**Topics: Descriptive Statistics and Probability**

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

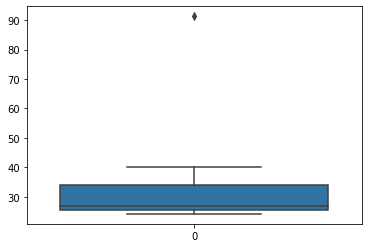
|  |  |
| --- | --- |
| **Name of company** | **Measure X** |
| Allied Signal | 24.23% |
| Bankers Trust | 25.53% |
| General Mills | 25.41% |
| ITT Industries | 24.14% |
| J.P.Morgan & Co. | 29.62% |
| Lehman Brothers | 28.25% |
| Marriott | 25.81% |
| MCI | 24.39% |
| Merrill Lynch | 40.26% |
| Microsoft | 32.95% |
| Morgan Stanley | 91.36% |
| Sun Microsystems | 25.99% |
| Travelers | 39.42% |
| US Airways | 26.71% |
| Warner-Lambert | 35.00% |

For given data:

**Mean**:33.27

**Variance**: 287.14

**Standard** **Deviation**:16.94





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.

**Answer: (Approximate estimated values)**

**Median = 7**

**Q1 = 5**

**Q3 = 12**

**IQR = 12-5 = 7**

**Minimum = Q1 - 1.5(IQR)**

**= -3.5**

**Maximum = Q3 + 1.5(IQR) = 22.5**

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

**Answer: Positively skewed or 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?

**Answer: Box plot is plotted by using the median of the data. Median is not affected by the lower extreme or higher extreme values.**

**In above box plot 25 is the outlier if it is 2.5 then median value moves to the very next positive data point of the data.**



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

**Answer: In given histogram, two peaked points are the same so mode=2nd,3rd bar in histogram**

1. Comment on the skewness of the dataset.

**Answer: Right skewed most data points are lied on the left side of the histogram**

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.

**Answer: Both the sets are right skewed**

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.)

Answer: 

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?

**Answer: - P (2000) = 0.3**

1. Is the venture likely to be successful? Explain

**Answer: Yes, P(x>0) =0.2+0.3+0.1= 0.6 (60% chance)**

**The Venture is likely to be successful**

1. What is the long-term average earning of business ventures of this kind? Explain

**Answer: µ = x1.p(x1)+x2.p(x2)+x3.p(x3)+x4.p(x4)+x5.p(x5)+x6.p(x6)**

**= (-2000) \*0.1+(-1000)\*0.1+0\*0.2+1000\*0.2+2000\*0.3+3000\*0.1**

**µ = 800**

1. What is the good measure of the risk involved in a venture of this kind? Compute this measure?

**Risk measure = P(-1000) + P(-2000) = 0.1 + 0.1 = 0.2**