STATISTICAL DISCRIPTION

```
# Aim: To Perform Statistical Description on Data
#Name: Samruddhi Umap
#Roll no. :67
#Sub : ET1
#section:3C
#Date: 26/07/2024
import pandas as pd
import os
os.getcwd()
'C:\\Users\\HP'
os.chdir("C:\\Users\\HP\\Desktop")
data=pd.read_csv("framingham.csv")
data.head()
   male age education currentSmoker cigsPerDay BPMeds
prevalentStroke
                    4.0
                                               0.0
      1
          39
                                                       0.0
0
                    2.0
                                               0.0
1
      0
          46
                                                       0.0
0
2
      1
          48
                    1.0
                                              20.0
                                                       0.0
0
3
                    3.0
      0
          61
                                              30.0
                                                       0.0
0
4
      0
          46
                    3.0
                                     1
                                              23.0
                                                       0.0
   prevalentHyp
                 diabetes totChol sysBP diaBP
                                                    BMI
                                                        heartRate
glucose \
                                            70.0 26.97
              0
                             195.0 106.0
                                                              80.0
77.0
              0
                             250.0 121.0
                                            81.0 28.73
                                                              95.0
76.0
              0
                        0
                             245.0 127.5 80.0 25.34
                                                              75.0
70.0
              1
                             225.0 150.0
                                            95.0 28.58
                                                              65.0
103.0
              0
                             285.0 130.0
                                            84.0 23.10
                                                              85.0
85.0
```

0 1 2 3 4	nYear	0 0 0 1 0							
data.	tail	()							
4233 4234 4235 4236 4237	male 1 0 0	L 50 L 51 O 48 O 44		. 0 . 0 . 0	moker 1 1 1 1 0	cigs	PerDay 1.0 43.0 20.0 15.0 0.0	BPMeds 0.0 0.0 NaN 0.0	\
	prev	/alentS	troke pr	revalentHyp	diabe	tes	totChol	sysBP	diaBP
BMI 4233 25.97	\		0	1		0	313.0	179.0	92.0
4234			0	0		0	207.0	126.5	80.0
19.71 4235			0	0		0	248.0	131.0	72.0
22.00									
4236 19.16			0	0		0	210.0	126.5	87.0
4237			0	0		0	269.0	133.5	83.0
21.47									
4233 4234 4235 4236 4237	hear	66.0 65.0 84.0 86.0 80.0	glucose 86.0 68.0 86.0 NaN 107.0	TenYearCHD 1 0 0 0					
data.head(10)									
ma	le a		ucation \	currentSmok	er ci	gsPer	Day BP	Meds	
0	1	39	4.0		0		0.0	0.0	
0 1	0	46	2.0		0		0.0	0.0	
0									
2 0	1	48	1.0		1	2	0.0	0.0	
3	0	61	3.0		1	3	0.0	0.0	
0 4	0	46	3.0		1	7	3.0	0.0	
4	U	40	3.0		1		5.0	0.0	

0 5	0	43	2.0		0	0.	0 0	. 0
0 6	0	63	1.0		0	0.	0 0	. 0
0								
7 0	0	45	2.0		1	20.	0 0	. 0
8	1	52	1.0		Θ	0.	0 0	.0
0 9 0	1	43	1.0		1	30.	0 0	. 0
pre glucos		entHyp \	diabetes	totChol	sysBP	diaBP	BMI	heartRate
0		` 0	0	195.0	106.0	70.0	26.97	80.0
77.0 1 76.0		0	0	250.0	121.0	81.0	28.73	95.0
2		0	0	245.0	127.5	80.0	25.34	75.0
70.0 3 103.0		1	0	225.0	150.0	95.0	28.58	65.0
4		0	0	285.0	130.0	84.0	23.10	85.0
85.0		1	Θ	228.0	180.0	110.0	30.30	77.0
99.0 6		0	0	205.0	138.0	71.0	33.11	60.0
85.0 7		0	0	313.0	100.0	71.0	21.68	79.0
78.0 8		1	0	260.0	141.5	89.0	26.36	76.0
79.0 9		1	0	225.0	162.0	107.0	23.61	93.0
88.0		1	U	223.0	102.0	107.0	23.01	93.0
	าYea	rCHD 0						
0 1		0						
1 2 3 4		0 1						
4		0						
5 6 7		0 1						
		0						
8 9		0 0						
data.1	tail	(10)						

Mark Mark											
BMI	4229 4230 4231 4232 4233 4234 4235 4236	0 0 1 1 1 1 0	50 51 56 58 68 50 51 48 44	1 3 1 3 1 1 3 2 1	. 0 . 0 . 0 . 0 . 0 . 0 . 0	ntSmo	0 1 0 0 1 1 1	cigs	0.0 20.0 3.0 0.0 0.0 1.0 43.0 20.0 15.0	0.0 0.0 0.0 0.0 0.0 0.0 NaN	
4228	DMT '		lentS	troke p	revalentH	ур с	diabet	tes	totChol	sysBP	diaBP
4229 0 1 0 251.0 140.0 80.0 25.60 4230 0 1 0 268.0 170.0 102.0 22.89 4231 0 1 0 187.0 141.0 81.0 24.96 4232 0 1 0 176.0 168.0 97.0 23.14 4233 0 1 0 313.0 179.0 92.0 25.97 4234 0 0 0 207.0 126.5 80.0 19.71 4235 0 0 0 248.0 131.0 72.0 22.00 4236 0 0 0 210.0 126.5 87.0 19.16 4237 0 0 0 269.0 133.5 83.0 21.47 heartRate glucose TenYearCHD 4228 85.0 260.0 0 0 249.0 133.5 83.0 2423 66.0 86.0 1 2423 60.0 79.0 1 2423 60.0 79.0 1 2423 60.0 79.0 1 2423 60.0 79.0 1 2423 60.0 79.0 1 2423 60.0 79.0 1 2423 60.0 79.0 1 2423 60.0 79.0 1 2423 60.0 79.0 1 2423 60.0 79.0 1 2423 60.0 79.0 1 2423 60.0 79.0 1 2423 60.0 79.0 1 2423 60.0 79.0 1 2423 60.0 86.0 1 2423 60.0 86.0 1 2423 60.0 86.0 0 2423 60.0 86.0 0 2423 60.0 86.0 0 2423 60.0 86.0 0 2423 60.0 86.0 0 2423 60.0 86.0 0 2423 60.0 86.0 0 2423 60.0 86.0 0 2423 60.0 86.0 0 2423 60.0 86.0 0 0 2423 60.0 86.0 0 0 2423 60.0 86.0 0 0 2423 60.0 86.0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	4228			0		1		1	260.0	190.0	130.0
4230 0 1 0 187.0 170.0 102.0 22.89 4231 0 1 0 187.0 141.0 81.0 24.96 4232 0 1 0 176.0 168.0 97.0 23.14 4233 0 1 0 313.0 179.0 92.0 25.97 4234 0 0 0 207.0 126.5 80.0 19.71 4235 0 0 0 0 248.0 131.0 72.0 22.00 4236 0 0 0 0 210.0 126.5 87.0 19.16 4237 0 0 0 0 269.0 133.5 83.0 21.47 heartRate glucose TenYearCHD 4228 85.0 260.0 0 4229 75.0 NaN 0 4230 57.0 NaN 0 4231 80.0 81.0 0 4232 60.0 79.0 1 4233 66.0 86.0 1 4234 65.0 68.0 0 4235 84.0 86.0 0 4236 86.0 NaN 0 4237 80.0 107.0 0	4229			0		1		0	251.0	140.0	80.0
4231 0 1 0 176.0 141.0 81.0 24.96 4232 0 1 0 176.0 168.0 97.0 23.14 4233 0 1 0 313.0 179.0 92.0 25.97 4234 0 0 0 0 207.0 126.5 80.0 19.71 4235 0 0 0 0 248.0 131.0 72.0 22.00 4236 0 0 0 0 210.0 126.5 87.0 19.16 4237 0 0 0 269.0 133.5 83.0 21.47 heartRate glucose TenYearCHD 4228 85.0 260.0 NaN 0 4230 57.0 NaN 0 4231 80.0 81.0 0 4231 80.0 81.0 0 4232 60.0 79.0 1 4233 66.0 86.0 1 4234 65.0 68.0 0 1 4234 65.0 68.0 0 0 4235 84.0 86.0 0 0 0 4236 86.0 NaN 0 4237 80.0 107.0 0	4230			0		1		0	268.0	170.0	102.0
4232 0 1 0 176.0 168.0 97.0 23.14 4233 0 1 0 313.0 179.0 92.0 25.97 4234 0 0 0 207.0 126.5 80.0 19.71 4235 0 0 0 248.0 131.0 72.0 22.00 4236 0 0 0 210.0 126.5 87.0 19.16 4237 0 0 0 269.0 133.5 83.0 21.47 heartRate glucose TenYearCHD 4228 85.0 260.0 0 4229 75.0 NaN 0 4230 57.0 NaN 0 4231 80.0 81.0 0 4232 60.0 79.0 1 4233 66.0 86.0 1 4234 65.0 68.0 0 4235 84.0 86.0 0 4236 86.0 NaN 0 4237 80.0 107.0 0	4231			0		1		0	187.0	141.0	81.0
4233 0 1 0 313.0 179.0 92.0 25.97 4234 0 0 0 207.0 126.5 80.0 19.71 4235 0 0 0 248.0 131.0 72.0 22.00 4236 0 0 0 210.0 126.5 87.0 19.16 4237 0 0 0 269.0 133.5 83.0 21.47 heartRate glucose TenYearCHD 4228 85.0 260.0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	4232			0		1		0	176.0	168.0	97.0
4234 0 0 0 207.0 126.5 80.0 19.71 4235 0 0 0 248.0 131.0 72.0 22.00 4236 0 0 0 210.0 126.5 87.0 19.16 4237 0 0 0 269.0 133.5 83.0 21.47 heartRate glucose TenYearCHD 4228 85.0 260.0 0 4230 57.0 NaN 0 4231 80.0 81.0 0 4232 60.0 79.0 1 4233 66.0 86.0 1 4234 65.0 68.0 0 4235 84.0 86.0 0 4236 86.0 NaN 0 4237 80.0 107.0 0	4233			0		1		0	313.0	179.0	92.0
4235 0 0 0 248.0 131.0 72.0 22.00 4236 0 0 0 210.0 126.5 87.0 19.16 4237 0 0 0 269.0 133.5 83.0 21.47 heartRate glucose TenYearCHD 4228 85.0 260.0 0 4229 75.0 NaN 0 4230 57.0 NaN 0 4231 80.0 81.0 0 4232 60.0 79.0 1 4233 66.0 86.0 1 4234 65.0 68.0 0 4235 84.0 86.0 0 4236 86.0 NaN 0 4237 80.0 107.0 0	4234			0		0		0	207.0	126.5	80.0
4236 0 0 0 210.0 126.5 87.0 19.16 4237 0 0 0 269.0 133.5 83.0 21.47 heartRate glucose TenYearCHD 4228 85.0 260.0 0 4229 75.0 NaN 0 4230 57.0 NaN 0 4231 80.0 81.0 0 4232 60.0 79.0 1 4233 66.0 86.0 1 4234 65.0 68.0 0 4235 84.0 86.0 0 4236 86.0 NaN 0 4237 80.0 107.0 0	4235			0		0		0	248.0	131.0	72.0
4237 0 0 0 269.0 133.5 83.0 21.47 heartRate glucose TenYearCHD 4228 85.0 260.0 0 4229 75.0 NaN 0 4230 57.0 NaN 0 4231 80.0 81.0 0 4232 60.0 79.0 1 4233 66.0 86.0 1 4234 65.0 68.0 0 4235 84.0 86.0 0 4236 86.0 NaN 0 4237 80.0 107.0 0	4236			0		0		0	210.0	126.5	87.0
heartRate glucose TenYearCHD 4228 85.0 260.0 0 4229 75.0 NaN 0 4230 57.0 NaN 0 4231 80.0 81.0 0 4232 60.0 79.0 1 4233 66.0 86.0 1 4234 65.0 68.0 0 4235 84.0 86.0 0 4236 86.0 NaN 0 4237 80.0 107.0 0	4237			0		0		0	269.0	133.5	83.0
<pre>data.describe()</pre>	4228 4229 4230 4231 4232 4233 4234 4235 4236		85.0 75.0 57.0 80.0 60.0 66.0 65.0 84.0	260.0 NaN NaN 81.0 79.0 86.0 68.0 86.0 NaN	TenYear	0 0 0 0 1 1 0 0					
	data.	descri	be()								

				a	_
cigsPe	male rDay \	age	education	currentSmoke	T
count 4209.0	4238.000000	4238.000000	4133.000000	4238.00000	0
mean 9.0030	0.429212	49.584946	1.978950	0.49410	1
std	0.495022	8.572160	1.019791	0.50002	4
11.920 min	0.000000	32.000000	1.000000	0.00000	0
0.0000 25%	0.000000	42.000000	1.000000	0.00000	0
0.0000 50%	0.000000	49.000000	2.000000	0.00000	0
0.0000 75%	1.000000	56.000000	3.000000	1.00000	0
20.000	000				
max 70.000	1.000000	70.000000	4.000000	1.00000	U
	BPMeds	prevalentStr	oke prevalen	tHyp diab	etes
totCho count	l \ 4185.000000	4238.000	000 4238.00	0000 4238.00	0000
4188.0 mean	00000	0.005	899 0.31	0524 0.02	5720
236.72	1585				
std 44.590	0.169584 334	0.076	587 0.46	2/63 0.15	8316
min 107.00	0.000000	0.000	000 0.00	0000 0.00	0000
25%	0.000000	0.000	000 0.00	0000 0.00	0000
206.00 50%	0.000000	0.000	000 0.00	0000 0.00	0000
234.00 75%	0000	0.000	000 1.00	0000 0.00	0000
263.00	0000 1.000000	1.000	000 1.00	0000 1 00	0000
max 696.00		1.000	1.00	1.00	0000
	sysBP	diaBP	BMI	heartRate	glucose
\ count	4238.000000	4238.000000	4219.000000	4237.000000	3850.000000
mean	132.352407	82.893464	25.802008	75.878924	81.966753
std	22.038097	11.910850	4.080111	12.026596	23.959998
min	83.500000	48.000000	15.540000	44.000000	40.000000
25%	117.000000	75.000000	23.070000	68.000000	71.000000

```
50%
        128.000000
                       82.000000
                                    25.400000
                                                  75.000000
                                                                78.000000
75%
        144.000000
                       89.875000
                                    28.040000
                                                  83.000000
                                                                87.000000
        295.000000
                      142.500000
                                    56.800000
                                                 143,000000
                                                               394.000000
max
        TenYearCHD
       4238.000000
count
          0.151958
mean
          0.359023
std
          0.000000
min
25%
          0.000000
          0.000000
50%
75%
          0.000000
          1.000000
max
data.ndim
2
data.info()
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 4238 entries, 0 to 4237
Data columns (total 16 columns):
#
                       Non-Null Count
     Column
                                       Dtype
- - -
 0
     male
                       4238 non-null
                                        int64
 1
     age
                       4238 non-null
                                        int64
 2
                       4133 non-null
                                       float64
     education
 3
                       4238 non-null
     currentSmoker
                                        int64
 4
     cigsPerDay
                       4209 non-null
                                       float64
 5
     BPMeds
                       4185 non-null
                                       float64
 6
     prevalentStroke
                       4238 non-null
                                        int64
 7
                       4238 non-null
     prevalentHyp
                                        int64
 8
     diabetes
                       4238 non-null
                                       int64
 9
     totChol
                       4188 non-null
                                       float64
     sysBP
 10
                       4238 non-null
                                        float64
 11
     diaBP
                       4238 non-null
                                       float64
 12
     BMI
                       4219 non-null
                                       float64
 13
     heartRate
                       4237 non-null
                                       float64
     glucose
                       3850 non-null
                                       float64
 14
 15
     TenYearCHD
                       4238 non-null
                                       int64
dtypes: float64(9), int64(7)
memory usage: 529.9 KB
data.size
67808
```

data.shape

(4238, 16)