▾ To import **libraries**

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
```

```
df=pd.read_csv("/content/sample_data/pr 6-Dataset.csv")
```

```
df
```

	Enroll	Name	TOC	DA&V	WP	C&NS	AJP	Address	SPI	Branch
0	1.910000e+11	AWTANI MANISH DINESH	24	19	24	22	15	Surat	6.93	CSE
1	1.910000e+11	BHATIYA PARTH NILESHKUMAR	21	18	20	19	15	Mandvi	6.20	CSE
2	1.910000e+11	BHAVSAR YASH ALPESHKUMAR	23	20	22	26	17	Vyara	7.20	CSE
3	1.910000e+11	CHHATRANI DIVYESHKUMAR ROSHANKUMAR	26	19	26	24	16	Surat	7.40	CSE
4	1.910000e+11	DARJI VEDANT MUKESHBHAI	21	14	13	7	10	Navsari	4.33	CSE
5	1.910000e+11	DESAI PRIT AMISHKUMAR	27	21	27	21	18	Surat	7.60	CSE
6	1.910000e+11	GAMIT SUNIL BALUBHAI	26	15	19	14	15	Bardoli	5.93	CSE
7	1.910000e+11	GUPTA NIDHI HARISHCHANDRA	23	20	19	14	15	Bardoli	6.07	CSE
8	1.910000e+11	JARIWALA YASH HIRENBHAI	27	25	28	28	18	Mandvi	8.40	CSE
9	1.910000e+11	KATARIA SHIV SHASHIBHAI	23	24	21	19	15	Navsari	6.80	CSE
10	1.910000e+11	KHARVASIYA VISHAKHABEN ASHVINSINH	24	23	24	25	17	Surat	7.53	Elec. Engg.
11	1.910000e+11	KOSAMBIYA RUCHITKUMAR RAMESHBHAI	22	20	24	23	17	Bardoli	7.07	Elec. Engg.
12	1.910000e+11	LAD ISHA NILESH	13	4	8	8	13	Vyara	3.07	Elec. Engg.
13	1.910000e+11	MARVADI YASHLAL DIPAKBHAI	17	2	18	12	15	Navsari	4.27	Elec. Engg.
14	1.910000e+11	MISTRY DINKEL HEMANT	19	16	22	22	16	Bardoli	6.33	Elec. Engg.
15	1.910000e+11	MISTRY SMIT JAYESHKUMAR	15	10	20	15	15	Bardoli	5.00	Elec.

15	1.910000e+11	MISTRY SMIT JAYESHKUMAR	18	18	20	15	15	Bardoli	5.00	Engg.
16	1.910000e+11	PANDYA MAHIMNA DHARMESHBHAI	8	6	12	2	13	Surat	2.73	Elec. Engg.
17	1.910000e+11	PARMAR ARYAN PRAKASHBHAI	12	11	4	15	11	Vyara	3.53	Elec. Engg.
18	1.910000e+11	PATEL HITESH PRAKASHBHAI	28	23	20	25	17	Bardoli	7.53	Elec. Engg.
19	1.910000e+11	PATEL RUSHIL KAUSHIKKUMAR	21	12	20	16	13	Navsari	5.47	Elec. Engg.
20	1.910000e+11	Adeline Jackson	22	25	26	24	16	Bardoli	7.53	Civil Engg.
21	1.910000e+11	Anthony Barclay	26	23	21	23	17	Bardoli	7.33	Civil Engg.
22	1.910000e+11	Bryce Jones	15	13	15	8	15	Surat	4.40	Civil Engg.
23	1.910000e+11	Cara Wilton	27	28	20	20	15	Bardoli	7.33	Civil Engg.
24	1.910000e+11	Cherish Rodgers	26	11	13	19	16	Vyara	5.67	Civil Engg.
25	1.910000e+11	Clint Hogg	12	4	10	7	15	Bardoli	3.20	Civil Engg.
26	1.910000e+11	Dasha Driscoll	26	25	23	22	16	Navsari	7.47	Civil Engg.
27	1.910000e+11	Emmanuelle Yoman	18	18	17	15	15	Bardoli	5.53	Civil Engg.
28	1.910000e+11	Fred Addley	22	23	23	24	17	Bardoli	7.27	Civil

										Engg.
29	1.910000e+11	Jacob Donnelly	23	17	14	25	16	Vyara	6.33	Civil Engg.
30	1.910000e+11	Joseph Clark	16	5	12	12	15	Surat	4.00	Mech. Engg.
31	1.910000e+11	Julian Forester	18	5	9	9	12	Bardoli	3.53	Mech. Engg.

▾ Region Wise Results comparison

| **33** | 1.910000e+11 | Kurt Rigg | 12 | 10 | 5 | 4 | 14 | Mandvi | 3.00 | Engg. |

```
mean1=list(df.groupby(['Address'])['SPI'].mean())
```

| **34** | 1.910000e+11 | Mason Rixon | 30 | 20 | 23 | 28 | 18 | Bardoli | 6.33 | Engg. |

