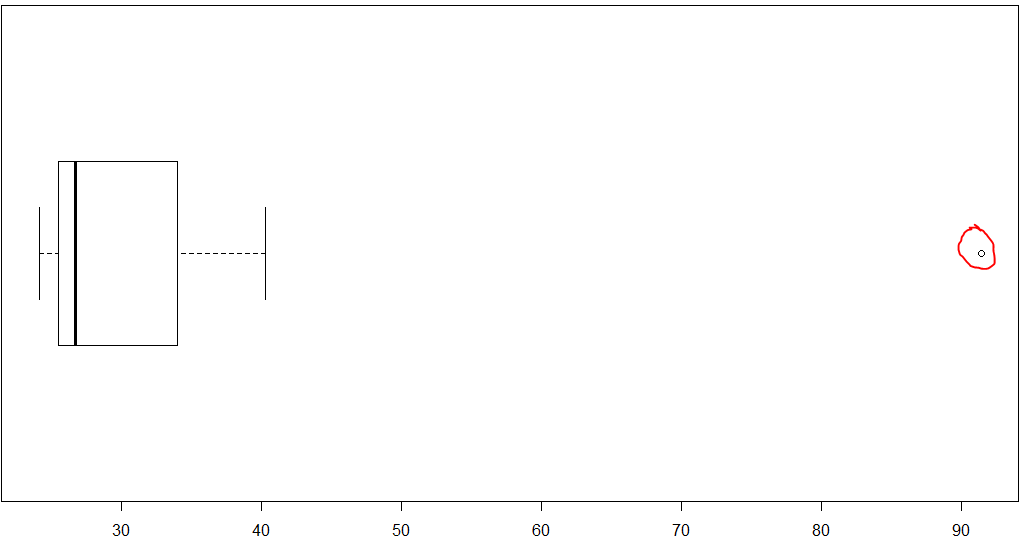
**Topics: Descriptive Statistics and Probability**

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% |

**Ans.**

1. ****
2. **Outlier = Morgan Stanley - 91.36%**
3. **33.27133**
4. **= 16.9454**
5. **= 287.1466**



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.
2. What can we say about the skewness of this dataset?
3. If it was found that the data point with the value 25 is actually 2.5, how would the new box-plot be affected?

**Ans.**

1. **IQR = 12-5=7;This implies middle portion of data is away of 7 units**
2. **Data followed skewness towards right**
3. **Median line will move little RIGHT (i.e., decreases) and no outliers will be there in boxplot.**



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?
2. Comment on the skewness of the dataset.
3. 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.

**Ans.**

1. **Mode lies between 4 to 8 range and there are 2 modes.**
2. **Data skewed towards right with few number of higher values**
3. **Histogram confirm that this data set follows Symmetric multi-modal distribution with skewness to the right while box plot confirms outlier, median,IQR,1st,2nd,3rd quartile, min, max. Using both plots we can identify central tendencys of data, outliers,distribution and etc..**
4. 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.)

**Ans.** 1 wrong call out of 200 is = (1/200)\*100=0.5%

Probability of 1 wrong call out of 5 = 0.5\*5=2.5%

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? 2000 (approx.) with highest probability
2. Is the venture likely to be successful? Explain Positive values probability=0.6 is higher than negative values and neutral value probability Hence, it will be successful in long term.
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

Ans.

c) Long-term average earning of business ventures = $800

|  |  |  |
| --- | --- | --- |
| p(x) | x | Expected value=p(x)\*x |
| 0.1 | -2,000 | -200 |
| 0.1 | -1,000 | -100 |
| 0.2 | 0 | 0 |
| 0.2 | 1000 | 200 |
| 0.3 | 2000 | 600 |
| 0.1 | 3000 | 300 |
|  |  | **$800** |

1. Variance is the good measure to find out risk involved in the venture.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| p(x) | x | XP | X^2 | (X^2)\*P |
| 0.1 | -2,000 | -200 | 4000000 | 400000 |
| 0.1 | -1,000 | -100 | 1000000 | 100000 |
| 0.2 | 0 | 0 | 0 | 0 |
| 0.2 | 1000 | 200 | 1000000 | 200000 |
| 0.3 | 2000 | 600 | 4000000 | 1200000 |
| 0.1 | 3000 | 300 | 9000000 | 900000 |
| **Expected Value** | | **800** | Sum of (X^2)\*P = | 2800000 |
|  |  |  | Variance(X)= Sum of (X^2)\*P - Expected value ^2 | 2160000 |
|  |  |  | Standard Deviation = sqrt(variance) | 1469.694 |