**Topics: Descriptive Statistics and Probability**

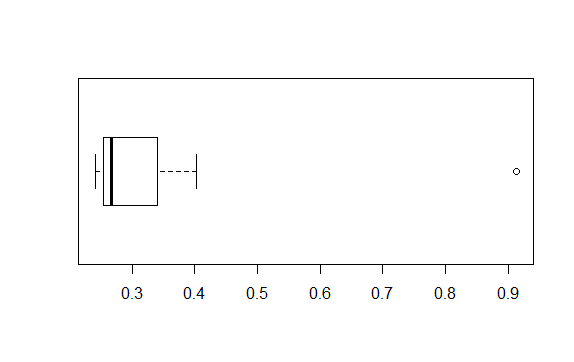
1. Look at the data given below. Plot the data, find the outliers and find out

|  |  |
| --- | --- |
| **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% |

First copy all the data in excel and use formulas in excel to find out all values

|  |  |
| --- | --- |
| Mean | 33.27% |
| standard deviation | 0.163708 |
| Variance | 0.0268 |

For finding outliers we have to use Jupyter Notebook and in that seaborn library





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.

**Ans. IQR = Q3 – Q1 = 12-5=7**

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

**Ans. It represents Positive Skewness**

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?

**Ans. The data point with the value 25 is suppose actually at 2.5 then it lies left side of the box in box plot. And it does not contain any outlier**



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

**Ans. Mode is 20. And lies in between 4-6 and 6-8**

1. Comment on the skewness of the dataset.

**Ans. It represents Positive Skewness**

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.

**Ans. Both diagrams shows the positive skewness of data**

**But**

**1) Boxplot provides outliers but it can not be provided by Histogram**

**2) Histogram provides frequency of data but it cant be provided by** **boxplot**

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.)

**Ans**. One call in 200 long-distance telephone calls is misdirected

So find the probability of at least one in five telephone calls reaches the wrong number

So, Probability of call misdirecting (p) = 1/200

Probability of call not misdirecting (q) = 1-1/200 = 199/200

Given, n=5

So, by using Binomial Distribution P(x) = ⁿCₓpˣqⁿ⁻ˣ

Where n=5, p=1/200, q=199/200

So for at least one misdirected calls probability is P(x=1)+P(x=2)+P(x=3)+P(x=4)+P(x=5)

To find this first we have to find P(x=0)

=  ⁵C₀(1/200)⁰(199/200)⁵⁻⁰

= 0.9752487531

at least one misdirected calls probability is

P(x=1)+P(x=2)+P(x=3)+P(x=4)+P(x=5)

**= 1-P(x=0)**

= 1- 0.9752487531

= 0.0247512469

**probability that at least one in five attempted telephone calls reaches the wrong number = 0.0247512469**

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?

**Ans . Highest probability is 0.3 corresponding to value 2000**

1. Is the venture likely to be successful? Explain

**Ans. Profit is corresponding to positive values is 0.2+0.3+0.1 = 0.6 that is 60%**

**And loss is corresponding to negative values is 0.1+0.1 = 0.2 that is 20%**

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

**Ans. Long term return =**

**(-2000\*0.1)+(-1000\*0.1)+(0\*0.2) +(1000\*0.2)+(2000\*0.3)+(3000\*0.1) = 800**

**Long term return average = 800/6 = 133.33**

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

**Ans. Good measure is positive returns.**

**So there will be 60 % profit it is more the loss which is 20%**