```
mean1
```

```
[6.226363636363636,
 6.467499999999999,
 5.771111111111111,
 5.429090909090909,
 6.177777777777778]
```
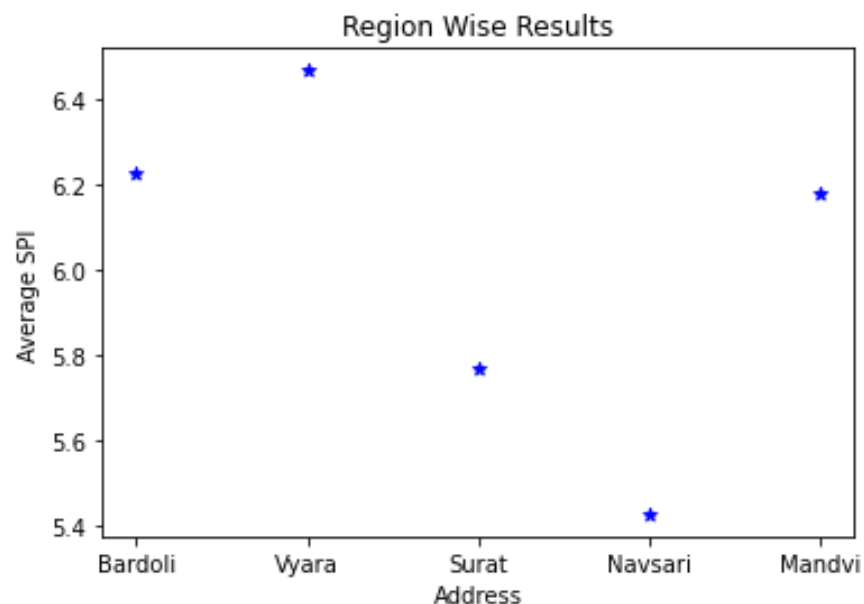
```
df['Address'].value_counts()
```

```
Bardoli    22
Surat      11
Vyara       9
Navsari     9
Mandvi      4
Name: Address, dtype: int64
```

```
cities=['Bardoli','Vyara','Surat','Navsari','Mandvi']
plt.scatter(cities, mean1, color='b',marker='*')
plt.title("Region Wise Results")
plt.xlabel("Address")
plt.ylabel("Average SPI")
```

Text(0, 0.5, 'Average SPI')


```
mtoc=df['TOC'].mean()
mdav=df['DA&V'].mean()
mwp=df['WP'].mean()
majp=df['AJP'].mean()
mcns=df['C&NS'].mean()
```

```python
mean1=[]
mean1.append(mtoc)
mean1.append(mdav)
mean1.append(mwp)
mean1.append(majp)
mean1.append(mcns)
mean1
```

```
[21.854545454545455,
 16.21818181818182,
 18.6,
 15.327272727272728,
 18.036363636363635]
```

```python
Subjects = ['TOC','DA&V','WP','C&NS','AJP']
plt.scatter(Subjects, mean1, color='b', marker='*')
plt.title("Subject Wise Results")
plt.xlabel("Subjects")
plt.ylabel("Average Score")
```

```
Text(0, 0.5, 'Average Score')
```

Subject Wise Results
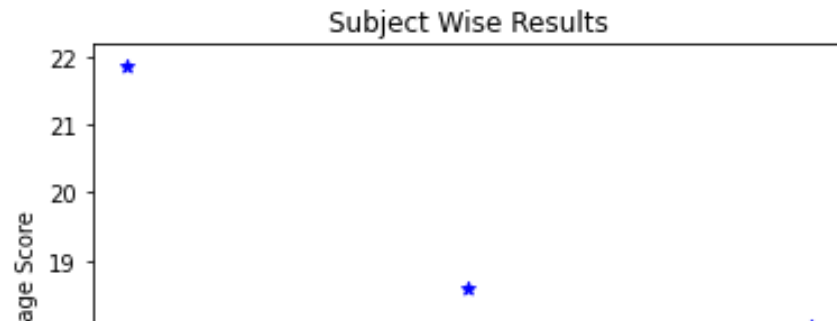

```
mean3 = list(df.groupby(['Branch'])['SPI'].mean())
print(mean3)
```

```
[6.686, 6.436000000000001, 6.206, 5.253, 5.212000000000001]
```

```
df['Branch'].value_counts()
```

```
Chem. Engg.    15
CSE            10
Elec. Engg.    10
Civil Engg.    10
Mech. Engg.    10
Name: Branch, dtype: int64
```

```
branches = ['Chem. Engg.','CSE','Elec. Engg.','Civil Engg.','Mech. Engg.']
plt.scatter(branches, mean3, color='b',marker='*')
plt.title("Branch Wise Results")
plt.xlabel("Branch")
plt.ylabel("Average SPI")
```

Text(0, 0.5, 'Average SPI')