Descriptive Statistics Review

Descriptive Measures

Central Tendency

Variation

Relative Standing

Central Tendency

Mean

Median

Mode

Mean

Feature 1
3
5
5
1
7
2
6
7
0
4
40

[3, 5, 5, 1, 7, 2, 6, 7, 0, 4]

Sum =

40/10 = **4**

Median

Feature 1
0
1
2
3
4
5
5
6
7
7
Sum = 40

Put the numbers in order

Half of measures are above

= 4.5

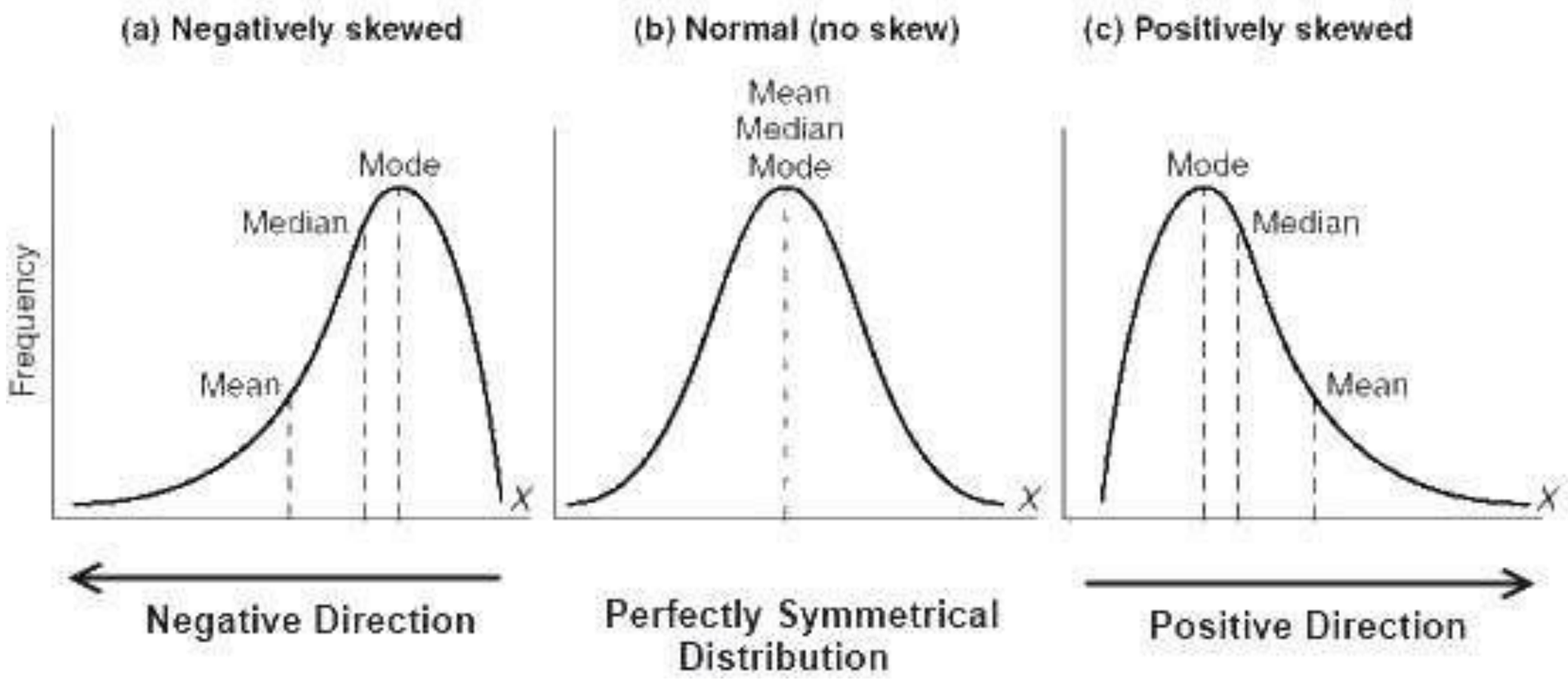
Half of measures are below

Mode

Feature 1
0
1
2
3
4
5
5
6
7
7
Sum = 40

Number(s) which appears most often

= 5 and 7


(a) Negatively skewed
(b) Normal (no skew)
(c) Positively skewed
Frequency
Mode
Median
Mean
Mean
Median
Mode
Mode
Median
Mean
X
X
X
Negative Direction
Perfectly Symmetrical Distribution
Positive Direction

Variation

Variance

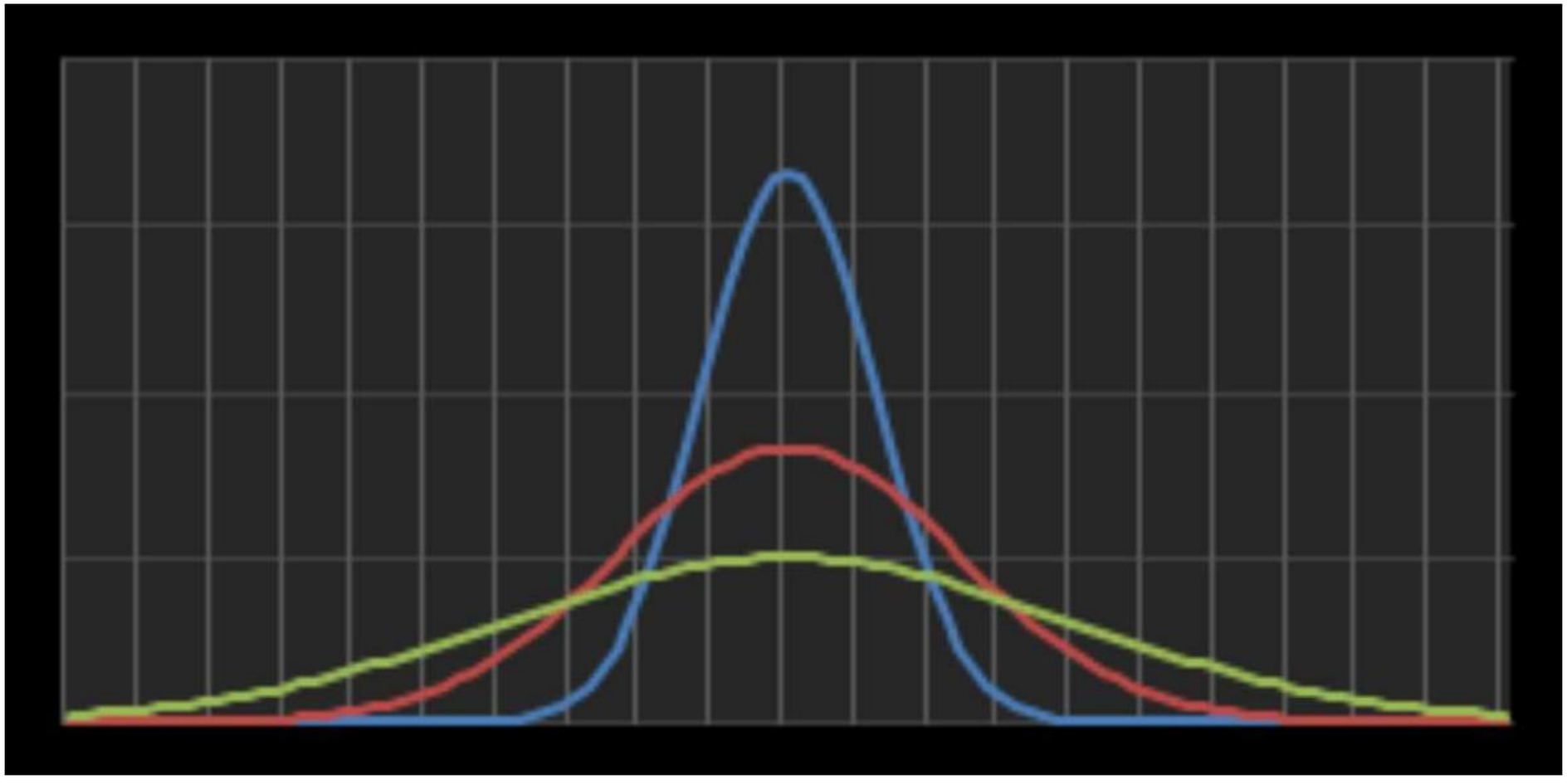
Standard Deviation

Range

Quartiles

Interquartile Range

The spread of the data



Variance

Feature 1	Deviations	Squared Deviations
0	-4	16
1	-3	9
2	-2	4
3	-1	1
4	0	0
5	1	1
5	1	1
6	2	4
7	3	9
7	3	9
40	0	54

Mean = 4

54/9 = **6**

Standard Deviation

Feature 1	Deviations	Squared Deviations
0	-4	16
1	-3	9
2	-2	4
3	-1	1
4	0	0
5	1	1
5	1	1
6	2	4
7	3	9
7	3	9
40	0	54

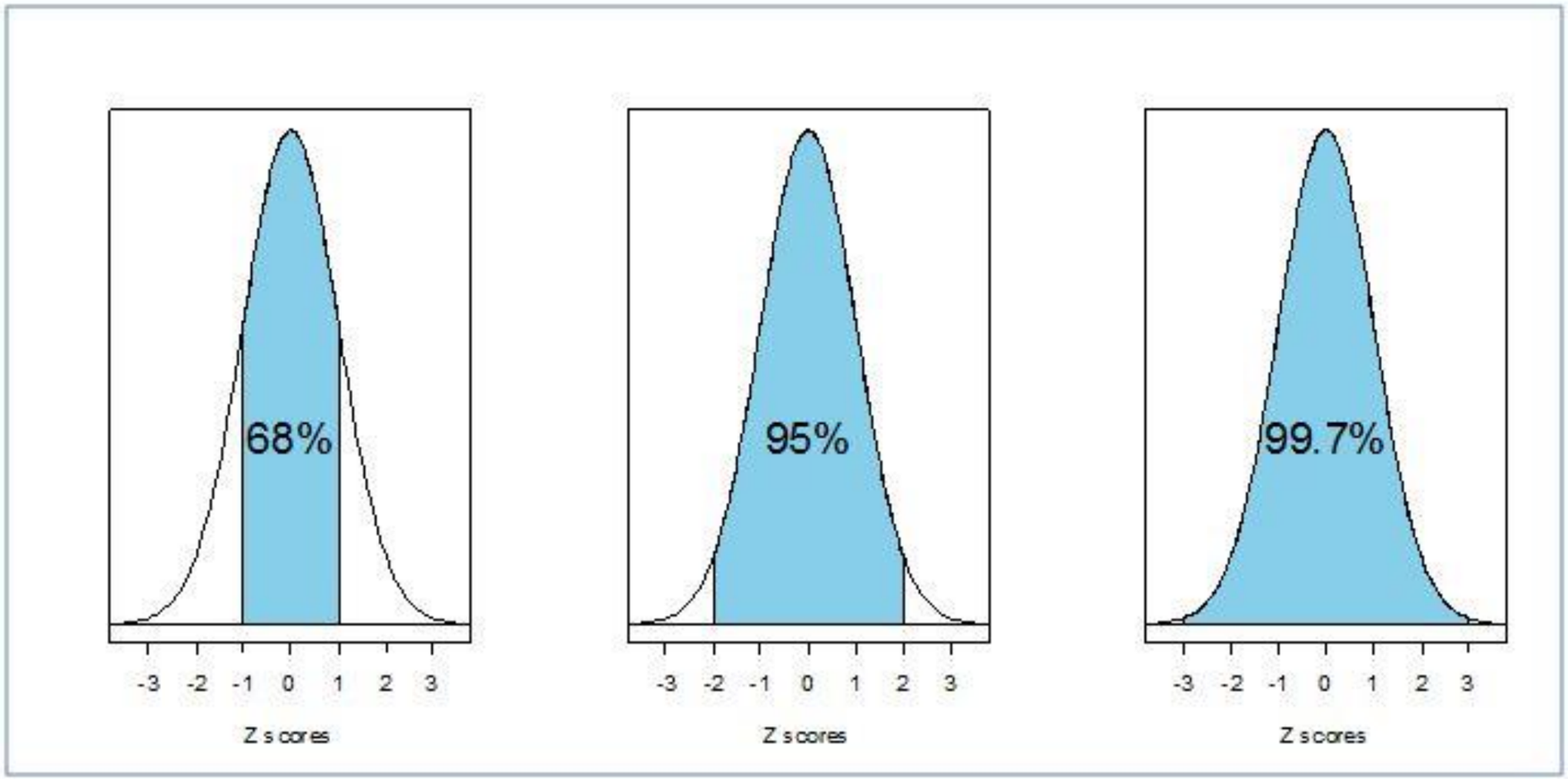
Mean = 4

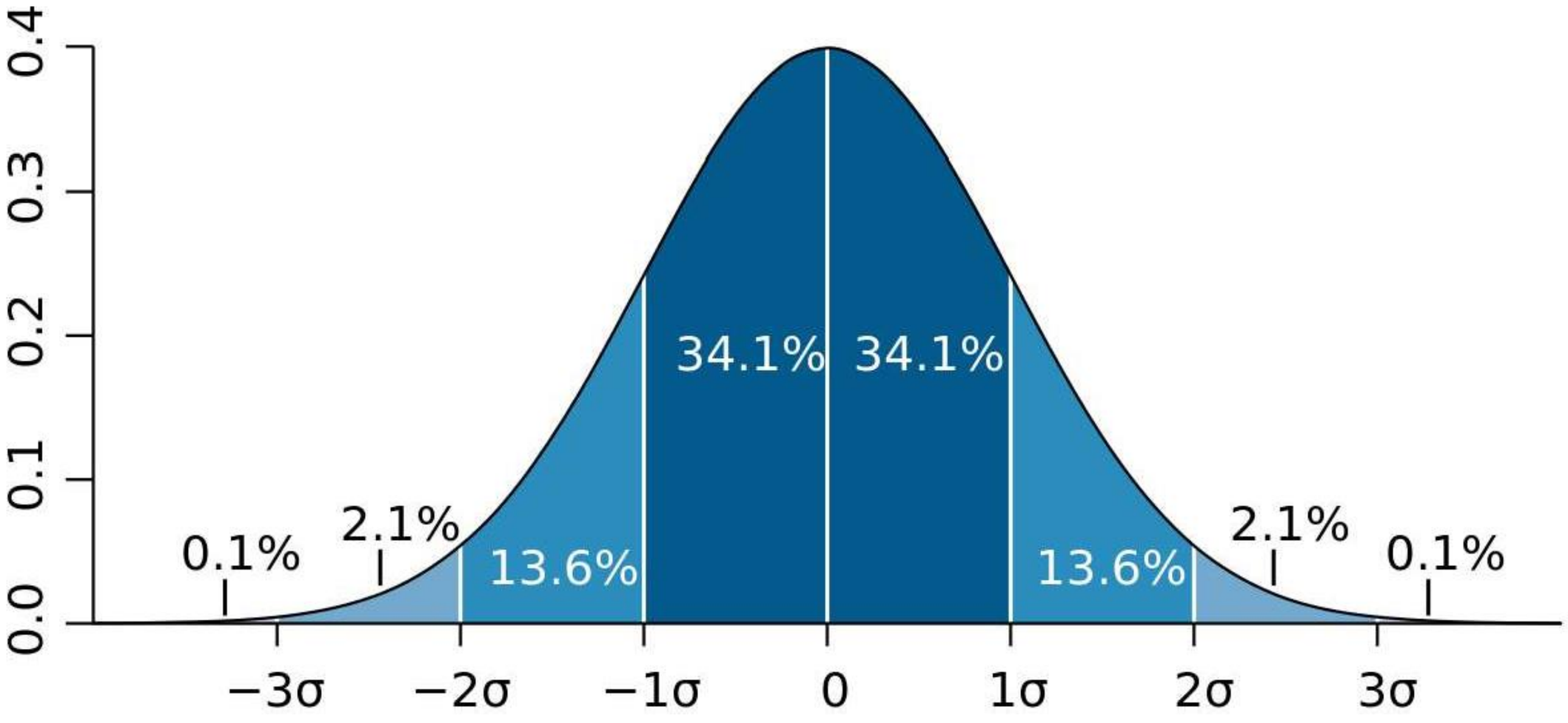
$\sqrt{6} = \mathbf{2.45}$

Standard Score
(standardization)

z-score = (x - mean)/std

Empirical Rule




0.4
0.3
0.2
0.1
0.0
34.1%
34.1%
13.6%
13.6%
2.1%
2.1%
0.1%
0.1%
−3σ
−2σ
−1σ
0
1σ
2σ
3σ

Measure	Qty
Price	1 million
Sq. Ft.	2,500
Neighborhood Rating	79
# Bedrooms	4

Measure	Qty
Price	1 million
Sq. Ft.	2,500
Neighborhood Rating	79
# Bedrooms	4

4 79 2500 1 m

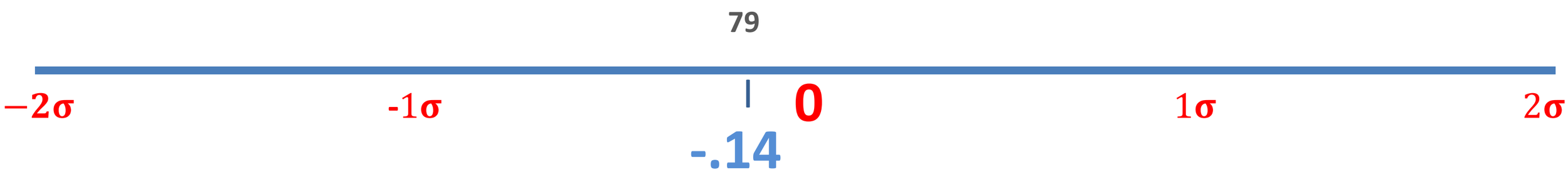
0

Measure	Qty
Price	1 million
Sq. Ft.	2,500
Neighborhood Rating	79
# Bedrooms	4

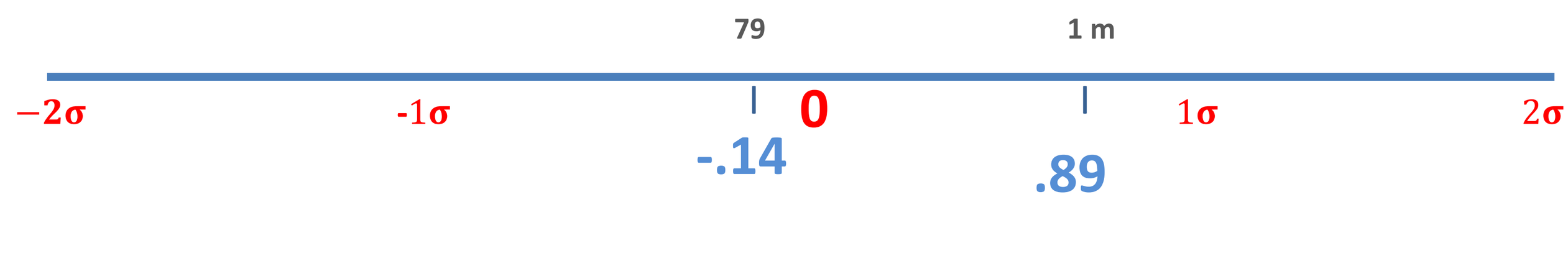
Measure	Qty	Standardized (z-score)
Price	1 million	.89
Sq. Ft.	2,500	.96
Neighborhood Rating	79	-.14
# Bedrooms	4	1.32

Measure	Qty	Standardized (z-score)
Price	1 million	.89
Sq. Ft.	2,500	.96
Neighborhood Rating	79	-.14
# Bedrooms	4	1.32

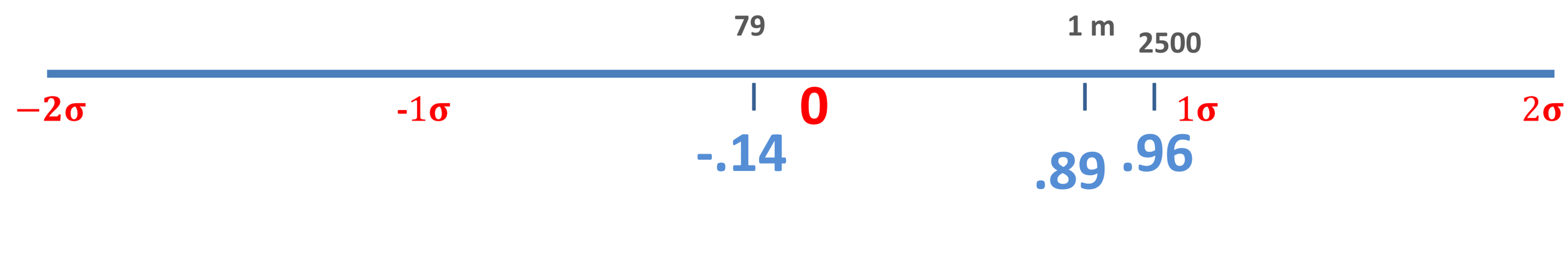
Measure	Qty	Standardized (z-score)
Price	1 million	.89
Sq. Ft.	2,500	.96
Neighborhood Rating	79	-.14
# Bedrooms	4	1.32



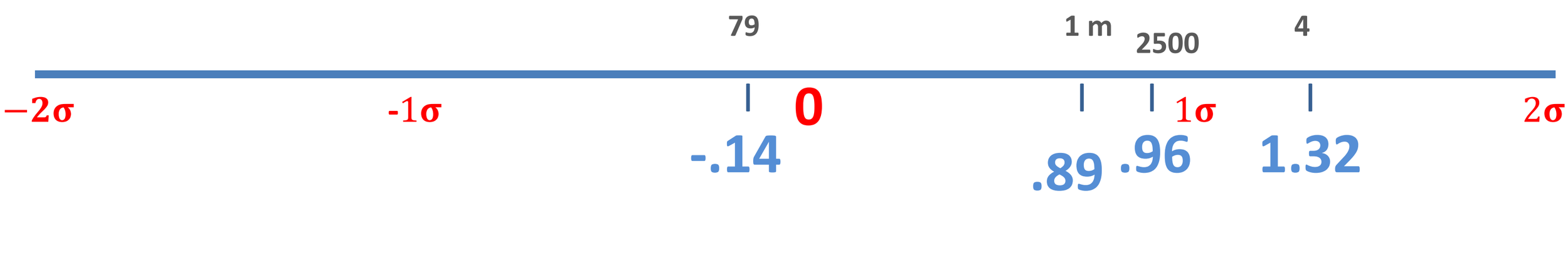
Measure	Qty	Standardized (z-score)
Price	1 million	.89
Sq. Ft.	2,500	.96
Neighborhood Rating	79	-.14
# Bedrooms	4	1.32



Measure	Qty	Standardized (z-score)
Price	1 million	.89
Sq. Ft.	2,500	.96
Neighborhood Rating	79	-.14
# Bedrooms	4	1.32



Measure	Qty	Standardized (z-score)
Price	1 million	.89
Sq. Ft.	2,500	.96
Neighborhood Rating	79	-.14
# Bedrooms	4	1.32



Range

Feature 1	Deviations	Squared Deviations
0	-4	16
1	-3	9
2	-2	4
3	-1	1
4	0	0
5	1	1
5	1	1
6	2	4
7	3	9
7	3	9
40	0	54

Max value = 7

Min value = 0

7-0 = **7**

Percentiles

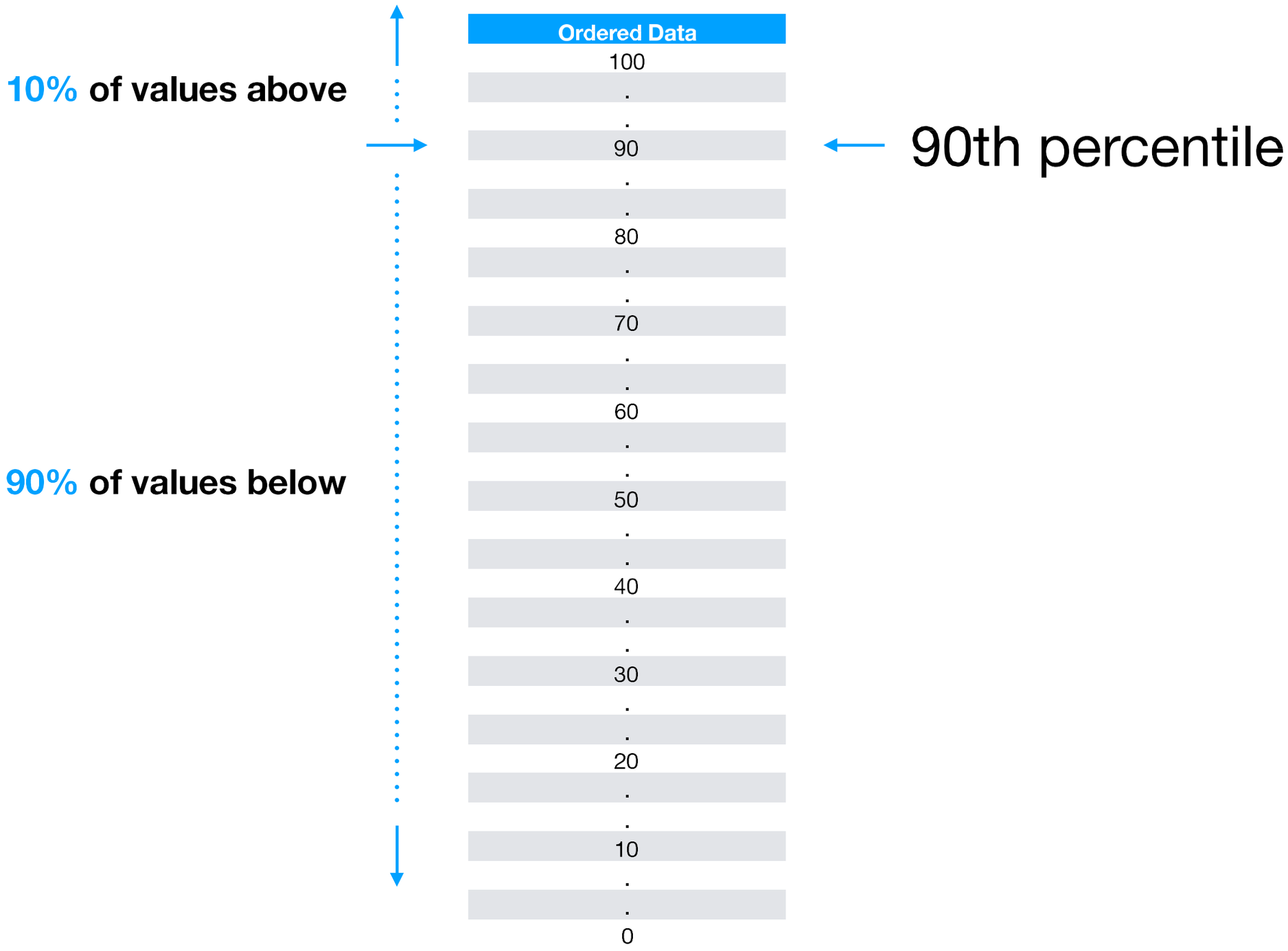
Ordered Data
100
.
.
90
.
.
80
.
.
70
.
.
60
.
.
50
.
.
40
.
.
30
.
.
20
.
.
10
.
.
0

Percentiles

Ordered Data
100
.
.
90
.
.
80
.
.
70
.
.
60
.
.
50
.
.
40
.
.
30
.
.
20
.
.
10
.
.
0

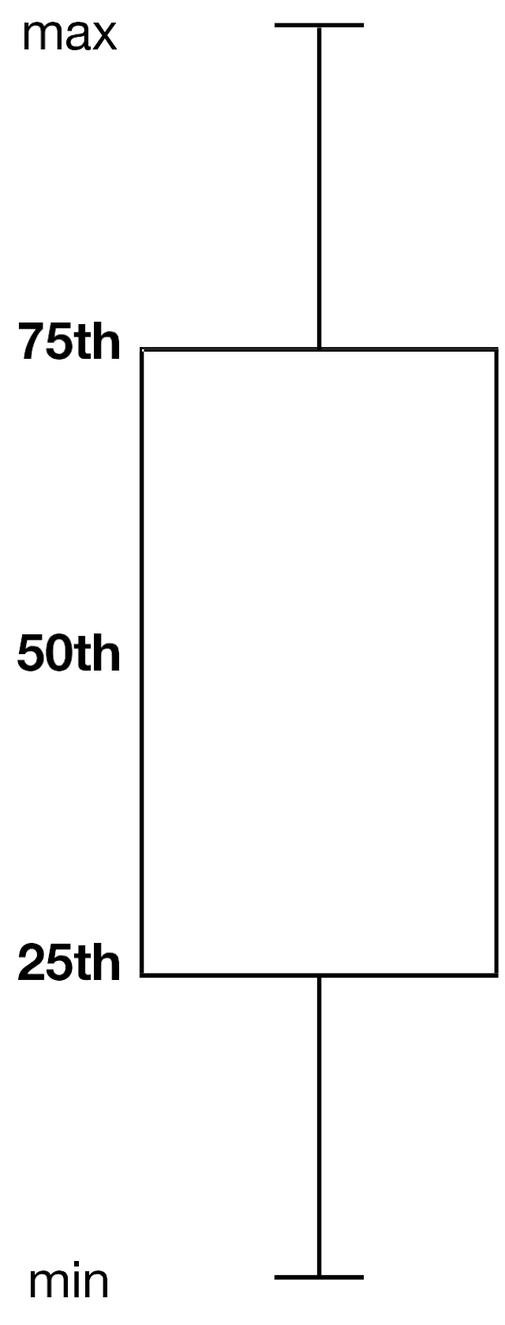
← 90th percentile

Percentiles



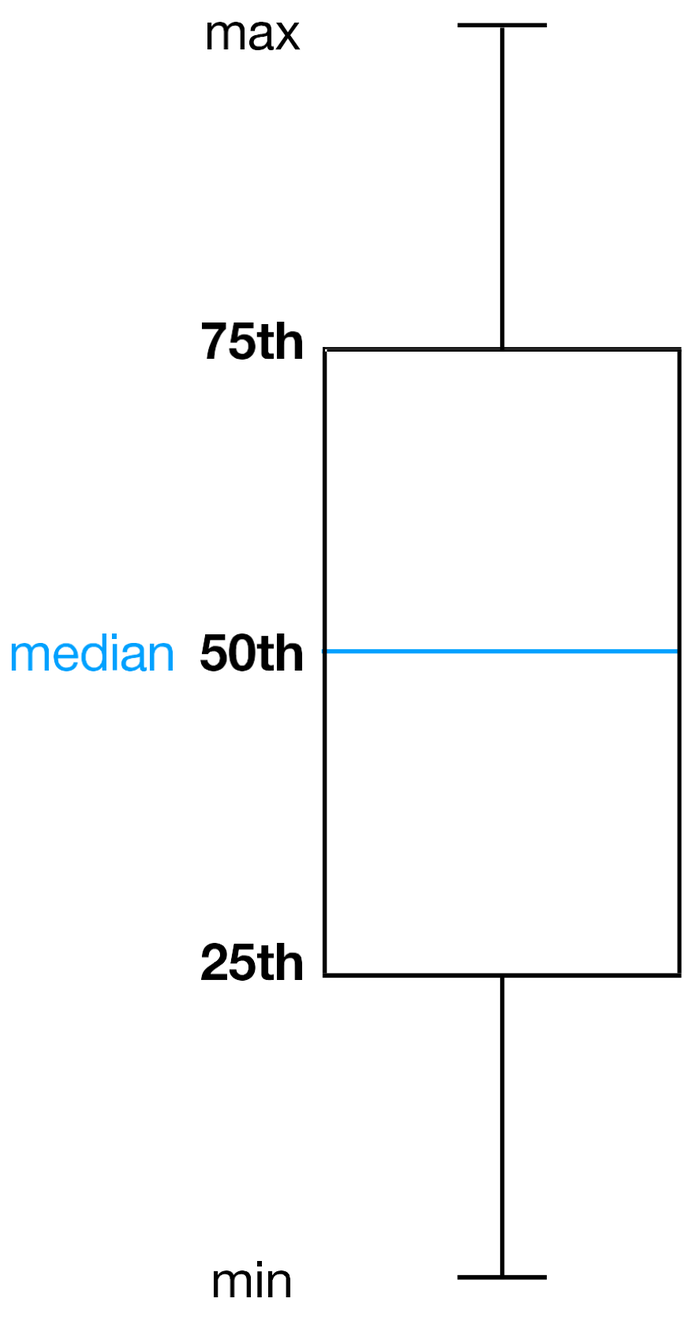
Quartiles

Ordered Data
100
.
.
90
.
.
80
.
.
70
.
.
60
.
.
50
.
.
40
.
.
30
.
.
20
.
.
10
.
.
0



Quartiles

Ordered Data
100
.
.
90
.
.
80
.
.
70
.
.
60
.
.
50
.
.
40
.
.
30
.
.
20
.
.
10
.
.
0



Quartiles

Ordered Data
100
.
.
90
.
.
80
.
.
70
.
.
60
.
.
50
.
.
40
.
.
30
.
.
20
.
.
10
.
.
0