4) What kind of relation between salary and mba_p

In [1]:
```
import pandas as pd
import numpy as np
```

In [2]:
```
dataset=pd.read_csv("Preplacementdata.csv")
dataset
```

Out[2]:

	sl_no	ssc_p	hsc_p	degree_p	etest_p	mba_p	salary	gender	ssc_b	hsc_b	hsc_s	degree_t	workex	special
0	1.0	67.00	91.00	58.00	55.0	58.80	270000.0	M	Others	Others	Commerce	Sci&Tech	No	M
1	2.0	79.33	78.33	77.48	86.5	66.28	200000.0	M	Central	Others	Science	Sci&Tech	Yes	M
2	3.0	65.00	68.00	64.00	75.0	57.80	250000.0	M	Central	Central	Arts	Comm&Mgmt	No	M
3	4.0	56.00	52.00	52.00	66.0	59.43	265000.0	M	Central	Central	Science	Sci&Tech	No	M
4	5.0	85.80	73.60	73.30	96.8	55.50	425000.0	M	Central	Central	Commerce	Comm&Mgmt	No	M
...	...	...	...	...	...	...	...	...	...	...	...	...	...	
210	211.0	80.60	82.00	77.60	91.0	74.49	400000.0	M	Others	Others	Commerce	Comm&Mgmt	No	M
211	212.0	58.00	60.00	72.00	74.0	53.62	275000.0	M	Others	Others	Science	Sci&Tech	No	M
212	213.0	67.00	67.00	73.00	59.0	69.72	295000.0	M	Others	Others	Commerce	Comm&Mgmt	Yes	M
213	214.0	74.00	66.00	58.00	70.0	60.23	204000.0	F	Others	Others	Commerce	Comm&Mgmt	No	M
214	215.0	62.00	58.00	53.00	89.0	60.22	265000.0	M	Central	Others	Science	Comm&Mgmt	No	M

215 rows × 15 columns

In [3]:
```
dataset.isnull().sum()
```

Out[3]:
```
sl_no             0
ssc_p             0
hsc_p             0
degree_p          0
etest_p           0
mba_p             0
salary            0
gender            0
ssc_b             0
hsc_b             0
hsc_s             0
degree_t          0
workex            0
specialisation    0
status            0
dtype: int64
```

In [4]:
```
dataset.drop('sl_no',inplace=True,axis=1)
```

In [5]:
```
dataset
```

Out[5]:

	ssc_p	hsc_p	degree_p	etest_p	mba_p	salary	gender	ssc_b	hsc_b	hsc_s	degree_t	workex	specialisation
0	67.00	91.00	58.00	55.0	58.80	270000.0	M	Others	Others	Commerce	Sci&Tech	No	Mkt&HR
1	79.33	78.33	77.48	86.5	66.28	200000.0	M	Central	Others	Science	Sci&Tech	Yes	Mkt&Fin
2	65.00	68.00	64.00	75.0	57.80	250000.0	M	Central	Central	Arts	Comm&Mgmt	No	Mkt&Fin
3	56.00	52.00	52.00	66.0	59.43	265000.0	M	Central	Central	Science	Sci&Tech	No	Mkt&HR
4	85.80	73.60	73.30	96.8	55.50	425000.0	M	Central	Central	Commerce	Comm&Mgmt	No	Mkt&Fin
...	...	...	...	...	...	...	...	...	...	...	...	...	...
210	80.60	82.00	77.60	91.0	74.49	400000.0	M	Others	Others	Commerce	Comm&Mgmt	No	Mkt&Fin
211	58.00	60.00	72.00	74.0	53.62	275000.0	M	Others	Others	Science	Sci&Tech	No	Mkt&Fin
212	67.00	67.00	73.00	59.0	69.72	295000.0	M	Others	Others	Commerce	Comm&Mgmt	Yes	Mkt&Fin
213	74.00	66.00	58.00	70.0	60.23	204000.0	F	Others	Others	Commerce	Comm&Mgmt	No	Mkt&HR
214	62.00	58.00	53.00	89.0	60.22	265000.0	M	Central	Others	Science	Comm&Mgmt	No	Mkt&HR

215 rows × 14 columns

In [6]:
```
dataset.corr(numeric_only=True)
```

Out[6]:

	ssc_p	hsc_p	degree_p	etest_p	mba_p	salary
ssc_p	1.000000	0.511472	0.538404	0.261993	0.388478	0.108669
hsc_p	0.511472	1.000000	0.434206	0.245113	0.354823	0.122921
degree_p	0.538404	0.434206	1.000000	0.224470	0.402364	0.053352
etest_p	0.261993	0.245113	0.224470	1.000000	0.218055	0.169233
mba_p	0.388478	0.354823	0.402364	0.218055	1.000000	0.155673
salary	0.108669	0.122921	0.053352	0.169233	0.155673	1.000000

The MBA pass mark and salary of the student are only 15% correlated.Because MBA pass marks only 100 and salry is lakhs.

In [ ]:

In [ ]:

In [ ]: