**Topics: Descriptive Statistics and Probability**

1. Look at the data given below. Plot the data, find the outliers and find out

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Name of company** | | | **Measure X** | |
| Allied Signal | | | 24.23% | |
| Bankers Trust | | | 25.53% | |
| General Mills | | | 25.41% | |
| ITT Industries | | | 24.14% | |
| J.P.Morgan& Co. | | | 29.62% | |
| Lehman Brothers | | | 28.25% | |
| Marriott | | | 25.81% | |
| MCI | | | 24.39% | |
| Merrill Lynch | | | 40.26% | |
| Microsoft | | | 32.95% | |
| Morgan Stanley | | | 91.36% | |
| Sun Microsystems | | | 25.99% | |
| Travelers | | | 39.42% | |
| US Airways | | | 26.71% | |
| Warner-Lambert | | | 35.00% | |
|  | | |  | |
|  |  |  | |  | |  |  |  |  |  |
| x |  |  | | x-m | | x-m^2 |  |  |  |  |
| 24.23% |  |  | | -9.04% | | 0.008175 |  |  |  |  |
| 25.53% |  |  | | -7.74% | | 0.005993 |  |  |  |  |
| 25.41% |  |  | | -7.86% | | 0.00618 |  |  |  |  |
| 24.14% |  |  | | -9.13% | | 0.008338 |  |  |  |  |
| 29.62% |  |  | | -3.65% | | 0.001333 |  |  |  |  |
| 28.25% |  |  | | -5.02% | | 0.002521 |  |  |  |  |
| 25.81% |  |  | | -7.46% | | 0.005567 |  |  |  |  |
| 24.39% |  |  | | -8.88% | | 0.007888 |  | Mean = | 33.27% | 0.3327 |
| 40.26% |  |  | | 6.99% | | 0.004884 |  | variance= | 2.87% | 0.028 |
| 32.95% |  |  | | -0.32% | | 1.03E-05 |  |  |  |  |
| 91.36% |  |  | | 58.09% | | 0.337429 |  | Stdev = | 91.36% | 0.9136 |
| 25.99% |  |  | | -7.28% | | 0.005302 |  |  |  |  |
| 39.42% |  |  | | 6.15% | | 0.003781 |  |  |  |  |
| 26.71% |  |  | | -6.56% | | 0.004305 |  |  |  |  |
| 35.00% |  |  | | 1.73% | | 0.000299 |  |  |  |  |
|  |  |  | |  | |  |  |  |  |  |
|  |  |  | | Sum= | | 0.402005 |  |  |  |  |



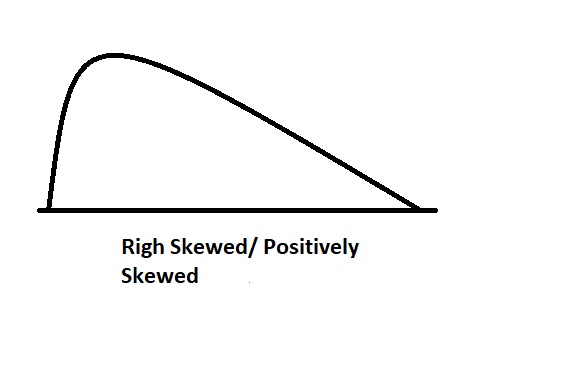
Answer the following three questions based on the box-plot above.

1. What is inter-quartile range of this dataset? (please approximate the numbers) In one line, explain what this value implies.

* IQR= 5to 12, 50% data lies in this range.

1. What can we say about the skewness of this dataset?

* Right skewed



1. If it was found that the data point with the value 25 is actually 2.5, how would the new box-plot be affected?

* 2.5 is in the starting of the boxplot it will not effect the representation but if the value changes then 2.5 will be neglected, and the box plot will be from 0 to 20.



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

* 5 -9

1. Comment on the skewness of the dataset.

* Right skewed

1. Suppose that the above histogram and the box-plot in question 2 are plotted for the same dataset. Explain how these graphs complement each other in providing information about any dataset.

* Mode is represented in histogram , median will be represented in Boxplot,
* 50% of data was lying 5-12,here it lies between 5-9 therefore data might lie between 5-9 range.
* A larger tail well occur on the edge.

1. AT & T was running commercials in 1990 aimed at luring back customers who had switched to one of the other long-distance phone service providers. One such commercial shows a businessman trying to reach Phoenix and mistakenly getting Fiji, where a half-naked native on a beach responds incomprehensibly in Polynesian. When asked about this advertisement, AT&T admitted that the portrayed incident did not actually take place but added that this was an enactment of something that “could happen.” Suppose that one in 200 long-distance telephone calls is misdirected. What is the probability that at least one in five attempted telephone calls reaches the wrong number? (Assume independence of attempts.

Probability of call being misdirected p = 1/200 = .005

No of attemped phone call n = 5

1-((199/200)\*(199/200)\*(199/200)\*(199/200)\*(199/200))

= 0.024751

1. Returns on a certain business venture, to the nearest $1,000, are known to follow the following probability distribution

|  |  |
| --- | --- |
| x | P(x) |
| -2,000 | 0.1 |
| -1,000 | 0.1 |
| 0 | 0.2 |
| 1000 | 0.2 |
| 2000 | 0.3 |
| 3000 | 0.1 |

1. What is the most likely monetary outcome of the business venture?
2. Is the venture likely to be successful? Explain
3. What is the long-term average earning of business ventures of this kind? Explain
4. What is the good measure of the risk involved in a venture of this kind? Compute this measure

solution

|  |  |  |
| --- | --- | --- |
| x | P(x) | x\*p(x) |
| -2,000 | 0.1 | -200 |
| -1,000 | 0.1 | -100 |
| 0 | 0.2 | 0 |
| 1000 | 0.2 | 200 |
| 2000 | 0.3 | 600 |
| 3000 | 0.1 | 300 |

Sum of (x\*p(x))=800

1. Max probability is 0.3 =2000
2. the venture likely to be

80% of profit &20% loss soo the venture is successful.

1. 800, Sum of (x\*p(x))=800 will be the long term earning of business

|  |  |
| --- | --- |
| x | P(x) |
| -2,000 | 0.1 |
| -1,000 | 0.1 |

It is the risk involved i.e .1+.1 =.2 20% risk