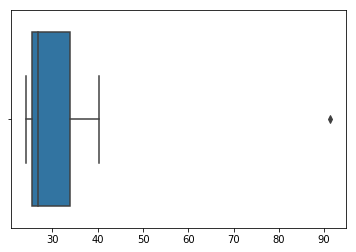
**Topics: Descriptive Statistics and Probability**

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% |

Solution-



**From x.mean()- 33.27133333333333**

**From x.std()- 16.945400921222028**

**From x.var()- 287.1466123809524**



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.

**Solution- By Calculation IQR**

**Q3-Q1=12-5= 7 where Q3(approximately 12)**

**It represents range which has 50% of the data points covered inside the plot.**

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

**Solution- Since the longer side of the box is to the right, therefore it is 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?

**Solution- It will not affect the boxplot, It will just start from 0 and end at 20, 2.5 will not be and outlier in this case.**



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

**Solution- From the given representation, it clearly show the mode is between 4 and 8.**

1. Comment on the skewness of the dataset.

**Solution- Since the tail trails towards right side of the histogram, it is 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.

**Solution- 1) Histogram provides frequency distribution, whereas boxplot provides Quantile distribution.**

**2) Outliers can be Determined easily with the help of Whiskers in boxplot, whereas in histogram it gives a rough plot of the data distributed in outliers.**

**3) Histograms are preferred to determine the underlying distribution of a data. Box plots on the other hand are more useful when comparing between several data sets.**

**4) Skewness and Symmetry is easier to understand in boxplot than histogram.**

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.)

**Solution- Probability of call misdirecting=1/200=**

**Probability of call not misdirecting=1-1/200=199/200**

**Number of calls=5**

**By using Binomial Distribution formula**

**P(x) = ⁿCₓpˣqⁿ⁻ˣ**

**n=5,**

**p=1/200**

**q=199/200**

**X=0, since we want none of the calls reaches the wrong number**

**Therefore the answer is 0.97524**

**But we want, at least one in five attempted telephone calls reaches the wrong number**

**Therefore, 1-0.97524=0.0247**

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?

**Solution-The most likely monetary outcome is P(x), the max value from table is**

**0.3 I.e. for P (2000). So most likely outcome is 2000.**

1. Is the venture likely to be successful? Explain

**Solution- the Profit of the venture has P(X>0) =0.6, means 60 % chance to earn profit,**

**Whereas the loss is of P(X<0) = 0.2, means of 20%.**

**Therefore the venture is likely to be successful.**

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

**Solution- BY calculating the expected value we can find the long term average earning**

|  |  |  |
| --- | --- | --- |
| x | P(x) | E(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 |
|  |  | Total= 800 |

**Therefore the Total Expected value (Ex) = 800$ (Long –term average earning)**

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

**Solution- P (-2000) + P (-1000) =0.2**

**So the Risk associated with the venture is 20%**