

```
In [1]: import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
import os
import warnings
warnings.filterwarnings('ignore')
```

```
In [2]: data=pd.read_csv("C:/Users/Admin/Downloads/supermarket_purchase.csv")
```

In [3]: data

Out[3]:

	Age	Job Type	Marital Status	Education	Metro City	Signed in since(Days)	Purchasemade
0	28.0	Unemployed	No	Graduate	No	32	261.4689
1	57.0	Retired	Yes	Graduate	Yes	31	273.9025
2	27.0	Student	Yes	Graduate	Yes	31	275.8921
3	23.0	Unemployed	Yes	Secondry	Yes	34	295.1524
4	23.0	Employed	Yes	Graduate	Yes	42	325.0809
5	58.0	Retired	Yes	Graduate	Yes	31	327.6100
6	26.0	Employed	No	Graduate	No	40	338.9281
7	NaN	Unemployed	No	Graduate	No	42	355.3225
8	56.0	Retired	No	Graduate	No	31	362.9025
9	23.0	Unemployed	Yes	Secondry	Yes	31	364.8100
10	57.0	Retired	Yes	Secondry	Yes	37	446.4769
11	29.0	Unemployed	Yes	Secondry	Yes	45	475.6761
12	59.0	Retired	No	Graduate	No	31	484.0000
13	NaN	Unemployed	Yes	Graduate	Yes	41	512.1169
14	61.0	Retired	Yes	Graduate	Yes	39	554.1316
15	60.0	Retired	No	Graduate	No	31	576.0000
16	58.0	Retired	Yes	Graduate	Yes	35	577.9216
17	65.0	Retired	Yes	Graduate	Yes	34	577.9216
18	59.0	Retired	No	Graduate	No	40	590.4900
19	55.0	Employed	Yes	Secondry	Yes	52	1313.3376
20	46.0	Retired	Yes	Graduate	Yes	42	1327.8736
21	30.0	Employed	Yes	Graduate	Yes	33	1343.9556
22	53.0	Employed	No	Graduate	No	31	1369.0000
23	34.0	Employed	No	Graduate	No	31	1369.0000
24	51.0	Employed	No	Graduate	Yes	34	1397.2644
25	34.0	Employed	Yes	Graduate	Yes	31	1410.0025
26	46.0	Employed	Yes	Graduate	Yes	44	2217.4681
27	29.0	Employed	Yes	Secondry	Yes	50	2226.8961
28	26.0	Unemployed	Yes	Graduate	Yes	64	2244.8644
29	30.0	Employed	No	Graduate	No	58	2270.5225
...	...	...	...	...	...	...	...
295	36.0	Employed	Yes	Graduate	Yes	98	12989.1609
296	48.0	Employed	Yes	Graduate	Yes	94	13011.9649
297	41.0	Employed	Yes	Graduate	Yes	98	13039.3561
298	33.0	Employed	Yes	Graduate	Yes	95	13043.9241
299	36.0	Employed	Yes	Graduate	Yes	91	13121.7025
300	35.0	Employed	Yes	Graduate	Yes	93	13261.8256

	Age	Job Type	Marital Status	Education	Metro City	Signed in since(Days)	Purchasemade
301	39.0	Employed	Yes	Graduate	Yes	97	13275.6484
302	37.0	Employed	Yes	Graduate	Yes	96	13291.7841
303	48.0	Employed	No	Graduate	Yes	100	13411.9561
304	45.0	Employed	No	Graduate	Yes	93	13500.1161
305	40.0	Employed	Yes	Graduate	Yes	99	13514.0625
306	32.0	Employed	Yes	Graduate	Yes	93	13560.6025
307	41.0	Employed	Yes	Graduate	Yes	100	13691.3401
308	45.0	Employed	No	Graduate	Yes	95	13691.3401
309	38.0	Employed	Yes	Graduate	Yes	95	13707.7264
310	35.0	Employed	Yes	Graduate	Yes	94	14042.2500
311	47.0	Employed	Yes	Graduate	Yes	98	14232.4900
312	32.0	Employed	No	Graduate	Yes	96	14390.4016
313	42.0	Employed	No	Graduate	Yes	94	14395.2004
314	33.0	Employed	Yes	Graduate	Yes	95	14400.0000
315	40.0	Employed	Yes	Graduate	Yes	100	14604.7225
316	36.0	Employed	No	Graduate	Yes	97	14718.5424
317	48.0	Employed	Yes	Graduate	Yes	97	15160.9969
318	45.0	Employed	No	Graduate	Yes	97	15212.7556
319	44.0	Employed	Yes	Graduate	Yes	98	15331.3924
320	38.0	Employed	Yes	Graduate	Yes	100	15645.0064
321	44.0	Employed	Yes	Graduate	Yes	99	16184.9284
322	41.0	Employed	Yes	Graduate	Yes	98	16187.4729
323	41.0	Employed	Yes	Graduate	Yes	99	16993.7296
324	39.0	Employed	No	Graduate	Yes	99	17355.4276

325 rows × 7 columns

In [4]: data.head()

Out[4]:

	Age	Job Type	Marital Status	Education	Metro City	Signed in since(Days)	Purchasemade
0	28.0	Unemployed	No	Graduate	No	32	261.4689
1	57.0	Retired	Yes	Graduate	Yes	31	273.9025
2	27.0	Student	Yes	Graduate	Yes	31	275.8921
3	23.0	Unemployed	Yes	Secondry	Yes	34	295.1524
4	23.0	Employed	Yes	Graduate	Yes	42	325.0809

```
In [5]: data.tail()
```

```
Out[5]:
```

	Age	Job Type	Marital Status	Education	Metro City	Signed in since(Days)	Purchasemade
320	38.0	Employed	Yes	Graduate	Yes	100	15645.0064
321	44.0	Employed	Yes	Graduate	Yes	99	16184.9284
322	41.0	Employed	Yes	Graduate	Yes	98	16187.4729
323	41.0	Employed	Yes	Graduate	Yes	99	16993.7296
324	39.0	Employed	No	Graduate	Yes	99	17355.4276

```
In [6]: data.isnull().sum()
```

```
Out[6]: Age                13
Job Type                0
Marital Status          0
Education                0
Metro City              0
Signed in since(Days)    0
Purchasemade            0
dtype: int64
```

```
In [7]: data['Age'].fillna(data['Age'].mean(),inplace=True)
```

```
In [8]: data.isnull().sum()
```

```
Out[8]: Age                0
Job Type                0
Marital Status          0
Education                0
Metro City              0
Signed in since(Days)    0
Purchasemade            0
dtype: int64
```

```
In [9]: Job_Type_dum=data['Job Type']=pd.get_dummies(data['Job Type'],drop_first=True)
```

In [10]:

Job\_Type\_dum

Out[10]:

	Retired	Student	Unemployed
0	0	0	1
1	1	0	0
2	0	1	0
3	0	0	1
4	0	0	0
5	1	0	0
6	0	0	0
7	0	0	1
8	1	0	0
9	0	0	1
10	1	0	0
11	0	0	1
12	1	0	0
13	0	0	1
14	1	0	0
15	1	0	0
16	1	0	0
17	1	0	0
18	1	0	0
19	0	0	0
20	1	0	0
21	0	0	0
22	0	0	0
23	0	0	0
24	0	0	0
25	0	0	0
26	0	0	0
27	0	0	0
28	0	0	1
29	0	0	0
...	...	...	...
295	0	0	0
296	0	0	0
297	0	0	0
298	0	0	0
299	0	0	0
300	0	0	0

	Retired	Student	Unemployed
301	0	0	0
302	0	0	0
303	0	0	0
304	0	0	0
305	0	0	0
306	0	0	0
307	0	0	0
308	0	0	0
309	0	0	0
310	0	0	0
311	0	0	0
312	0	0	0
313	0	0	0
314	0	0	0
315	0	0	0
316	0	0	0
317	0	0	0
318	0	0	0
319	0	0	0
320	0	0	0
321	0	0	0
322	0	0	0
323	0	0	0
324	0	0	0

325 rows × 3 columns

```
In [11]: Marital_Status_dum=data['Marital Status']=pd.get_dummies(data['Marital Status'])
```

```
In [12]: Marital_Status_dum
```

Out[12]:

	No	Yes
0	1	0
1	0	1
2	0	1
3	0	1
4	0	1
5	0	1
6	1	0
7	1	0
8	1	0
9	0	1
10	0	1
11	0	1
12	1	0
13	0	1
14	0	1
15	1	0
16	0	1
17	0	1
18	1	0
19	0	1
20	0	1
21	0	1
22	1	0
23	1	0
24	1	0
25	0	1
26	0	1
27	0	1
28	0	1
29	1	0
...	...	...
295	0	1
296	0	1
297	0	1
298	0	1
299	0	1
300	0	1

	No	Yes
301	0	1
302	0	1
303	1	0
304	1	0
305	0	1
306	0	1
307	0	1
308	1	0
309	0	1
310	0	1
311	0	1
312	1	0
313	1	0
314	0	1
315	0	1
316	1	0
317	0	1
318	1	0
319	0	1
320	0	1
321	0	1
322	0	1
323	0	1
324	1	0

325 rows × 2 columns

```
In [13]: Edu_dum=data['Education']=pd.get_dummies(data['Education'],drop_first=True)
```



In [14]: Edu\_dum

Out[14]:

	Primary	Secondry
0	0	0
1	0	0
2	0	0
3	0	1
4	0	0
5	0	0
6	0	0
7	0	0
8	0	0
9	0	1
10	0	1
11	0	1
12	0	0
13	0	0
14	0	0
15	0	0
16	0	0
17	0	0
18	0	0
19	0	1
20	0	0
21	0	0
22	0	0
23	0	0
24	0	0
25	0	0
26	0	0
27	0	1
28	0	0
29	0	0
...	...	...
295	0	0
296	0	0
297	0	0
298	0	0
299	0	0
300	0	0

	Primary	Secondry
301	0	0
302	0	0
303	0	0
304	0	0
305	0	0
306	0	0
307	0	0
308	0	0
309	0	0
310	0	0
311	0	0
312	0	0
313	0	0
314	0	0
315	0	0
316	0	0
317	0	0
318	0	0
319	0	0
320	0	0
321	0	0
322	0	0
323	0	0
324	0	0

325 rows × 2 columns

```
In [15]: Metro_City_dum=data['Metro City']=pd.get_dummies(data['Metro City'],drop_first=True)
```

In [16]: Metro\_City\_dum

Out[16]:

	Yes
0	0
1	1
2	1
3	1
4	1
5	1
6	0
7	0
8	0
9	1
10	1
11	1
12	0
13	1
14	1
15	0
16	1
17	1
18	0
19	1
20	1
21	1
22	0
23	0
24	1
25	1
26	1
27	1
28	1
29	0
...	...
295	1
296	1
297	1
298	1
299	1
300	1

	Yes
301	1
302	1
303	1
304	1
305	1
306	1
307	1
308	1
309	1
310	1
311	1
312	1
313	1
314	1
315	1
316	1
317	1
318	1
319	1
320	1
321	1
322	1
323	1
324	1

325 rows × 1 columns

```
In [17]: data=pd.concat([data,Job_Type_dum,Marital_Status_dum,Edu_dum,Metro_City_dum],axis=1)
```

In [18]: data

Out[18]:

	Age	Job Type	Marital Status	Education	Metro City	Signed in since(Days)	Purchasemade	Retired	Student	Unemployed	N
0	28.000000	0	1	0	0	32	261.4689	0	0	1	
1	57.000000	1	0	0	1	31	273.9025	1	0	0	
2	27.000000	0	0	0	1	31	275.8921	0	1	0	
3	23.000000	0	0	0	1	34	295.1524	0	0	1	
4	23.000000	0	0	0	1	42	325.0809	0	0	0	
5	58.000000	1	0	0	1	31	327.6100	1	0	0	
6	26.000000	0	1	0	0	40	338.9281	0	0	0	
7	38.349359	0	1	0	0	42	355.3225	0	0	1	
8	56.000000	1	1	0	0	31	362.9025	1	0	0	
9	23.000000	0	0	0	1	31	364.8100	0	0	1	
10	57.000000	1	0	0	1	37	446.4769	1	0	0	
11	29.000000	0	0	0	1	45	475.6761	0	0	1	
12	59.000000	1	1	0	0	31	484.0000	1	0	0	
13	38.349359	0	0	0	1	41	512.1169	0	0	1	
14	61.000000	1	0	0	1	39	554.1316	1	0	0	
15	60.000000	1	1	0	0	31	576.0000	1	0	0	
16	58.000000	1	0	0	1	35	577.9216	1	0	0	
17	65.000000	1	0	0	1	34	577.9216	1	0	0	
18	59.000000	1	1	0	0	40	590.4900	1	0	0	
19	55.000000	0	0	0	1	52	1313.3376	0	0	0	
20	46.000000	1	0	0	1	42	1327.8736	1	0	0	
21	30.000000	0	0	0	1	33	1343.9556	0	0	0	
22	53.000000	0	1	0	0	31	1369.0000	0	0	0	
23	34.000000	0	1	0	0	31	1369.0000	0	0	0	
24	51.000000	0	1	0	1	34	1397.2644	0	0	0	
25	34.000000	0	0	0	1	31	1410.0025	0	0	0	
26	46.000000	0	0	0	1	44	2217.4681	0	0	0	
27	29.000000	0	0	0	1	50	2226.8961	0	0	0	
28	26.000000	0	0	0	1	64	2244.8644	0	0	1	
29	30.000000	0	1	0	0	58	2270.5225	0	0	0	
...	...	...	...	...	...	...	...	...	...	...	
295	36.000000	0	0	0	1	98	12989.1609	0	0	0	
296	48.000000	0	0	0	1	94	13011.9649	0	0	0	
297	41.000000	0	0	0	1	98	13039.3561	0	0	0	
298	33.000000	0	0	0	1	95	13043.9241	0	0	0	
299	36.000000	0	0	0	1	91	13121.7025	0	0	0	

	Age	Job Type	Marital Status	Education	Metro City	Signed in since(Days)	Purchase made	Retired	Student	Unemployed	N
300	35.000000	0	0	0	1	93	13261.8256	0	0	0	
301	39.000000	0	0	0	1	97	13275.6484	0	0	0	
302	37.000000	0	0	0	1	96	13291.7841	0	0	0	
303	48.000000	0	1	0	1	100	13411.9561	0	0	0	
304	45.000000	0	1	0	1	93	13500.1161	0	0	0	
305	40.000000	0	0	0	1	99	13514.0625	0	0	0	
306	32.000000	0	0	0	1	93	13560.6025	0	0	0	
307	41.000000	0	0	0	1	100	13691.3401	0	0	0	
308	45.000000	0	1	0	1	95	13691.3401	0	0	0	
309	38.000000	0	0	0	1	95	13707.7264	0	0	0	
310	35.000000	0	0	0	1	94	14042.2500	0	0	0	
311	47.000000	0	0	0	1	98	14232.4900	0	0	0	
312	32.000000	0	1	0	1	96	14390.4016	0	0	0	
313	42.000000	0	1	0	1	94	14395.2004	0	0	0	
314	33.000000	0	0	0	1	95	14400.0000	0	0	0	
315	40.000000	0	0	0	1	100	14604.7225	0	0	0	
316	36.000000	0	1	0	1	97	14718.5424	0	0	0	
317	48.000000	0	0	0	1	97	15160.9969	0	0	0	
318	45.000000	0	1	0	1	97	15212.7556	0	0	0	
319	44.000000	0	0	0	1	98	15331.3924	0	0	0	
320	38.000000	0	0	0	1	100	15645.0064	0	0	0	
321	44.000000	0	0	0	1	99	16184.9284	0	0	0	
322	41.000000	0	0	0	1	98	16187.4729	0	0	0	
323	41.000000	0	0	0	1	99	16993.7296	0	0	0	
324	39.000000	0	1	0	1	99	17355.4276	0	0	0	

325 rows × 15 columns



```
In [19]: data=data.drop(['Job Type','Marital Status','Education','Metro City'],axis=1)
```

In [20]: data

Out[20]:

	Age	Signed in since(Days)	Purchasemade	Retired	Student	Unemployed	No	Yes	Primary	Secondry	Yes
0	28.000000	32	261.4689	0	0	1	1	0	0	0	0
1	57.000000	31	273.9025	1	0	0	0	1	0	0	1
2	27.000000	31	275.8921	0	1	0	0	1	0	0	1
3	23.000000	34	295.1524	0	0	1	0	1	0	1	1
4	23.000000	42	325.0809	0	0	0	0	1	0	0	1
5	58.000000	31	327.6100	1	0	0	0	1	0	0	1
6	26.000000	40	338.9281	0	0	0	1	0	0	0	0
7	38.349359	42	355.3225	0	0	1	1	0	0	0	0
8	56.000000	31	362.9025	1	0	0	1	0	0	0	0
9	23.000000	31	364.8100	0	0	1	0	1	0	1	1
10	57.000000	37	446.4769	1	0	0	0	1	0	1	1
11	29.000000	45	475.6761	0	0	1	0	1	0	1	1
12	59.000000	31	484.0000	1	0	0	1	0	0	0	0
13	38.349359	41	512.1169	0	0	1	0	1	0	0	1
14	61.000000	39	554.1316	1	0	0	0	1	0	0	1
15	60.000000	31	576.0000	1	0	0	1	0	0	0	0
16	58.000000	35	577.9216	1	0	0	0	1	0	0	1
17	65.000000	34	577.9216	1	0	0	0	1	0	0	1
18	59.000000	40	590.4900	1	0	0	1	0	0	0	0
19	55.000000	52	1313.3376	0	0	0	0	1	0	1	1
20	46.000000	42	1327.8736	1	0	0	0	1	0	0	1
21	30.000000	33	1343.9556	0	0	0	0	1	0	0	1
22	53.000000	31	1369.0000	0	0	0	1	0	0	0	0
23	34.000000	31	1369.0000	0	0	0	1	0	0	0	0
24	51.000000	34	1397.2644	0	0	0	1	0	0	0	1
25	34.000000	31	1410.0025	0	0	0	0	1	0	0	1
26	46.000000	44	2217.4681	0	0	0	0	1	0	0	1
27	29.000000	50	2226.8961	0	0	0	0	1	0	1	1
28	26.000000	64	2244.8644	0	0	1	0	1	0	0	1
29	30.000000	58	2270.5225	0	0	0	1	0	0	0	0
...	...	...	...	...	...	...	...	...	...	...	...
295	36.000000	98	12989.1609	0	0	0	0	1	0	0	1
296	48.000000	94	13011.9649	0	0	0	0	1	0	0	1
297	41.000000	98	13039.3561	0	0	0	0	1	0	0	1
298	33.000000	95	13043.9241	0	0	0	0	1	0	0	1
299	36.000000	91	13121.7025	0	0	0	0	1	0	0	1

	Age	Signed in since(Days)	Purchasemade	Retired	Student	Unemployed	No	Yes	Primary	Secondry	Yes
300	35.000000	93	13261.8256	0	0	0	0	1	0	0	1
301	39.000000	97	13275.6484	0	0	0	0	1	0	0	1
302	37.000000	96	13291.7841	0	0	0	0	1	0	0	1
303	48.000000	100	13411.9561	0	0	0	1	0	0	0	1
304	45.000000	93	13500.1161	0	0	0	1	0	0	0	1
305	40.000000	99	13514.0625	0	0	0	0	1	0	0	1
306	32.000000	93	13560.6025	0	0	0	0	1	0	0	1
307	41.000000	100	13691.3401	0	0	0	0	1	0	0	1
308	45.000000	95	13691.3401	0	0	0	1	0	0	0	1
309	38.000000	95	13707.7264	0	0	0	0	1	0	0	1
310	35.000000	94	14042.2500	0	0	0	0	1	0	0	1
311	47.000000	98	14232.4900	0	0	0	0	1	0	0	1
312	32.000000	96	14390.4016	0	0	0	1	0	0	0	1
313	42.000000	94	14395.2004	0	0	0	1	0	0	0	1
314	33.000000	95	14400.0000	0	0	0	0	1	0	0	1
315	40.000000	100	14604.7225	0	0	0	0	1	0	0	1
316	36.000000	97	14718.5424	0	0	0	1	0	0	0	1
317	48.000000	97	15160.9969	0	0	0	0	1	0	0	1
318	45.000000	97	15212.7556	0	0	0	1	0	0	0	1
319	44.000000	98	15331.3924	0	0	0	0	1	0	0	1
320	38.000000	100	15645.0064	0	0	0	0	1	0	0	1
321	44.000000	99	16184.9284	0	0	0	0	1	0	0	1
322	41.000000	98	16187.4729	0	0	0	0	1	0	0	1
323	41.000000	99	16993.7296	0	0	0	0	1	0	0	1
324	39.000000	99	17355.4276	0	0	0	1	0	0	0	1

325 rows × 11 columns

In [21]: data.head()

Out[21]:

	Age	Signed in since(Days)	Purchasemade	Retired	Student	Unemployed	No	Yes	Primary	Secondry	Yes
0	28.0	32	261.4689	0	0	1	1	0	0	0	0
1	57.0	31	273.9025	1	0	0	0	1	0	0	1
2	27.0	31	275.8921	0	1	0	0	1	0	0	1
3	23.0	34	295.1524	0	0	1	0	1	0	1	1
4	23.0	42	325.0809	0	0	0	0	1	0	0	1



```
In [22]: data.tail()
```

```
Out[22]:
```

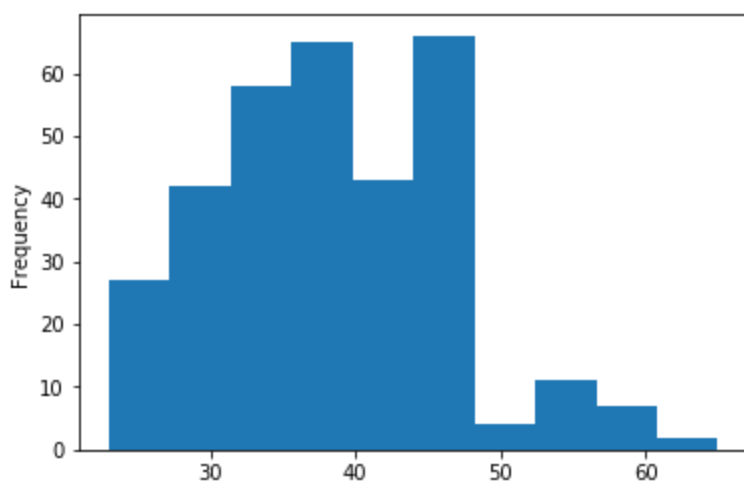
	Age	Signed in since(Days)	Purchasemade	Retired	Student	Unemployed	No	Yes	Primary	Secondry	Yes
320	38.0	100	15645.0064	0	0	0	0	1	0	0	1
321	44.0	99	16184.9284	0	0	0	0	1	0	0	1
322	41.0	98	16187.4729	0	0	0	0	1	0	0	1
323	41.0	99	16993.7296	0	0	0	0	1	0	0	1
324	39.0	99	17355.4276	0	0	0	1	0	0	0	1

```
In [23]: data.isnull().sum()
```

```
Out[23]: Age                                0
Signed in since(Days)                       0
Purchasemade                               0
Retired                                     0
Student                                    0
Unemployed                                 0
No                                          0
Yes                                         0
Primary                                    0
Secondry                                  0
Yes                                         0
dtype: int64
```

```
In [24]: data['Age'].plot.hist()
```

```
Out[24]: <matplotlib.axes._subplots.AxesSubplot at 0xafe3198>
```

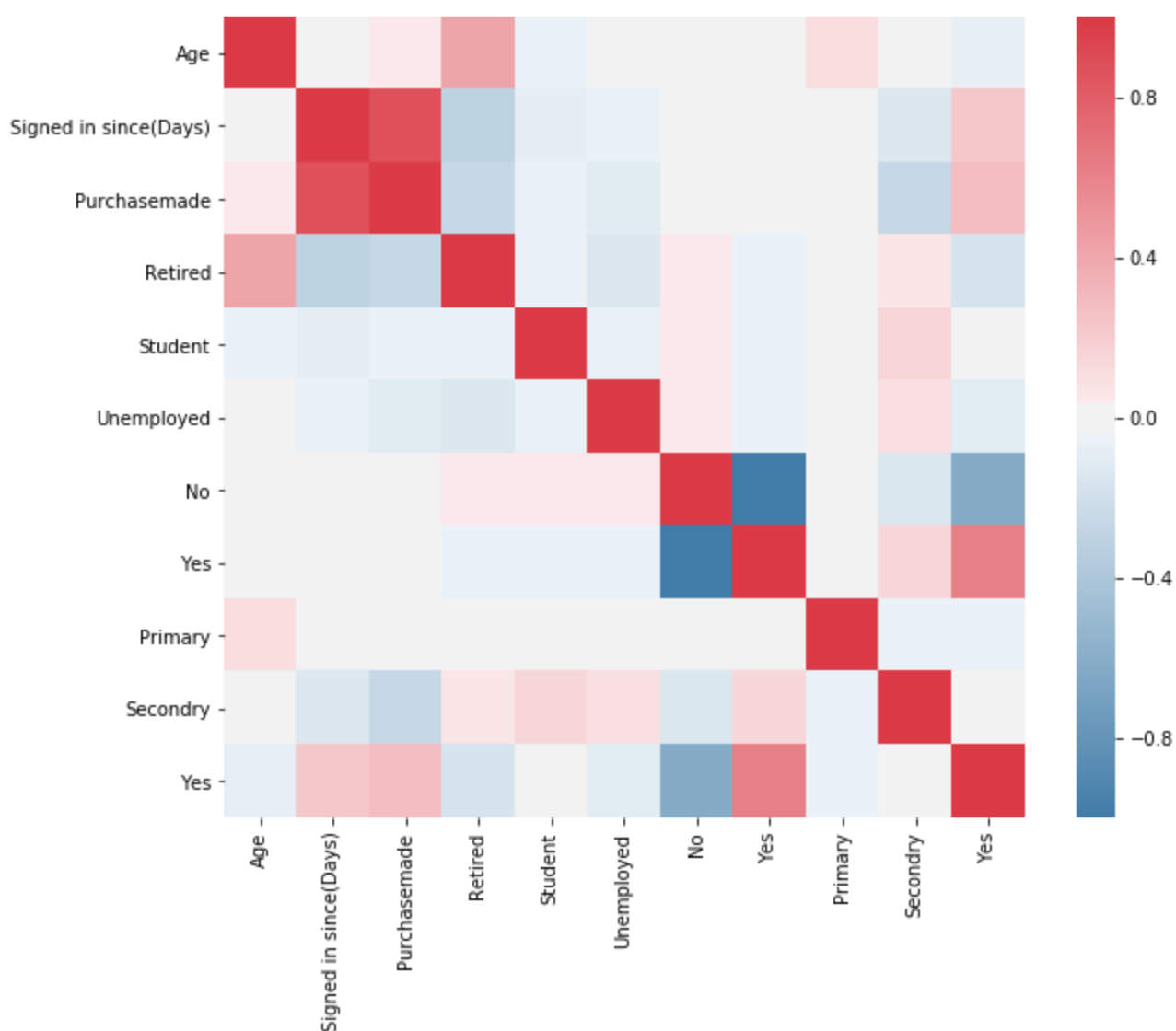


```
In [25]: data.corr()['Purchasemade'].sort_values()
```

```
Out[25]: Secondary          -0.262507  
Retired                    -0.262287  
Unemployed                 -0.112446  
Student                   -0.051403  
Primary                   -0.017975  
Yes                       -0.009934  
No                        0.009934  
Age                       0.054033  
Yes                       0.285400  
Signed in since(Days)     0.877485  
Purchasemade              1.000000  
Name: Purchasemade, dtype: float64
```

```
In [26]: f, ax=plt.subplots(figsize=(10,8))  
corr=data.corr()  
sns.heatmap(corr,mask=np.zeros_like(corr,dtype=np.bool),cmap=sns.diverging_palette(240,10,
```

