23-10-2025

On insurance data

1)Each region wise totalexpenses

2)Each region,each children and smoker class wise total expenses

3)Gender wise avg bmi and expenses

1)inu = pd.read_csv(r"D:\MLP\insurance.csv")

inu

	age	sex	bmi	children	smoker	region	expenses
0	19	female	27.9	0	yes	southwest	16884.92
1	18	male	33.8	1	no	southeast	1725.55
2	28	male	33.0	3	no	southeast	4449.46
3	33	male	22.7	0	no	northwest	21984.47
4	32	male	28.9	0	no	northwest	3866.86
...	...	...	...	...	...	...	...
1333	50	male	31.0	3	no	northwest	10600.55
1334	18	female	31.9	0	no	northeast	2205.98
1335	18	female	36.9	0	no	southeast	1629.83

1336	21	female	25.8	0	no	southwest	2007.95
1337	61	female	29.1	0	yes	northwest	29141.36

2)inu.groupby(by='region')['expenses'].sum()

```
region
northeast    4343668.64
northwest    4035711.93
southeast    5363689.80
southwest    4012754.82
Name: expenses, dtype: float64
```

3) inu['region'].value_counts()

```
region
southeast    364
southwest    325
northwest    325
northeast    324
Name: count, dtype: int64
```

4)inu.groupby(by='region')['expenses'].mean()

```
region
northeast    13406.384691
northwest    12417.575169
southeast    14735.411538
southwest    12346.937908
Name: expenses, dtype: float64
```

5) inu.groupby(by='sex')['expenses'].mean()

```
ex
female    12569.578897
male      13956.751420
Name: expenses, dtype: float64
```