task2

June 1, 2024

```
import pandas as pd
[2]:
     data = pd.read_csv('C:\\Users\\kumar\\Desktop\\heart.csv')
     type(data)
[4]:
[4]: pandas.core.frame.DataFrame
[5]: data.info
[5]: <bound method DataFrame.info of
                                                age
                                                      sex
                                                                 trtbps
                                                                          chol
                                                                                 fbs restecg
     thalachh
                        oldpeak slp
                 exng
                        3
                                      233
                                                        0
                                                                                   2.3
     0
            63
                               145
                                              1
                                                                  150
                                                                           0
                                                                                           0
     1
            37
                        2
                                                                                   3.5
                   1
                               130
                                      250
                                              0
                                                        1
                                                                  187
                                                                           0
                                                                                           0
     2
                                                                                   1.4
            41
                   0
                        1
                               130
                                      204
                                              0
                                                        0
                                                                  172
                                                                           0
                                                                                           2
     3
            56
                   1
                        1
                               120
                                      236
                                              0
                                                        1
                                                                  178
                                                                           0
                                                                                   0.8
                                                                                           2
     4
                        0
            57
                   0
                               120
                                      354
                                              0
                                                         1
                                                                  163
                                                                           1
                                                                                   0.6
                                                                                           2
                                                        •••
     298
            57
                   0
                        0
                               140
                                      241
                                              0
                                                         1
                                                                  123
                                                                           1
                                                                                   0.2
                                                                                           1
     299
            45
                        3
                               110
                                      264
                                              0
                                                         1
                                                                  132
                                                                           0
                                                                                   1.2
                                                                                            1
     300
                        0
                               144
                                      193
                                                         1
                                                                  141
                                                                           0
                                                                                   3.4
                                                                                            1
            68
                                              1
     301
            57
                   1
                        0
                               130
                                      131
                                              0
                                                        1
                                                                  115
                                                                           1
                                                                                   1.2
                                                                                           1
     302
                                                                  174
            57
                   0
                        1
                               130
                                      236
                                              0
                                                        0
                                                                           0
                                                                                   0.0
                                                                                            1
                 thall
                         output
           caa
     0
                      1
             0
                               1
                     2
                               1
     1
             0
                      2
     2
                               1
             0
     3
                      2
             0
                               1
                      2
     4
             0
                               1
                      3
                               0
     298
             0
     299
             0
                     3
                               0
     300
             2
                     3
                               0
     301
              1
                      3
                               0
                      2
                               0
     302
              1
```

[6]: data.shape [6]: (303, 14)data.describe() [7]: [7]: trtbps chol fbs age sex ср count 303.000000 303.000000 303.000000 303.000000 303.000000 303.000000 131.623762 54.366337 0.683168 0.966997 246.264026 0.148515 mean std 9.082101 0.466011 1.032052 17.538143 51.830751 0.356198 min 29.000000 0.000000 0.00000 94.000000 126.000000 0.000000 25% 47.500000 0.000000 0.00000 120.000000 211.000000 0.000000 50% 55.000000 1.000000 1.000000 130.000000 240.000000 0.000000 75% 61.000000 1.000000 2.000000 140.000000 274.500000 0.000000 3.000000 77.000000 1.000000 200.000000 564.000000 1.000000 maxoldpeak restecg thalachh exng slp caa 303.000000 303.000000 303.000000 303.000000 303.000000 count 303.000000 0.528053 149.646865 0.326733 1.039604 1.399340 0.729373 mean 22.905161 0.469794 1.161075 std 0.525860 0.616226 1.022606 min 0.000000 71.000000 0.00000 0.00000 0.00000 0.00000 133.500000 25% 0.000000 0.00000 0.00000 1.000000 0.00000 50% 1.000000 153.000000 0.000000 0.800000 1.000000 0.00000 75% 1.000000 166.000000 1.000000 1.600000 2.000000 1.000000 2.000000 202.000000 1.000000 6.200000 2.000000 4.000000 maxthall output 303.000000 303.000000 count mean 2.313531 0.544554 std 0.612277 0.498835 min 0.000000 0.000000 25% 2.000000 0.00000 50% 2.000000 1.000000 75% 3.000000 1.000000 3.000000 1.000000 max[8]: data=data.drop_duplicates() data [8]: oldpeak slp \ trtbps chol fbs restecg thalachh exng age sex ср 2.3 0 63 1 3 145 233 1 0 150 0 0 2 1 37 1 0 1 0 3.5 0 130 250 187 2 0 0 2 41 0 1 130 204 172 0 1.4 3 56 1 1 120 236 0 1 178 0 0.8 2 4 57 0 0 120 354 0 1 163 1 0.6 2

