CASE STUDY 2

CHOLESTEROL DATA SEGREGATION INTO GOOD AND BAD

CHOLESTEROL DATA

DATA 1	
BODY FAT	Z SCORE
12.3	0.962149752
6.1	-0.716648127
10.4	0.447679434
12.4	0.989227137
4.1	-1.25819583
11.7	0.799685441
7.1	-0.445874275
7.8	-0.256332579
15.2	1.747393921
14	1.422465299
3.7	-1.36650537
3.7	-1.36650537
8.8	0.014441272
5.7	-0.824957667
2	-1.826820918

7.7	-0.283409964
13.9	1.395387914
10.8	0.555988975
5.6	-0.852035052
13.6	1.314155759
9	0.068596042
10.2	0.393524664
6.6	-0.581261201
8	-0.202177809
6.3	-0.662493356
3.9	-1.3123506
13.8	1.368310529
6.3	-0.662493356
12.9	1.124614063
8.8	0.014441272

DATA 2	
BODY FAT	Z SCORE
60	2.019248179
20.9	-0.240414208
19.2	-0.338660399
15	-0.581386282
21.2	-0.223076645
22.1	-0.171063956
20.9	-0.240414208
51	1.499121287
22.9	-0.124830454
16	-0.523594405
16.5	-0.494698466
19.1	-0.344439586
15.6	-0.546711156
17.7	-0.425348214
59	1.961456303
40.6	0.898085767
38.6	0.782502013
37.9	0.7420477
36.4	0.655359884
39.1	0.811397952
35.3	0.59178882
45.1	1.158149213
48	1.325745656
3.7	-1.234434491
0.5	-1.419368497
20	-0.292426897
5.3	-1.141967488
1.7	-1.350018245
2	-1.332680682
0.5	-1.419368497

SUMMARIZATION

- DATA 1

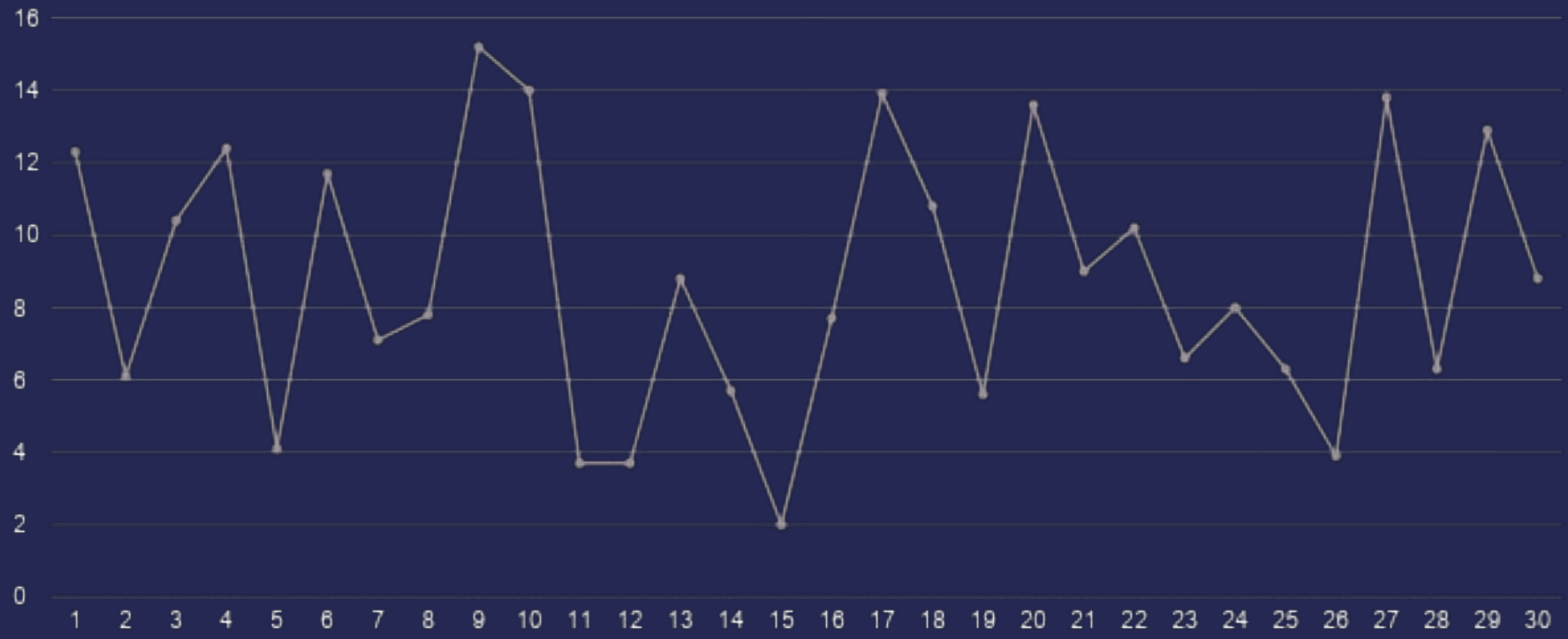
CENTRAL TENDENCY	
mean	8.74666666666667
median	8.4
mode	3.7
trimmed mean	8.747
harmonic mean	6.886
95% upper mean	8.845
95% lower mean	8.649
DISPERSIONS	
standard deviation	3.69311879537899
standard error	0.674
variance	13.639
cv	0.422
DQ	0.787
minimum	2
maximum	15.2
range	13.2

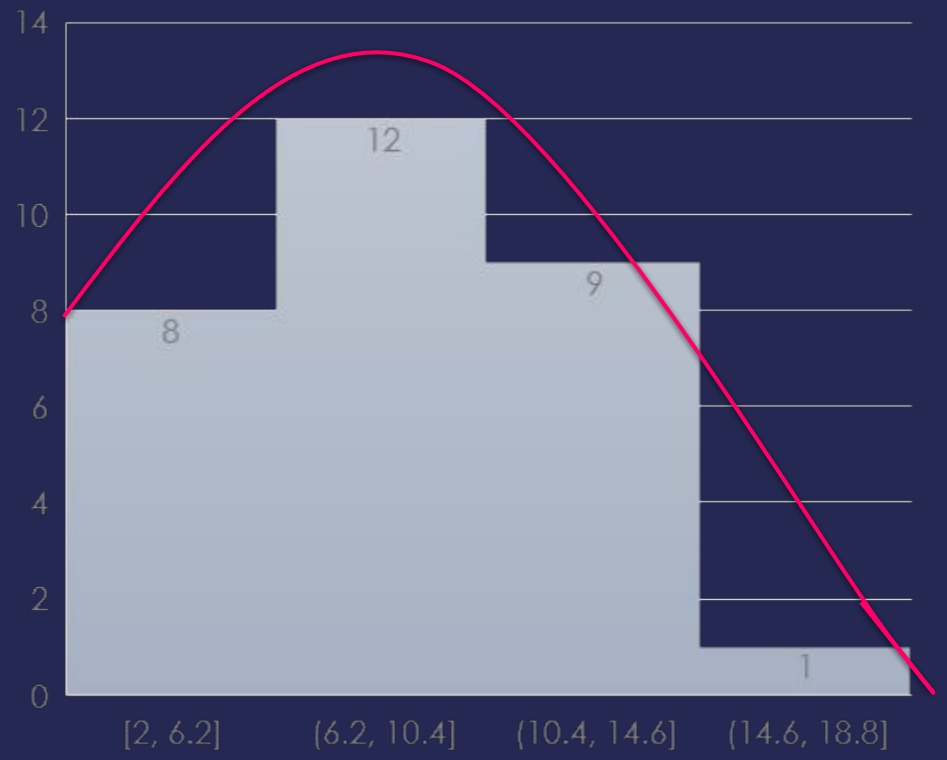
- DATA 2

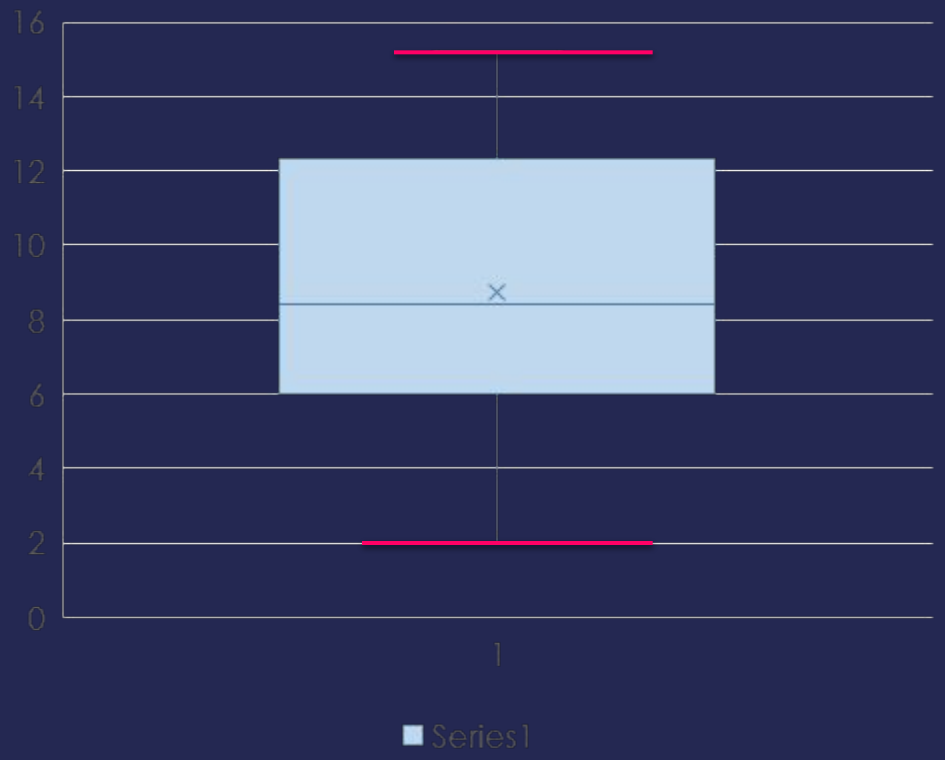
CENTRAL TENDENCY	
mean	25.06
median	20.9
mode	20.9
trimmed mean	25.06
harmonic mean	4.616587594
95% upper mean	25.158
95% lower mean	24.962
DISPERSIONS	
standard deviation	17.30346985
standard error	3.159166921
variance	299.410069
cv	1.448264436
DQ	0.184221372
minimum	0.5
maximum	60
range	59.5

VISUALIZATION

DATA 1

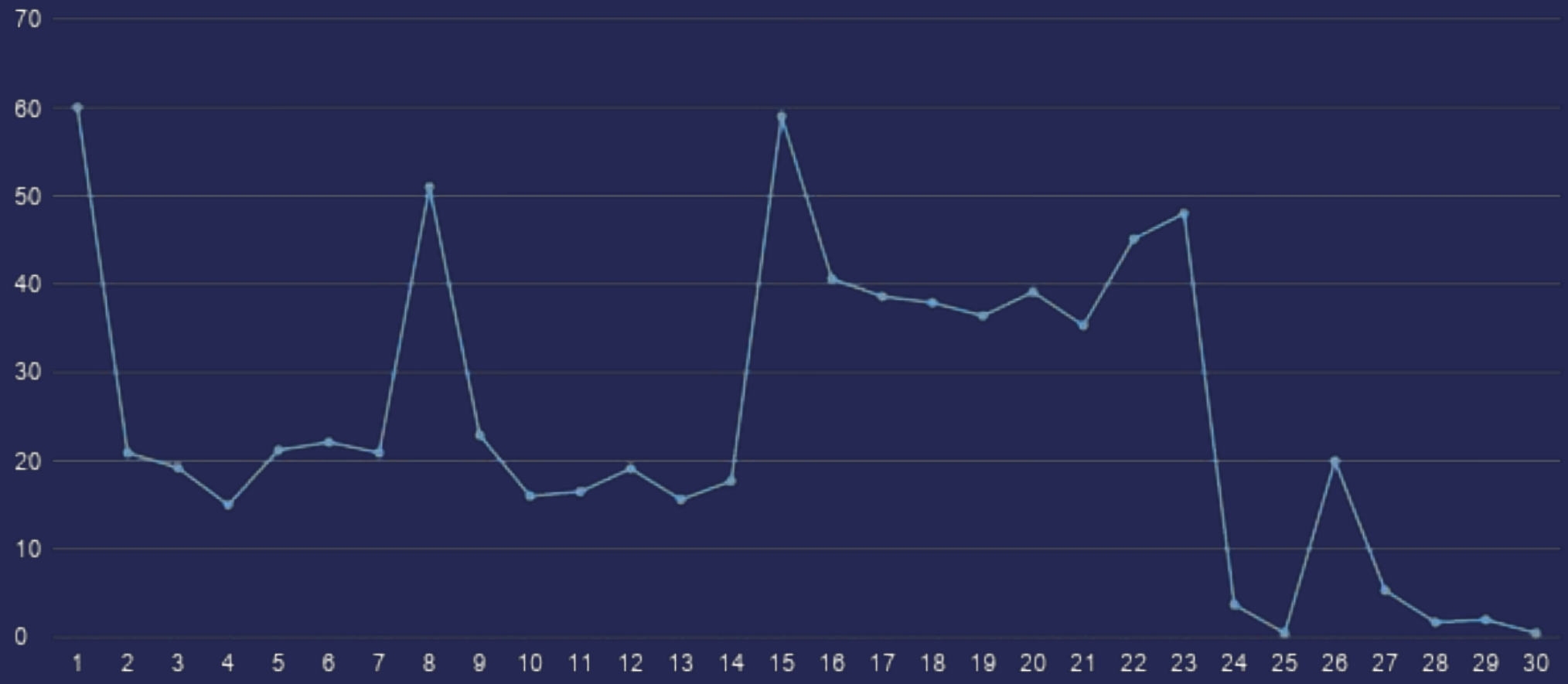


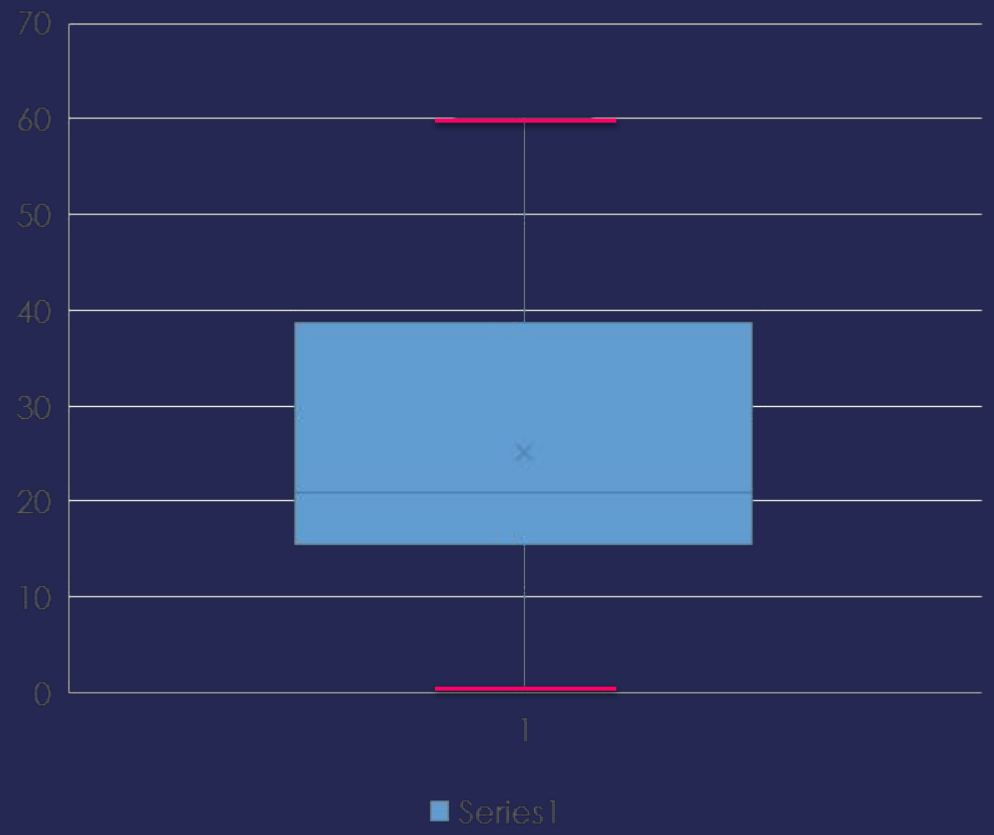


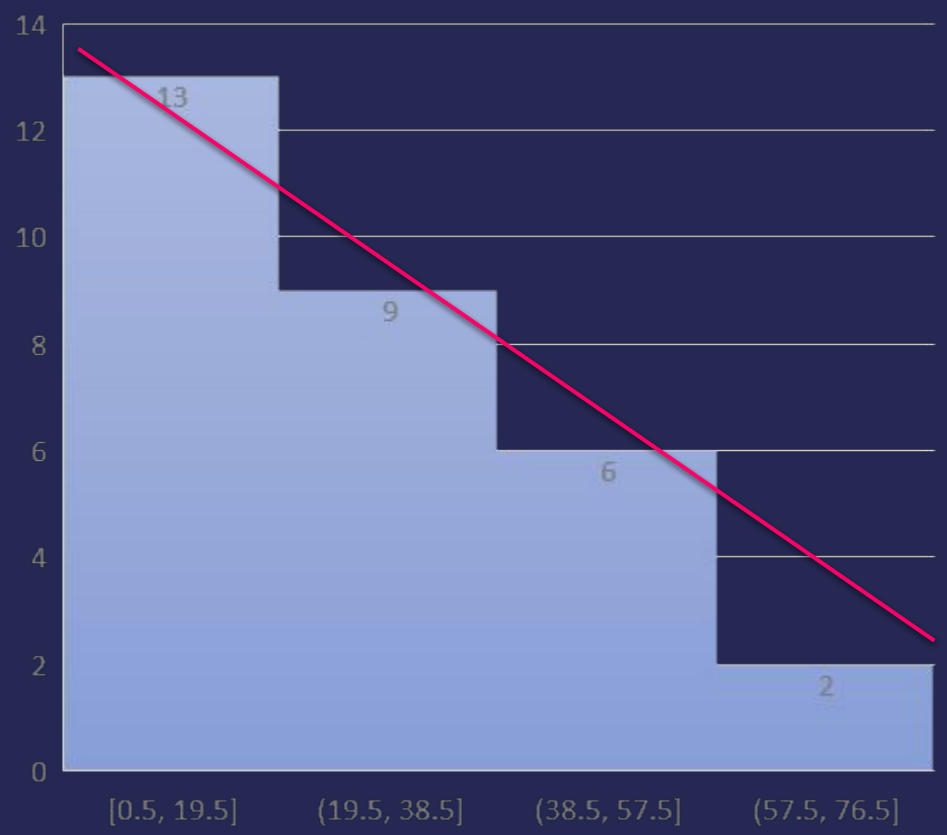


Since the data is following normal distribution, the data is good and there are no outliers. The data quality is 78.7%.

DATA 2


70
60
50
40
30
20
10
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30





Since the data is not following normal distribution, and there are two outliers.
The data quality is 18.4%.