Q.11) Take the heart disease dataset & calculate the mean, variance, kurtosis, and skewness. Plot the graph of mean, variance, kurtosis, and skewness for the dataset.

#heart stats_mvks
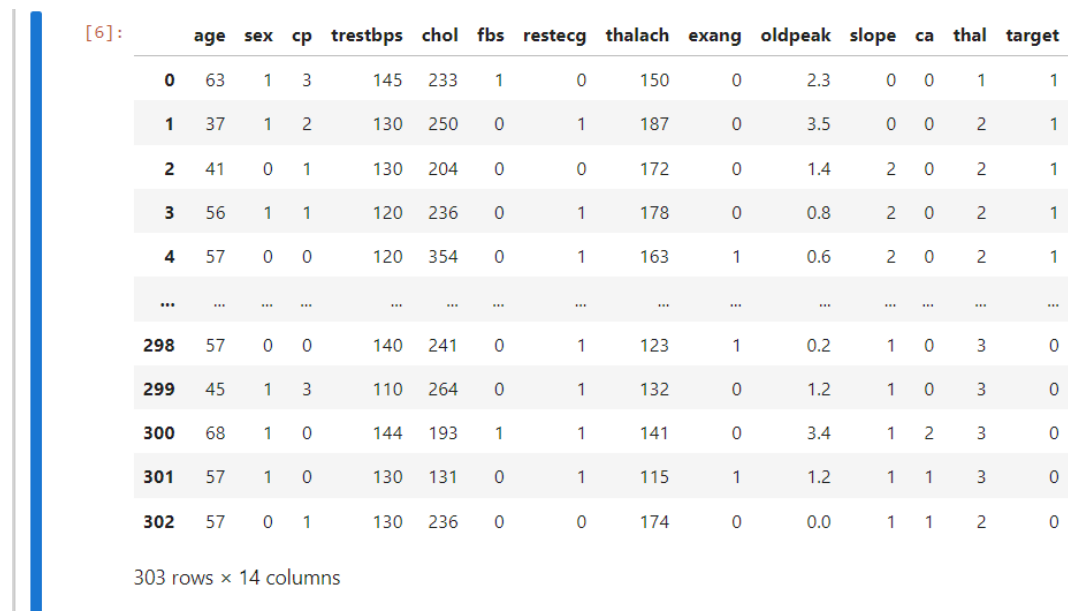
import pandas as pd

import numpy as np

import scipy.stats as stats

import matplotlib.pyplot as plt

#load the dataset

data = pd.read_csv('C:/Users/HP/OneDrive/Desktop/ml 7th sem codes/datasets/heart.csv')

data

#data.head()

[6]:
	age	sex	cp	trestbps	chol	fbs	restecg	thalach	exang	oldpeak	slope	ca	thal	target
0	63	1	3	145	233	1	0	150	0	2.3	0	0	1	1
1	37	1	2	130	250	0	1	187	0	3.5	0	0	2	1
2	41	0	1	130	204	0	0	172	0	1.4	2	0	2	1
3	56	1	1	120	236	0	1	178	0	0.8	2	0	2	1
4	57	0	0	120	354	0	1	163	1	0.6	2	0	2	1
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
298	57	0	0	140	241	0	1	123	1	0.2	1	0	3	0
299	45	1	3	110	264	0	1	132	0	1.2	1	0	3	0
300	68	1	0	144	193	1	1	141	0	3.4	1	2	3	0
301	57	1	0	130	131	0	1	115	1	1.2	1	1	3	0
302	57	0	1	130	236	0	0	174	0	0.0	1	1	2	0

303 rows × 14 columns

#drop 'target' column as it is class label & we don't need it in this case

data = data.drop('target', axis=1)   # axis=0 is default'=> drop rows ; axis=1 => drop cols

data

	age	sex	cp	trestbps	chol	fbs	restecg	thalach	exang	oldpeak	slope	ca	thal
0	63	1	3	145	233	1	0	150	0	2.3	0	0	1
1	37	1	2	130	250	0	1	187	0	3.5	0	0	2
2	41	0	1	130	204	0	0	172	0	1.4	2	0	2
3	56	1	1	120	236	0	1	178	0	0.8	2	0	2
4	57	0	0	120	354	0	1	163	1	0.6	2	0	2
...	...	...	...	...	...	...	...	...	...	...	...	...	...
298	57	0	0	140	241	0	1	123	1	0.2	1	0	3
299	45	1	3	110	264	0	1	132	0	1.2	1	0	3
300	68	1	0	144	193	1	1	141	0	3.4	1	2	3
301	57	1	0	130	131	0	1	115	1	1.2	1	1	3
302	57	0	1	130	236	0	0	174	0	0.0	1	1	2

303 rows × 13 columns

```
# Calculate mean, variance, kurtosis, and skewness for the entire dataset
mean = data.mean()  #colwise mean
variance = np.var(data)
kurtosis = stats.kurtosis(data)
skewness = stats.skew(data)


# Print the results for each cols
print("Mean:")
print(mean)
print("\nVariance:")
print(variance)
print("\nKurtosis:")
print(kurtosis)
print("\nSkewness:")
print(skewness)


#print the overall results
print("\nOverall Mean:")
print(np.mean(mean))
print("\nOverall Variance:")
```

```python
print(np.var(variance))

print("\nOverall Kurtosis:")

print(stats.kurtosis(kurtosis))

print("\nOverall Skewness:")

print(stats.skew(skewness))
```

```
Mean:
age            54.366337
sex             0.683168
cp              0.966997
trestbps      131.623762
chol          246.264026
fbs             0.148515
restecg         0.528053
thalach       149.646865
exang           0.326733
oldpeak         1.039604
slope           1.399340
ca              0.729373
thal            2.313531
dtype: float64

Variance:
age            82.212332
sex             0.216449
cp              1.061617
trestbps      306.571317
chol         2677.560653
fbs             0.126458
restecg         0.275616
thalach       522.914899
exang           0.219978
oldpeak         1.343646
slope           0.378481
ca              1.042273
thal            0.373645
dtype: float64
```

```
Kurtosis:
[-0.55299857 -1.37998188 -1.19321094  0.89405696  4.41168212  1.90775194
 -1.36002744 -0.08068888 -1.45409982  1.53019258 -0.63695019  0.80573112
  0.27328459]

Skewness:
[-0.20145969 -0.78741229  0.4823294   0.71023006  1.13773262  1.97680346
  0.16171657 -0.53474554  0.73885058  1.26342552 -0.50579573  1.30392596
 -0.47435893]

Overall Mean:
45.38740797156639

Overall Variance:
503827.1673387734

Overall Kurtosis:
0.7739774040956542

Overall Skewness:
0.195811644687599
```

```python
#plot results

plt.figure(figsize=(12, 8))

plt.plot(mean,label='mean', marker='o')

plt.plot(variance,label='variance', marker='o')

plt.plot(kurtosis,label='kurtosis', marker='s')

plt.plot(skewness,label='skewness', marker='x')

plt.title('plotting of mean,variance,kurtosis,skewness')

plt.legend()

plt.grid(True)

plt.show()
```

plotting of mean,variance,kurtosis,skewness

[ ]: