.0
                                                                             360.0
                            . . .
                                                . . .
                                                            . . .
                                                                               . . .
                                                           71.0
          609
                           2900
                                                0.0
                                                                             360.0
          610
                           4106
                                                0.0
                                                           40.0
                                                                             180.0
         611
                           8072
                                              240.0
                                                          253.0
                                                                             360.0
         612
                          7583
                                                0.0
                                                          187.0
                                                                             360.0
                          4583
         613
                                                0.0
                                                          133.0
                                                                             360.0
               Credit_History Property_Area Loan_Status
          0
                          1.0
                                       Urban
          1
                          1.0
                                       Rural
                                                        Ν
                                       Urban
          2
                          1.0
                                                        Υ
          3
                           1.0
                                       Urban
                                                        Υ
          4
                           1.0
                                       Urban
                                                        Υ
                                         . . .
          . .
                           . . .
                                       Rural
          609
                           1.0
                                                        Υ
                                                        Υ
          610
                           1.0
                                       Rural
          611
                          1.0
                                       Urban
          612
                           1.0
                                       Urban
                                                        Υ
          613
                           0.0
                                   Semiurban
                                                        N
          [614 rows x 13 columns]
In [20]:
          print (df['Loan ID'].isnull())
          0
                 False
          1
                 False
          2
                 False
          3
                 False
          4
                 False
          609
                 False
```

```
610
                 False
                  False
          611
          612
                 False
                  False
          613
          Name: Loan_ID, Length: 614, dtype: bool
In [21]:
           print (df['Dependents'].notnull())
          0
                 True
          1
                  True
          2
                  True
          3
                  True
          4
                 True
                  . . .
          609
                 True
          610
                 True
          611
                  True
          612
                  True
          613
                  True
          Name: Dependents, Length: 614, dtype: bool
In [31]:
           print( df['Self_Employed'].isnull())
          1
                 False
          2
                  False
          3
                  False
          4
                  False
          5
                  False
          6
                  False
          7
                 False
          8
                 False
          9
                 False
          10
                 False
          12
                 False
          13
                  False
          14
                  False
          15
                  False
                 False
          17
                 False
          18
          20
                  False
          21
                 False
          22
                 False
          25
                  False
          26
                 False
          27
                 False
          28
                 False
          31
                  False
          32
                 False
          33
                  False
          34
                  False
          37
                  False
          38
                  False
          39
                  False
          575
                 False
          577
                  False
          578
                  False
          580
                 False
          581
                 False
          582
                 False
          584
                  False
          585
                  False
          586
                  False
          587
                  False
```

```
590
               False
        591
               False
               False
        593
        594
               False
        595
               False
        596
               False
        598
               False
        599
             False
        602
             False
        603
               False
        604
               False
        606
               False
        607
               False
        608
               False
        609
             False
        610
             False
              False
        611
        612
              False
        613
               False
        Name: Self_Employed, dtype: bool
In [8]:
         print(df)
         print ("NaN replaced with '0':")
         print( df.fillna(3))
              Loan_ID Gender Married Dependents
                                                     Education Self_Employed \
        0
             LP001002
                       Male
                                  No 0
                                                      Graduate
        1
             LP001003
                         Male
                                  Yes
                                              1
                                                      Graduate
                                                                          No
        2
             LP001005
                         Male
                                  Yes
                                              0
                                                      Graduate
                                                                         Yes
        3
                                              0 Not Graduate
             LP001006
                         Male
                                  Yes
                                                                          No
        4
             LP001008
                         Male
                                  No
                                              0
                                                      Graduate
                                             . . .
                                  . . .
                 . . .
                         . . .
                                                           . . .
        . .
        609 LP002978 Female
                                  No
                                              0
                                                      Graduate
                                                                          No
        610 LP002979
                                             3+
                       Male
                                  Yes
                                                      Graduate
                                                                          No
        611 LP002983
                         Male
                                  Yes
                                             1
                                                      Graduate
                                                                          No
        612 LP002984
                         Male
                                  Yes
                                              2
                                                      Graduate
                                                                          No
        613 LP002990 Female
                                  No
                                              0
                                                      Graduate
                                                                         Yes
             ApplicantIncome CoapplicantIncome LoanAmount Loan_Amount_Term
        0
                        5849
                                                       NaN
                                                                        360.0
                                            0.0
        1
                        4583
                                         1508.0
                                                      128.0
                                                                        360.0
        2
                        3000
                                            0.0
                                                       66.0
                                                                        360.0
        3
                        2583
                                         2358.0
                                                      120.0
                                                                        360.0
        4
                        6000
                                                      141.0
                                            0.0
                                                                        360.0
                         . . .
                                            . . .
                                                        . . .
                                                                          . . .
        . .
        609
                        2900
                                            0.0
                                                       71.0
                                                                        360.0
        610
                        4106
                                            0.0
                                                       40.0
                                                                        180.0
                                          240.0
                                                      253.0
        611
                        8072
                                                                        360.0
        612
                        7583
                                            0.0
                                                      187.0
                                                                        360.0
        613
                        4583
                                            0.0
                                                      133.0
                                                                        360.0
             Credit_History Property_Area Loan_Status
        0
                        1.0
                                    Urban
        1
                        1.0
                                    Rural
                                                    N
        2
                        1.0
                                    Urban
                                                    Υ
        3
                        1.0
                                    Urban
        4
                        1.0
                                    Urban
                                                    Υ
                        . . .
                                     . . .
        609
                        1.0
                                    Rural
        610
                        1.0
                                    Rural
        611
                        1.0
                                    Urban
                                                    Υ
                                                    Υ
        612
                        1.0
                                    Urban
        613
                        0.0
                                Semiurban
```

589

False

```
[614 rows x 13 columns]
         NaN replaced with '0':
              Loan ID Gender Married Dependents
                                                  Education Self_Employed \
         0
             LP001002
                        Male
                                 No 0
                                                    Graduate
                                             1
                         Male
                                 Yes
         1
             LP001003
                                                    Graduate
                                                                        No
                                             0
         2
             LP001005
                        Male
                                 Yes
                                                     Graduate
                                                                       Yes
                                             0 Not Graduate
         3
             LP001006
                        Male
                                 Yes
                                                                       No
         4
             LP001008
                         Male
                                 No
                                             0
                                                     Graduate
                                                                        No
                                 . . .
                                            . . .
                . . .
                        . . .
                                                                       . . .
         . .
         609 LP002978 Female
                                 No
                                             0
                                                     Graduate
                                                                        No
                                 Yes
         610 LP002979
                        Male
                                            3+
                                                     Graduate
                                                                        No
         611 LP002983
                         Male
                                 Yes
                                             1
                                                     Graduate
                                                                        No
                                             2
         612 LP002984
                        Male
                                 Yes
                                                     Graduate
                                                                        No
         613 LP002990 Female
                                 No
                                             0
                                                    Graduate
                                                                       Yes
             ApplicantIncome CoapplicantIncome LoanAmount Loan Amount Term
                        5849
                                           0.0
                                                     3.0
         1
                        4583
                                        1508.0
                                                    128.0
                                                                      360.0
         2
                        3000
                                           0.0
                                                     66.0
                                                                      360.0
         3
                        2583
                                        2358.0
                                                     120.0
                                                                      360.0
         4
                        6000
                                          0.0
                                                     141.0
                                                                      360.0
                        . . .
                                           . . .
                                                     . . .
         609
                        2900
                                          0.0
                                                     71.0
                                                                      360.0
         610
                        4106
                                          0.0
                                                     40.0
                                                                      180.0
         611
                        8072
                                         240.0
                                                     253.0
                                                                      360.0
         612
                        7583
                                           0.0
                                                     187.0
                                                                      360.0
         613
                        4583
                                           0.0
                                                     133.0
                                                                      360.0
             Credit_History Property_Area Loan_Status
                                  Urban
         0
                        1.0
         1
                        1.0
                                   Rural
         2
                        1.0
                                   Urban
                                                   Υ
         3
                        1.0
                                   Urban
         4
                        1.0
                                   Urban
                                                   Υ
                        . . .
                                    . . .
         . .
                                                 . . .
                                   Rural
         609
                        1.0
                                                   Υ
         610
                        1.0
                                   Rural
                                                   Υ
         611
                        1.0
                                   Urban
                                                   Υ
         612
                        1.0
                                   Urban
                                                   Υ
                               Semiurban
         613
                        0.0
         [614 rows x 13 columns]
In [9]:
         df = df.dropna() #drops rows with null values
In [24]:
         print(df)
              Loan ID Gender Married Dependents
                                                    Education Self_Employed \
         1
             LP001003
                       Male
                                 Yes 1
                                                    Graduate
                                                                       No
                         Male
                                                     Graduate
         2
             LP001005
                                 Yes
                                             0
                                                                       Yes
                                 Yes
                                             0 Not Graduate
         3
             LP001006
                         Male
                                                                        No
                                             0
         4
             LP001008
                         Male
                                 No
                                                     Graduate
                                                                        No
         5
             LP001011
                         Male
                                 Yes
                                             2
                                                     Graduate
                                                                       Yes
                                 . . .
                         . . .
                                            . . .
                                                        . . .
         . .
                 . . .
         609 LP002978 Female
                                 No
                                             0
                                                     Graduate
                                                                        No
         610 LP002979
                       Male
                                 Yes
                                             3+
                                                     Graduate
                                                                        No
         611 LP002983
                         Male
                                                     Graduate
                                 Yes
                                             1
                                                                        No
                         Male
                                 Yes
                                                     Graduate
         612 LP002984
                                             2
                                                                        No
         613 LP002990 Female
                                 No
                                             0
                                                     Graduate
                                                                       Yes
             ApplicantIncome CoapplicantIncome LoanAmount Loan_Amount_Term \
                                        1508.0
                                                                      360.0
         1
                        4583
                                                     128.0
```

```
2
                            3000
                                                              66.0
                                                  0.0
                                                                                 360.0
          3
                            2583
                                                             120.0
                                                                                 360.0
                                               2358.0
          4
                            6000
                                                  0.0
                                                             141.0
                                                                                 360.0
          5
                            5417
                                               4196.0
                                                             267.0
                                                                                 360.0
          . .
                                                  . . .
                                                                                   . . .
                                                              71.0
          609
                            2900
                                                  0.0
                                                                                 360.0
                            4106
                                                  0.0
                                                              40.0
                                                                                 180.0
          610
                            8072
                                                240.0
                                                             253.0
                                                                                 360.0
          611
          612
                            7583
                                                  0.0
                                                                                 360.0
                                                             187.0
          613
                            4583
                                                  0.0
                                                             133.0
                                                                                 360.0
                Credit_History Property_Area Loan_Status
          1
                            1.0
                                         Rural
          2
                            1.0
                                         Urban
                                                           Υ
          3
                            1.0
                                         Urban
                                                           Υ
                                                           Υ
          4
                            1.0
                                         Urban
          5
                            1.0
                                         Urban
                                                           Υ
                                            . . .
                            . . .
          . .
          609
                            1.0
                                         Rural
                                                           Υ
                                         Rural
          610
                            1.0
                                                           Υ
          611
                            1.0
                                         Urban
                                         Urban
          612
                            1.0
                                                           Υ
                                     Semiurban
          613
                            0.0
                                                           Ν
          [480 rows x 13 columns]
In [13]:
           import pandas as pd
           import numpy as np
           df = pd.read_csv("C:\\Users\\Suresh Jamadagni\\downloads\\train.csv")
           df.head()
              Loan_ID Gender Married Dependents
                                                     Education Self_Employed ApplicantIncome Coapplican
Out[13]:
          0 LP001002
                          Male
                                    No
                                                  0
                                                      Graduate
                                                                         No
                                                                                         5849
             LP001003
                          Male
                                    Yes
                                                      Graduate
                                                                          No
                                                                                         4583
            LP001005
                                                      Graduate
                                                                                         3000
                          Male
                                                  0
                                    Yes
                                                                         Yes
                                                          Not
          3 LP001006
                          Male
                                                                                         2583
                                    Yes
                                                                         No
                                                      Graduate
          4 LP001008
                          Male
                                    No
                                                      Graduate
                                                                          No
                                                                                         6000
In [14]:
           to_drop = ['Gender','Married']
           #df.drop(columns=to_drop, inplace=True)
           df.drop(to_drop, inplace=True, axis=1)
In [15]:
           df.head()
Out[15]:
              Loan ID
                       Dependents
                                    Education Self_Employed ApplicantIncome CoapplicantIncome
                                                                                                LoanAmo
          0 LP001002
                                     Graduate
                                                                        5849
                                 0
                                                         No
                                                                                            0.0
             LP001003
                                 1
                                     Graduate
                                                         No
                                                                        4583
                                                                                         1508.0
                                                                                                        1
            LP001005
                                 0
                                     Graduate
                                                         Yes
                                                                        3000
                                                                                            0.0
                                         Not
          3 LP001006
                                 0
                                                         No
                                                                        2583
                                                                                         2358.0
                                                                                                        1
                                     Graduate
```

	4	LP001008	0	Graduate	No	6000	0.0	1
In []:								
In []:								

Loan_ID Dependents Education Self_Employed ApplicantIncome CoapplicantIncome LoanAmo