C4_S3_Challenge

In [1]:

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

In [2]:

```
import warnings
warnings.filterwarnings('ignore')
```

In [3]:

```
pd.Series()
```

Out[3]:

```
Series([], dtype: float64)
```

In [4]:

```
ld= pd.read_csv(r'E:\Aishwarya official\Aishwarya Data Scince\Course 4\DS1_C4_S3_Loan_Data_Practice.csv.csv')
ld
```

Out[4]:

	Loan_ID	Gender	Married	Dependents	Education	Self_Employed	ApplicantIncome	CoapplicantIncome	LoanAmount	Loan_Amount_Term	Credit_H
0	LP001002	Male	No	0	Graduate	No	5849	0.0	146	360	
1	LP001003	Male	Yes	1	Graduate	No	4583	1508.0	128	360	
2	LP001005	Male	Yes	0	Graduate	Yes	3000	0.0	66	360	
3	LP001006	Male	Yes	0	Not Graduate	No	2583	2358.0	120	360	
4	LP001008	Male	No	0	Graduate	No	6000	0.0	141	360	
...	...	...	...	...	...	...	...	...	...	...	...
609	LP002978	Female	No	0	Graduate	No	2900	0.0	71	360	
610	LP002979	Male	Yes	3+	Graduate	No	4106	0.0	40	180	
611	LP002983	Male	Yes	1	Graduate	No	8072	240.0	253	360	
612	LP002984	Male	Yes	2	Graduate	No	7583	0.0	187	360	
613	LP002990	Female	No	0	Graduate	Yes	4583	0.0	133	360	

614 rows × 13 columns

Task 1 :

In [5]:

```
x= ld[(ld.ApplicantIncome <6000) & (ld.LoanAmount> 200) & (ld.Loan_Status=='Y')]
x
```

Out[5]:

	Loan_ID	Gender	Married	Dependents	Education	Self_Employed	ApplicantIncome	CoapplicantIncome	LoanAmount	Loan_Amount_Term	Credit_H
5	LP001011	Male	Yes	2	Graduate	Yes	5417	4196.0	267	360	
21	LP001046	Male	Yes	1	Graduate	No	5955	5625.0	315	360	
159	LP001552	Male	Yes	0	Graduate	No	4583	5625.0	255	360	
253	LP001843	Male	Yes	1	Not Graduate	No	2661	7101.0	279	180	
255	LP001846	Female	No	3+	Graduate	No	3083	0.0	255	360	
276	LP001903	Male	Yes	0	Graduate	No	3993	3274.0	207	360	
361	LP002170	Male	Yes	2	Graduate	No	5000	3667.0	236	360	
381	LP002229	Male	No	0	Graduate	No	5941	4232.0	296	360	
502	LP002615	Male	Yes	2	Graduate	No	4865	5624.0	208	360	
505	LP002622	Male	Yes	2	Graduate	No	3510	4416.0	243	360	
530	LP002717	Male	Yes	0	Graduate	No	1025	5500.0	216	360	
562	LP002820	Male	Yes	0	Graduate	No	5923	2054.0	211	360	

In [6]:

```
print("The no of applicant", x['Gender'].value_counts())
```

```
The no of applicant Male      11
Female     1
Name: Gender, dtype: int64
```

Task 2:

In [7]:

```
Semiurban = ld.value_counts(ld.Property_Area)[0]
Semiurban
```

Out[7]:

```
233
```

In [8]:

```
Urban = ld.value_counts(ld.Property_Area)[1]
Urban
```

Out[8]:

```
202
```

In [9]:

```
print()
```

In [10]:

```
print("Relative % in Urban",((( Semiurban - Urban)/Urban)  *100/2))
```

```
Relative % in Urban 7.673267326732673
```

Task 3 :

In [11]:

```
z = ld[(ld.Loan_Status=='Y')]
z
```

Out[11]:

	Loan_ID	Gender	Married	Dependents	Education	Self_Employed	ApplicantIncome	CoapplicantIncome	LoanAmount	Loan_Amount_Term	Credit_H
0	LP001002	Male	No	0	Graduate	No	5849	0.0	146	360	
2	LP001005	Male	Yes	0	Graduate	Yes	3000	0.0	66	360	
3	LP001006	Male	Yes	0	Not Graduate	No	2583	2358.0	120	360	
4	LP001008	Male	No	0	Graduate	No	6000	0.0	141	360	
5	LP001011	Male	Yes	2	Graduate	Yes	5417	4196.0	267	360	
...	...	...	...	...	...	...	...	...	...	...	...
608	LP002974	Male	Yes	0	Graduate	No	3232	1950.0	108	360	
609	LP002978	Female	No	0	Graduate	No	2900	0.0	71	360	
610	LP002979	Male	Yes	3+	Graduate	No	4106	0.0	40	180	
611	LP002983	Male	Yes	1	Graduate	No	8072	240.0	253	360	
612	LP002984	Male	Yes	2	Graduate	No	7583	0.0	187	360	

422 rows × 13 columns

In [12]:

```
sort = z.sort_values('ApplicantIncome',ascending = False)[:5]
sort
```

Out[12]:

	Loan_ID	Gender	Married	Dependents	Education	Self_Employed	ApplicantIncome	CoapplicantIncome	LoanAmount	Loan_Amount_Term	Credit_H
333	LP002101	Male	Yes	0	Graduate	No	63337	0.0	490	180	
171	LP001585	Male	Yes	3+	Graduate	No	51763	0.0	700	300	
155	LP001536	Male	Yes	3+	Graduate	No	39999	0.0	600	180	
185	LP001640	Male	Yes	0	Graduate	Yes	39147	4750.0	120	360	
443	LP002422	Male	No	1	Graduate	No	37719	0.0	152	360	

Task 4 :

In [13]:

```
ld[(ld.Gender=='Female')& (ld.Dependents == '2')].iloc[:,[1,6,8]]
```

Out[13]:

	Gender	ApplicantIncome	LoanAmount
29	Female	3750	120
82	Female	1378	167
146	Female	14866	70
219	Female	4283	127
251	Female	3427	138
293	Female	5417	143
468	Female	210	98
516	Female	2031	113

Task 5 :

In [14]:

```
ld.loc[ld.LoanAmount >200, 'Loan_Amount_Term'] = 180
ld
```

Out[14]:

	Loan_ID	Gender	Married	Dependents	Education	Self_Employed	ApplicantIncome	CoapplicantIncome	LoanAmount	Loan_Amount_Term	Credit_H
0	LP001002	Male	No	0	Graduate	No	5849	0.0	146	360	
1	LP001003	Male	Yes	1	Graduate	No	4583	1508.0	128	360	
2	LP001005	Male	Yes	0	Graduate	Yes	3000	0.0	66	360	
3	LP001006	Male	Yes	0	Not Graduate	No	2583	2358.0	120	360	
4	LP001008	Male	No	0	Graduate	No	6000	0.0	141	360	
...	...	...	...	...	...	...	...	...	...	...	...
609	LP002978	Female	No	0	Graduate	No	2900	0.0	71	360	
610	LP002979	Male	Yes	3+	Graduate	No	4106	0.0	40	180	
611	LP002983	Male	Yes	1	Graduate	No	8072	240.0	253	180	
612	LP002984	Male	Yes	2	Graduate	No	7583	0.0	187	360	
613	LP002990	Female	No	0	Graduate	Yes	4583	0.0	133	360	

614 rows × 13 columns

Task 6 :

In [15]:

```
newdata = ld
newdata
```

Out[15]:

	Loan_ID	Gender	Married	Dependents	Education	Self_Employed	ApplicantIncome	CoapplicantIncome	LoanAmount	Loan_Amount_Term	Credit_H
0	LP001002	Male	No	0	Graduate	No	5849	0.0	146	360	
1	LP001003	Male	Yes	1	Graduate	No	4583	1508.0	128	360	
2	LP001005	Male	Yes	0	Graduate	Yes	3000	0.0	66	360	
3	LP001006	Male	Yes	0	Not Graduate	No	2583	2358.0	120	360	
4	LP001008	Male	No	0	Graduate	No	6000	0.0	141	360	
...	...	...	...	...	...	...	...	...	...	...	...
609	LP002978	Female	No	0	Graduate	No	2900	0.0	71	360	
610	LP002979	Male	Yes	3+	Graduate	No	4106	0.0	40	180	
611	LP002983	Male	Yes	1	Graduate	No	8072	240.0	253	180	
612	LP002984	Male	Yes	2	Graduate	No	7583	0.0	187	360	
613	LP002990	Female	No	0	Graduate	Yes	4583	0.0	133	360	

614 rows × 13 columns

In [16]:

```
newdata['Loan_Amount_Term'] = newdata['Loan_Amount_Term'].replace([12,36],60)
newdata
```

Out[16]:

	Loan_ID	Gender	Married	Dependents	Education	Self_Employed	ApplicantIncome	CoapplicantIncome	LoanAmount	Loan_Amount_Term	Credit_H
0	LP001002	Male	No	0	Graduate	No	5849	0.0	146	360	
1	LP001003	Male	Yes	1	Graduate	No	4583	1508.0	128	360	
2	LP001005	Male	Yes	0	Graduate	Yes	3000	0.0	66	360	
3	LP001006	Male	Yes	0	Not Graduate	No	2583	2358.0	120	360	
4	LP001008	Male	No	0	Graduate	No	6000	0.0	141	360	
...	...	...	...	...	...	...	...	...	...	...	...
609	LP002978	Female	No	0	Graduate	No	2900	0.0	71	360	
610	LP002979	Male	Yes	3+	Graduate	No	4106	0.0	40	180	
611	LP002983	Male	Yes	1	Graduate	No	8072	240.0	253	180	
612	LP002984	Male	Yes	2	Graduate	No	7583	0.0	187	360	
613	LP002990	Female	No	0	Graduate	Yes	4583	0.0	133	360	

614 rows × 13 columns

In [22]:

```
newdata = newdata.rename(columns = {'Loan_Amount_Term' : 'LoanAmountTerm'})
newdata
```

Out[22]:

Dependents	Education	Self_Employed	ApplicantIncome	CoapplicantIncome	LoanAmount	LoanAmountTerm	Credit_History	Property_Area	Loan_Status
0	Graduate	No	5849	0.0	146	360	1	Urban	Y
1	Graduate	No	4583	1508.0	128	360	1	Rural	N
0	Graduate	Yes	3000	0.0	66	360	1	Urban	Y
0	Not Graduate	No	2583	2358.0	120	360	1	Urban	Y
0	Graduate	No	6000	0.0	141	360	1	Urban	Y
...	...	...	...	...	...	...	...	...	...
0	Graduate	No	2900	0.0	71	360	1	Rural	Y
3+	Graduate	No	4106	0.0	40	180	1	Rural	Y
1	Graduate	No	8072	240.0	253	180	1	Urban	Y
2	Graduate	No	7583	0.0	187	360	1	Urban	Y
0	Graduate	Yes	4583	0.0	133	360	0	Semiurban	N

Task 7:

In [17]:

```
credit = ld[(ld.Married == 'Yes') & (ld.Education == 'Graduate') & (ld.ApplicantIncome>10000)]
credit["Credit_Limit"]= credit.ApplicantIncome*2
credit
```

Out[17]:

	Loan_ID	Gender	Married	Dependents	Education	Self_Employed	ApplicantIncome	CoapplicantIncome	LoanAmount	Loan_Amount_Term	Credit_H
9	LP001020	Male	Yes	1	Graduate	No	12841	10968.0	349	180	
54	LP001186	Female	Yes	1	Graduate	Yes	11500	0.0	286	180	
67	LP001233	Male	Yes	1	Graduate	No	10750	0.0	312	180	
102	LP001350	Male	Yes	1	Graduate	No	13650	0.0	146	360	
106	LP001369	Male	Yes	2	Graduate	No	11417	1126.0	225	180	
115	LP001401	Male	Yes	1	Graduate	No	14583	0.0	185	180	
126	LP001448	Male	Yes	3+	Graduate	No	23803	0.0	370	180	
128	LP001451	Male	Yes	1	Graduate	Yes	10513	3850.0	160	180	
144	LP001508	Male	Yes	2	Graduate	No	11757	0.0	187	180	
146	LP001516	Female	Yes	2	Graduate	No	14866	0.0	70	360	
155	LP001536	Male	Yes	3+	Graduate	No	39999	0.0	600	180	
171	LP001585	Male	Yes	3+	Graduate	No	51763	0.0	700	180	
183	LP001637	Male	Yes	1	Graduate	No	33846	0.0	260	180	
185	LP001640	Male	Yes	0	Graduate	Yes	39147	4750.0	120	360	
258	LP001859	Male	Yes	0	Graduate	No	14683	2100.0	304	180	
271	LP001891	Male	Yes	0	Graduate	No	11146	0.0	136	360	
278	LP001907	Male	Yes	0	Graduate	No	14583	0.0	436	180	
284	LP001922	Male	Yes	0	Graduate	No	20667	0.0	146	360	
324	LP002065	Male	Yes	3+	Graduate	No	15000	0.0	300	180	
333	LP002101	Male	Yes	0	Graduate	No	63337	0.0	490	180	
369	LP002191	Male	Yes	0	Graduate	No	19730	5266.0	570	180	
409	LP002317	Male	Yes	3+	Graduate	No	81000	0.0	360	180	
424	LP002364	Male	Yes	0	Graduate	No	14880	0.0	96	360	
435	LP002393	Female	Yes	1	Graduate	No	10047	0.0	146	240	
467	LP002501	Male	Yes	0	Graduate	No	16692	0.0	110	360	
475	LP002527	Male	Yes	2	Graduate	Yes	16525	1014.0	150	360	
478	LP002531	Male	Yes	1	Graduate	Yes	16667	2250.0	86	360	
483	LP002541	Male	Yes	0	Graduate	No	10833	0.0	234	180	
487	LP002547	Male	Yes	1	Graduate	No	18333	0.0	500	180	
506	LP002624	Male	Yes	0	Graduate	No	20833	6667.0	480	180	
525	LP002699	Male	Yes	2	Graduate	Yes	17500	0.0	400	180	
557	LP002795	Male	Yes	3+	Graduate	Yes	10139	0.0	260	180	
561	LP002813	Female	Yes	1	Graduate	Yes	19484	0.0	600	180	
572	LP002855	Male	Yes	2	Graduate	No	16666	0.0	275	180	
594	LP002938	Male	Yes	0	Graduate	Yes	16120	0.0	260	180	
604	LP002959	Female	Yes	1	Graduate	No	12000	0.0	496	180	

Task 8 :

In [18]:

```
ld
```

Out[18]:

	Loan_ID	Gender	Married	Dependents	Education	Self_Employed	ApplicantIncome	CoapplicantIncome	LoanAmount	Loan_Amount_Term	Credit_H
0	LP001002	Male	No	0	Graduate	No	5849	0.0	146	360	
1	LP001003	Male	Yes	1	Graduate	No	4583	1508.0	128	360	
2	LP001005	Male	Yes	0	Graduate	Yes	3000	0.0	66	360	
3	LP001006	Male	Yes	0	Not Graduate	No	2583	2358.0	120	360	
4	LP001008	Male	No	0	Graduate	No	6000	0.0	141	360	
...	...	...	...	...	...	...	...	...	...	...	...
609	LP002978	Female	No	0	Graduate	No	2900	0.0	71	360	
610	LP002979	Male	Yes	3+	Graduate	No	4106	0.0	40	180	
611	LP002983	Male	Yes	1	Graduate	No	8072	240.0	253	180	
612	LP002984	Male	Yes	2	Graduate	No	7583	0.0	187	360	
613	LP002990	Female	No	0	Graduate	Yes	4583	0.0	133	360	

614 rows × 13 columns

In [24]:

```
Delete = ld[(ld.Dependents == '3+') & (ld.Education == 'Not Graduate')&(ld.Credit_History ==0)]
Delete
```

Out[24]:

	Loan_ID	Gender	Married	Dependents	Education	Self_Employed	ApplicantIncome	CoapplicantIncome	LoanAmount	Loan_Amount_Term	Credit_H
73	LP001250	Male	Yes	3+	Not Graduate	No	4755	0.0	95	342	
338	LP002113	Female	No	3+	Not Graduate	No	1830	0.0	146	360	
466	LP002500	Male	Yes	3+	Not Graduate	No	2947	1664.0	70	180	

In [25]:

```
Delete = Delete.drop([73,338,466], axis = 0)
Delete
```

Out[25]:

	Loan_ID	Gender	Married	Dependents	Education	Self_Employed	ApplicantIncome	CoapplicantIncome	LoanAmount	Loan_Amount_Term	Credit_Histor

Task 9 :

In [20]:

```
gra = ld[ld.Education == 'Graduate'].ApplicantIncome.mean()
Ngra = ld[ld.Education == 'Not Graduate'].ApplicantIncome.mean()
print("Average of difference of graduate and non graduate :- ", (gra-Ngra,2))
```

```
Average of difference of graduate and non graduate :-  (2080.1497512437813, 2)
```

In [ ]:

```

```