

5. Which Specialization is getting a minimum salary?

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

```
In [3]: dataset=pd.read_csv("Preplacementdata.csv")
```

```
In [4]: dataset
```

```
Out[4]:
```

	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 [5]: dataset.isnull().sum()
```

```
Out[5]: 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 [6]: average_salary_by_specialisation=dataset.groupby("specialisation")["salary"].mean()
```

```
In [7]: average_salary_by_specialisation
```

```
Out[7]: specialisation
Mkt&Fin      291800.0
Mkt&HR       268000.0
Name: salary, dtype: float64
```

```
In [13]: min_salary_specialization=average_salary_by_specialisation.idxmin()
```

```
In [14]: min_salary_specialization
```

```
Out[14]: 'Mkt&HR'
```

```
In [15]: print(f"The specialization with the minimum averagae salary is: {min_salary_specialization}")
```

The specialization with the minimum averagae salary is: Mkt&HR

Mkt and HR specalization is getting a minimum salary.

6. How many of them are getting above 500000 salaries?

```
In [29]: dataset=pd.read_csv("Preplacementdata.csv")
```

```
In [30]: df=dataset["salary"]
```

```
In [31]: df
```

```
Out[31]: 0      270000.0
         1      200000.0
         2      250000.0
         3      265000.0
         4      425000.0
         ...
        210     400000.0
        211     275000.0
        212     295000.0
        213     204000.0
        214     265000.0
        Name: salary, Length: 215, dtype: float64
```

```
In [32]: df.max()
```

```
Out[32]: 940000.0
```

```
In [33]: df[df>500000].count()
```

```
Out[33]: 3
```

Three students are getting the above 5,00,000 salaries.It is from the preprocessed dataset.

```
In [ ]:
```

```
In [ ]:
```

```
In [ ]:
```

Loading [MathJax]/jax/output/CommonHTML/fonts/TeX/fontdata.js