• •	•••	• •	•••		•	••	•••	••• •••			
298	57	0	0	140	241	0	1	123	1	0.2	1
299	45	1	3	110	264	0	1	132	0	1.2	1
300	68	1	0	144	193	1	1	141	0	3.4	1
301	57	1	0	130	131	0	1	115	1	1.2	1
302	57	0	1	130	236	0	0	174	0	0.0	1

	caa	thall	output
0	0	1	1
1	0	2	1
2	0	2	1
3	0	2	1
4	0	2	1
		•••	
298	0	3	0
299	0	3	0
300	2	3	0
301	1	3	0
302	1	2	0

[302 rows x 14 columns]

[9]: data.isnull()

[9]:		age	sex	-	trtbps			_	thalachh	_	\
	0	False	False	False	False	False	False	False	False	False	
	1	False	False	False	False	False	False	False	False	False	
	2	False	False	False	False	False	False	False	False	False	
	3	False	False	False	False	False	False	False	False	False	
	4	False	False	False	False	False	False	False	False	False	
		•••		•••		•••	•••				
	298	False	False	False	False	False	False	False	False	False	
	299	False	False	False	False	False	False	False	False	False	
	300	False	False	False	False	False	False	False	False	False	
	301	False	False	False	False	False	False	False	False	False	
	302	False	False	False	False	False	False	False	False	False	
		oldpea	k sl	p caa	thall	outpu	t				
	0	Fals	e Fals	e False	e False	Fals	е				
	1	Fals	e Fals	e False	e False	Fals	е				
	2	Fals	e Fals	e False	e False	Fals	е				
	3	Fals	e Fals	e False	False	Fals	е				
	4	Fals	e Fals	e False	False	Fals	е				
		•••	•••		•••						
	298	Fals	e Fals	e False	False	Fals	е				
	299	Fals	e Fals	e False	False	Fals	е				
	300	Fals	e Fals	e False	e False						

```
[302 rows x 14 columns]
[10]: data.isnull().sum()
[10]: age
                   0
      sex
                   0
      ср
                   0
                   0
      trtbps
      chol
                   0
      fbs
                   0
      restecg
                   0
      thalachh
                   0
                   0
      exng
      oldpeak
                   0
      slp
                   0
                   0
      caa
      thall
                   0
      output
      dtype: int64
[11]: data.notnull()
                                                                        exng oldpeak \
[11]:
                              trtbps
                                       chol
                                               fbs
                                                    restecg
                                                              thalachh
            age
                   sex
                          ср
                                                                  True
                                                                                  True
                                 True
                                       True
                                                                         True
      0
           True
                  True
                        True
                                              True
                                                       True
      1
           True
                 True
                        True
                                       True
                                                                        True
                                                                                  True
                                 True
                                              True
                                                       True
                                                                  True
                                                                                  True
      2
           True
                 True
                        True
                                 True
                                       True
                                             True
                                                       True
                                                                  True
                                                                        True
      3
           True True
                        True
                                 True
                                       True
                                             True
                                                       True
                                                                  True
                                                                        True
                                                                                  True
           True
                 True
                        True
                                 True
                                       True
                                              True
                                                       True
                                                                  True
                                                                        True
                                                                                  True
      298
           True
                 True
                        True
                                 True
                                       True
                                             True
                                                       True
                                                                  True
                                                                        True
                                                                                  True
      299
           True
                 True
                        True
                                 True
                                       True
                                             True
                                                       True
                                                                  True
                                                                        True
                                                                                  True
      300
           True
                  True
                        True
                                       True
                                              True
                                                                        True
                                                                                  True
                                 True
                                                       True
                                                                  True
      301
           True
                  True
                        True
                                 True
                                       True
                                              True
                                                       True
                                                                  True
                                                                         True
                                                                                  True
      302
           True
                  True
                        True
                                 True
                                       True
                                              True
                                                       True
                                                                        True
                                                                                  True
                                                                  True
                                output
            slp
                   caa
                        thall
      0
           True True
                         True
                                  True
      1
           True
                 True
                         True
                                  True
      2
           True
                 True
                         True
                                  True
      3
           True
                 True
                         True
                                  True
      4
           True True
                         True
                                  True
      298
           True True
                         True
                                  True
      299
           True True
                         True
                                  True
```

301

302

False

False

False False

False

False

False

False

False

False

```
True
      302
            True
                   True
                           True
                                    True
      [302 rows x 14 columns]
[12]: data.isnull().sum().sum()
[12]: 0
[13]: data1=data.fillna(value=00)
      data1
[13]:
                            trtbps
                                      chol
                                            fbs
                                                  restecg
                                                            thalachh
                                                                        exng
                                                                               oldpeak
                                                                                         slp
            age
                  sex
                        ср
                         3
                                       233
                                               1
                                                                  150
                                                                           0
                                                                                    2.3
             63
                                145
                                                         0
                         2
      1
             37
                    1
                                130
                                       250
                                               0
                                                         1
                                                                  187
                                                                           0
                                                                                    3.5
                                                                                            0
      2
             41
                    0
                         1
                                130
                                       204
                                               0
                                                         0
                                                                  172
                                                                           0
                                                                                    1.4
                                                                                            2
      3
             56
                         1
                                120
                                       236
                                               0
                                                         1
                                                                  178
                                                                           0
                                                                                    0.8
                                                                                            2
                    1
      4
             57
                    0
                         0
                                120
                                       354
                                               0
                                                         1
                                                                  163
                                                                                    0.6
                                                                                            2
                                                                            1
                                                         •••
                                                                                    0.2
      298
             57
                    0
                         0
                                140
                                       241
                                               0
                                                         1
                                                                  123
                                                                           1
                                                                                            1
      299
             45
                    1
                         3
                                       264
                                               0
                                                                  132
                                                                           0
                                                                                    1.2
                                                                                            1
                                110
                                                         1
      300
             68
                    1
                         0
                                144
                                       193
                                               1
                                                         1
                                                                  141
                                                                           0
                                                                                    3.4
                                                                                            1
      301
                    1
                                               0
                                                         1
                                                                  115
                                                                                    1.2
             57
                         0
                                130
                                       131
                                                                           1
                                                                                            1
      302
             57
                    0
                         1
                                130
                                       236
                                               0
                                                         0
                                                                  174
                                                                           0
                                                                                    0.0
                                                                                            1
                          output
                  thall
            caa
      0
              0
                       1
                       2
      1
              0
      2
              0
                      2
                                1
                      2
      3
              0
                                1
      4
              0
                       2
                                1
      298
              0
                       3
                                0
      299
                       3
                                0
              0
      300
                       3
               2
                                0
                       3
      301
                                0
               1
      302
      [302 rows x 14 columns]
[14]: data1.isnull().sum().sum()
[14]: 0
[15]: data2=data.fillna(method='pad')
      data2
```

True

True

True

True

True

True

True

C:\Users\kumar\AppData\Local\Temp\ipykernel_27712\1171554614.py:1: FutureWarning: DataFrame.fillna with 'method' is deprecated and will raise in a future version. Use obj.ffill() or obj.bfill() instead.

data2=data.fillna(method='pad')

[15]:		age	sex	ср	trtbps	chol	fbs	restecg	thalachh	exng	oldpeak	slp	\
	0	63	1	3	145	233	1	0	150	0	2.3	0	
	1	37	1	2	130	250	0	1	187	0	3.5	0	
	2	41	0	1	130	204	0	0	172	0	1.4	2	
	3	56	1	1	120	236	0	1	178	0	0.8	2	
	4	57	0	0	120	354	0	1	163	1	0.6	2	
							•••						
	298	57	0	0	140	241	0	1	123	1	0.2	1	
	299	45	1	3	110	264	0	1	132	0	1.2	1	
	300	68	1	0	144	193	1	1	141	0	3.4	1	
	301	57	1	0	130	131	0	1	115	1	1.2	1	
	302	57	0	1	130	236	0	0	174	0	0.0	1	

	caa	thall	output
0	0	1	1
1	0	2	1
2	0	2	1
3	0	2	1
4	0	2	1
	•••	•••	•••
298	0	3	0
299	0	3	0
300	2	3	0
301	1	3	0
302	1	2	0

[302 rows x 14 columns]

[16]: data3=data.fillna(method='bfill')
 data3

C:\Users\kumar\AppData\Local\Temp\ipykernel_27712\3254375401.py:1: FutureWarning: DataFrame.fillna with 'method' is deprecated and will raise in a future version. Use obj.ffill() or obj.bfill() instead.

data3=data.fillna(method='bfill')

[16]:	age	sex	ср	trtbps	chol	fbs	restecg	thalachh	exng	oldpeak	${ t slp}$	\
0	63	1	3	145	233	1	0	150	0	2.3	0	
1	37	1	2	130	250	0	1	187	0	3.5	0	
2	41	0	1	130	204	0	0	172	0	1.4	2	
3	56	1	1	120	236	0	1	178	0	0.8	2	
4	57	0	0	120	354	0	1	163	1	0.6	2	
						•••						

```
299
            45
                  1
                      3
                            110
                                   264
                                          0
                                                   1
                                                           132
                                                                   0
                                                                           1.2
                                                                                  1
      300
                  1
                      0
                            144
                                   193
                                                   1
                                                                           3.4
            68
                                          1
                                                           141
                                                                   0
                                                                                  1
      301
                  1
                      0
                                                                           1.2
            57
                            130
                                   131
                                          0
                                                   1
                                                           115
                                                                   1
                                                                                  1
      302
            57
                      1
                            130
                                   236
                                          0
                                                   0
                                                           174
                                                                   0
                                                                           0.0
                                                                                  1
           caa thall output
      0
             0
                    1
      1
             0
                    2
                            1
      2
             0
                    2
                            1
                    2
      3
             0
                            1
                    2
      4
             0
                            1
      . .
      298
             0
                    3
                            0
      299
             0
                    3
                            0
      300
             2
                    3
                            0
      301
                    3
                            0
             1
      302
                    2
             1
                            0
      [302 rows x 14 columns]
[17]: import numpy as np
      import matplotlib.pyplot as plt
      from scipy import stats
[18]: data1.columns
[18]: Index(['age', 'sex', 'cp', 'trtbps', 'chol', 'fbs', 'restecg', 'thalachh',
             'exng', 'oldpeak', 'slp', 'caa', 'thall', 'output'],
            dtype='object')
[19]: data1.drop(['oldpeak'], axis=1, inplace=True)
      data1.columns
[19]: Index(['age', 'sex', 'cp', 'trtbps', 'chol', 'fbs', 'restecg', 'thalachh',
             'exng', 'slp', 'caa', 'thall', 'output'],
            dtype='object')
[20]: Q1=data1.quantile(0.25)
      Q3=data1.quantile(0.75)
      IQR=Q3-Q1
      print(IQR)
                 13.00
     age
                  1.00
     sex
                  2.00
     ср
                 20.00
     trtbps
```

0.2

```
chol
            63.75
fbs
             0.00
restecg
             1.00
thalachh
            32.75
             1.00
exng
slp
             1.00
             1.00
caa
thall
             1.00
output
             1.00
dtype: float64
```

```
[23]: data1=data1[~((data1<(Q1-1.5*IQR))|(data1>(Q3+1.5*IQR))).any(axis=1)] data1
```

```
[23]:
                                     chol
                                                             thalachh
                            trtbps
                                            fbs
                                                  restecg
                                                                        exng
                                                                               slp
                                                                                     caa
            age
                  sex
                        ср
                         2
             37
                                130
                                       250
                                               0
                                                         1
                                                                   187
                                                                            0
                                                                                  0
                                                                                       0
      1
                    1
      2
                                                         0
                                                                   172
                                                                                  2
                                                                                       0
             41
                    0
                         1
                                130
                                       204
                                               0
                                                                            0
      3
             56
                                       236
                                               0
                                                         1
                                                                   178
                                                                            0
                                                                                       0
                         1
                                120
      4
             57
                         0
                                120
                                       354
                                               0
                                                         1
                                                                   163
                                                                            1
                                                                                  2
                                                                                       0
      5
             57
                         0
                                140
                                       192
                                                                   148
                    1
                                               0
                                                         1
                                                                                  1
                                                                                       0
                                                         •••
      296
             63
                    0
                         0
                                124
                                       197
                                               0
                                                         1
                                                                   136
                                                                            1
                                                                                  1
                                                                                       0
      298
             57
                    0
                         0
                                140
                                       241
                                               0
                                                         1
                                                                   123
                                                                            1
                                                                                  1
                                                                                       0
      299
                         3
                                                         1
                                                                   132
                                                                            0
                                                                                       0
             45
                    1
                                110
                                       264
                                               0
                                                                                  1
      301
             57
                    1
                         0
                                130
                                       131
                                               0
                                                         1
                                                                   115
                                                                            1
                                                                                  1
                                                                                        1
      302
             57
                         1
                                               0
                                                         0
                                                                   174
                                                                                  1
                                                                                        1
                                130
                                       236
                                                                            0
```

```
thall
             output
1
          2
                    1
2
          2
                    1
          2
3
                    1
          2
4
                    1
5
          1
                    1
. .
296
          2
                    0
298
          3
                    0
299
          3
                    0
301
          3
                    0
          2
302
                    0
```

[229 rows x 13 columns]

[]: