National University of Computer & Emerging Sciences, Karachi

MT-2005 Probability and Statistics

Quiz 1

Question: Air Quality Standards

The number of days that selected U.S. metropolitan areas failed to meet acceptable air quality standards is shown below for 1998. Construct grouped frequency distributions using 5 classes and a histogram for the set of data. What you can say about the mean, median and mode of data from the graph.
Calculate the variance and S.D.

1998

43	76	51	14	0	10
20	0	5	17	67	25
38	0	56	8	0	9
14	5	37	14	95	20
23	12	33	0	3	45

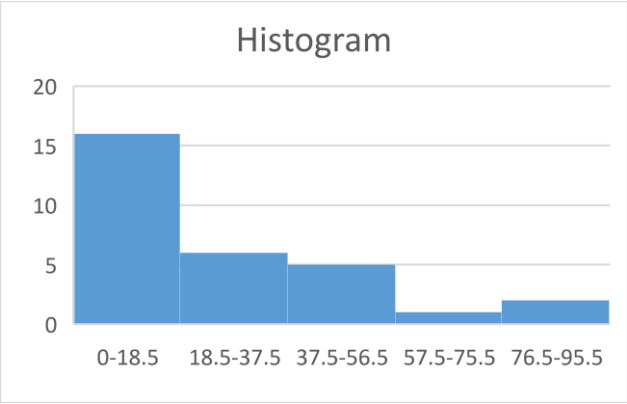
Solution:

Data Set of 1998

Lowest value = 0
Highest Value = 95
Range = 95
Class Width = 19

Class Limits	Class Boundaries	Frequency	m	$f * m$	$f * m^2$
0-18	0-18.5	16	9.1	145.6	1324.96
19-37	18.5-37.5	6	28	168	4704
38-56	37.5-56.5	5	47	235	11045
57-75	57.5-75.5	1	66.5	66.5	4422.25
76-95 Or 76 and above	76.5-95.5	2	86	172	14792
Sum		30		787.1	36288.21

Variance = 538.1
S.D = 23.2



Uni-modal and positive(right) skewed distribution
Mode > Median > Mean

The value of the mean is the largest, that of the mode is the smallest, and the value of the median lies between these two. The value of the mean is the largest in this case because it is sensitive to outliers that occur in the right tail. These outliers pull the mean to the right.

National University of Computer & Emerging Sciences, Karachi

MT-2005 Probability and Statistics

Quiz 1

Question: Protein Grams in Fast Food

The amount of protein (in grams) for a variety of fast-food sandwiches is reported here. Construct a frequency distribution using 6 classes. Draw a histogram using relative frequencies. Describe its shape, what you can say about the mean, median and mode of data from the graph. Calculate the variance and S.D.

23	30	20	27	44	26	35	20	29	29
25	15	18	27	19	22	12	26	34	15
27	35	26	43	35	14	24	12	23	31
40	35	38	57	22	42	24	21	27	33

Solution:

Lowest value = 12

Highest Value = 57

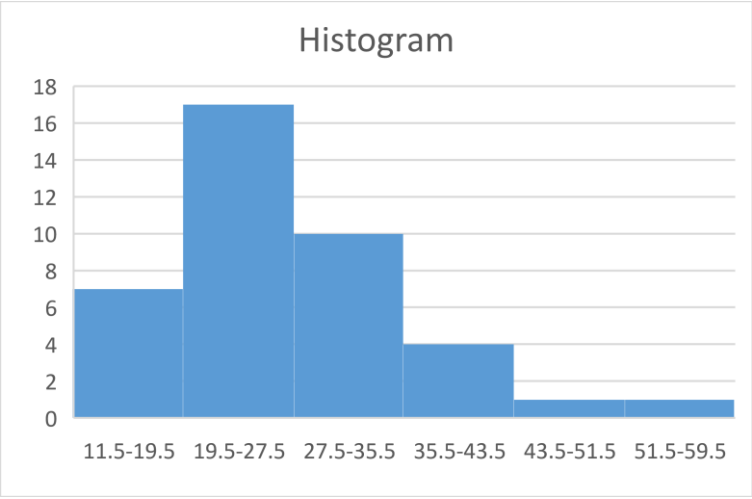
Range = 45

Class Width = 8

Class Limits	Class Boundaries	Frequency	m	$f * m$	$f * m^2$
12-19	11.5-19.5	7	15.5	108.5	1681.75
20-27	19.5-27.5	17	23.5	399.5	9388.25
28-35	27.5-35.5	10	31.5	315	9922.5
36-43	35.5-43.5	4	39.5	158	6241
44-51	43.5-51.5	1	47.5	47.5	2256.25
52-59	51.5-59.5	1	55.5	55.5	3080.25
sum		40	213	1084	32570

Variance = 81.9

S. D = 9.04



Conclusion:

The graph is unimodal and right skewed

Most of the data values are clustered in first three classes.

The mean is greater than median and the median is always greater than the mode.

National University of Computer & Emerging Sciences, Karachi

MT-2005 Probability and Statistics

Quiz 1

Question: Air Quality Standards

The number of days that selected U.S. metropolitan areas failed to meet acceptable air quality standards is shown below for 1998. Construct grouped frequency distributions using 5 classes. Calculate the mean, median, mode and quartiles of data. Construct a box plot.

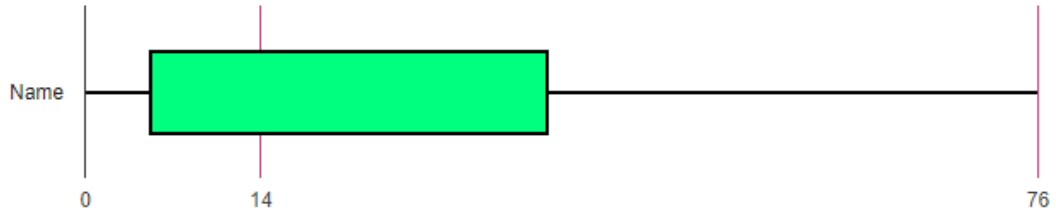
1998					
43	76	51	14	0	10
20	0	5	17	67	25
38	0	56	8	0	9
14	5	37	14	95	20
23	12	33	0	3	45

Solution:

Data Set of 1998

Lowest value = 0
Highest Value = 95
Range = 95
Class Width = 19
Mean = 26.2
Median Q2 =14
1st class is the mode class
Mode = 9.75
Q1 = 5
Q3=37
Outlier 95

Class Limits	Class Boundaries	Frequency
0-18	0-18.5	16
19-37	18.5-37.5	6
38-56	37.5-56.5	5
57-75	57.5-75.5	1
76-95 Or 76 and above	76.5-95.5	2
Sum		30



Conclusion:

The graph is positively skewed, have right tail. One outlier on the right.
A very large difference between the median and highest value of data set can be observed. There is one outlier in the data