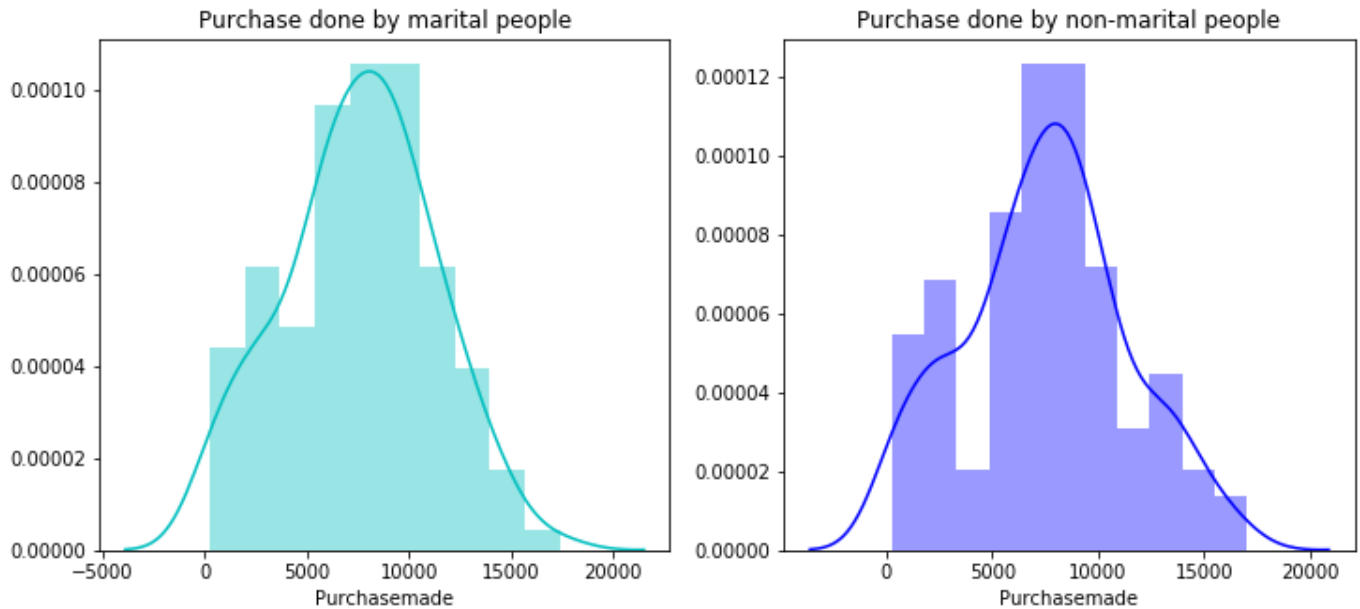
```
Out[26]: <matplotlib.axes._subplots.AxesSubplot at 0xaf7358>
```



```
In [27]: f=plt.figure(figsize=(12,5))

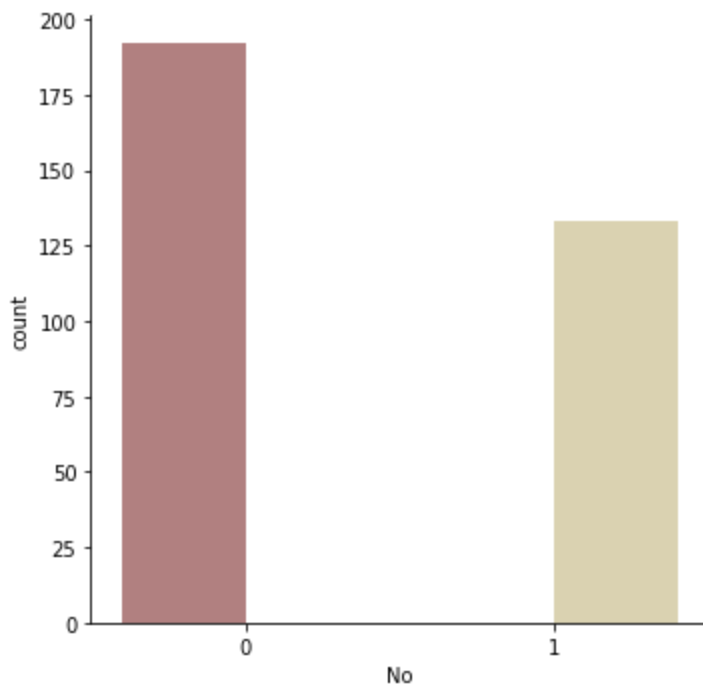
ax=f.add_subplot(121)
sns.distplot(data[(data.No==1)][ 'Purchasemade'],color='c',ax=ax)
ax.set_title('Purchase done by marital people')
ax=f.add_subplot(122)
sns.distplot(data[(data.No==0)][ 'Purchasemade'],color='b',ax=ax)
ax.set_title('Purchase done by non-marital people')
```

Out[27]: Text(0.5, 1.0, 'Purchase done by non-marital people')



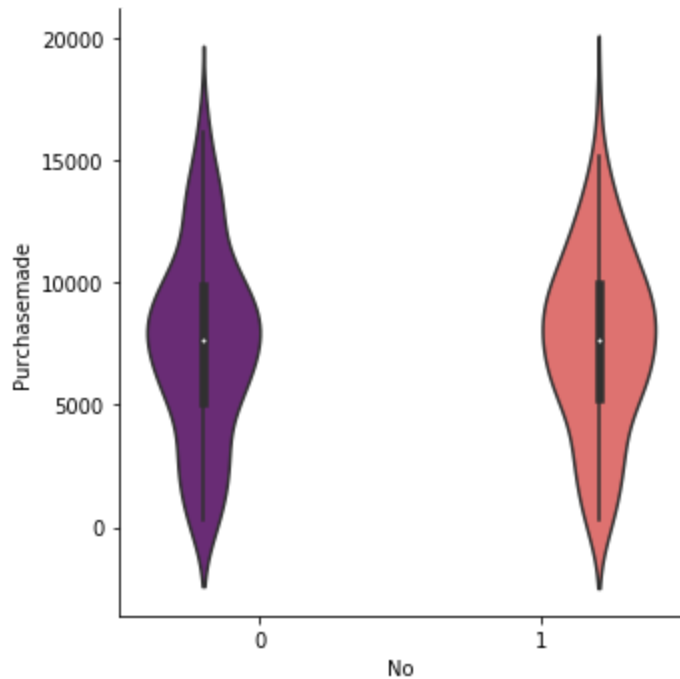
```
In [28]: sns.catplot(x="No",kind="count",hue="No",palette="pink",data=data)"""
```

Out[28]: <seaborn.axisgrid.FacetGrid at 0xb5adcc0>



```
In [29]: sns.catplot(x="No",y="Purchasemade",hue="No",kind="violin",data=data,palette='magma')
```

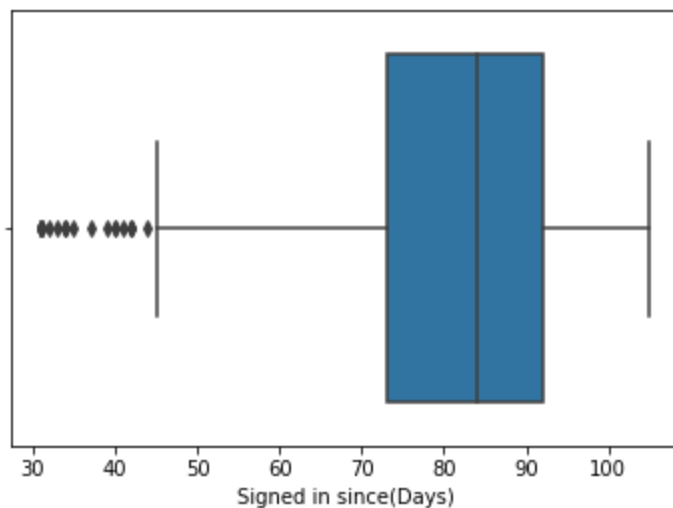
```
Out[29]: <seaborn.axisgrid.FacetGrid at 0xa7470>
```



```
In [30]: data['Signed in since(Days)']=np.where(data["Signed in since(Days)"]>95,105,data['Signed i
```

```
In [31]: sns.boxplot(data['Signed in since(Days)'])
```

```
Out[31]: <matplotlib.axes._subplots.AxesSubplot at 0xbaafd30>
```



```
In [32]: from sklearn.linear_model import LinearRegression
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import PolynomialFeatures
from sklearn.metrics import r2_score,mean_squared_error
```

```
In [33]: x=data.drop(['Purchasemade'],axis=1)
```

```
In [34]: y=data.Purchasemade
```

```
In [35]: x_train,x_test,y_train,y_test=train_test_split(x,y,random_state=0)
```

```
In [36]: lr=LinearRegression().fit(x_train,y_train)
```

```
In [37]: y_train_pred=lr.predict(x_train)
```

```
In [38]: y_test_pred = lr.predict(x_test)
```

```
In [39]: print(lr.score(x_test,y_test))
```

```
0.8120949635915722
```

```
In [40]: x=data.drop(['Purchased', 'Unemployed', 'Retired', 'Student'],axis=1)
```

```
In [41]: y=data.Purchased
```

```
In [42]: quad = PolynomialFeatures (degree = 2)
```

```
In [43]: x_quad=quad.fit_transform(x)
```

```
In [44]: x_train,x_test,y_train,y_test=train_test_split(x_quad,y,random_state=0)
```

```
In [45]: lr=LinearRegression().fit(x_train,y_train)
```

```
In [46]: y_train_pred=lr.predict(x_train)
```

```
In [47]: y_test_pred=lr.predict(x_test)
```

```
In [48]: print(lr.score(x_test,y_test))
```

```
0.8165784033755232
```

```
In [49]: y_test_pred
```

```
Out[49]: array([ 8709.66683968,  4158.51697769,   209.13499944,  7102.3307169 ,
  9692.49644887,  8254.72892789,  6909.27766953, 10834.34109266,
  3531.54596858,  7756.82111846,  6077.90962894,  6247.93391988,
  7693.93890166, 10816.9287749 ,  8380.63547179,  6338.51385171,
  9892.77181501, 10567.59877231,  3735.09231339, 10429.30549678,
 10147.94619   ,  8445.8835602 , 10136.92765088,  1776.06124019,
  7376.81189655,  6148.04092298, 13961.89478139,  8392.65696694,
  8924.15341375, 10263.4205839 ,  7542.31378641, 12378.57466816,
  2036.21715966,   967.95526418, 13362.38259027,  4875.18929389,
   895.8762512 ,  8507.74690377,  -299.63423092,  6968.79974003,
  5016.72851304,   327.93462056,  7857.49367533,  4748.20980416,
  4738.77720228,  6334.19963203,  8053.89687875,  -886.68259572,
  5968.05288829,  7310.17243424,  6894.96394123,  8429.46331932,
  7028.14252098,  9962.85889458,  5784.02096644,  5574.2937538 ,
  7960.4331085 ,  8438.43556854,  9958.25477689,  6662.05955279,
  2754.85379894,  9940.35352314,  7861.5614179 ,  4128.48367809,
 13833.94575372,  -327.4183685 , 10170.13126346,   642.94602788,
 10198.33692328, 10215.25763415,  7039.84374127,  9517.8507575 ,
 10069.94403369,  9016.50801865,  -417.3915152 ,  4225.96720322,
  8841.5103410 ,  8887.68070662,  7351.66777035,  7825.88070654]
```