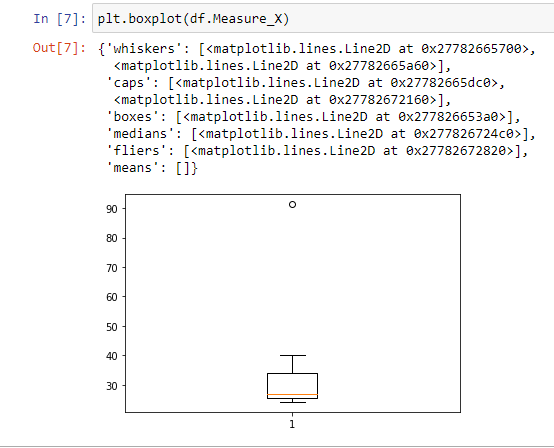
**Topics: Descriptive Statistics and Probability**

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)**

****

* Number of Outliers in the given data = 1
* Mean of the given data = 33.271
* Standard deviation of the given data = 16.945
* Variance of the given data = 287.147



**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. Lower Quartile Q1=5

Upper Quartile Q3=12

Inter Quartile Range IQR=Q3-Q1

IQR=12-5

IQR=7

1. Here we can observe that the given data have a **right skewness.** Because, we can see that the median line is situated towards lower quartile.
2. Then their won’t be any outliers and the point 2.5 lies in lower whisker



**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. Generally Mode decide the peak of the graph and By observing given histogram, mode will be at **the peak of the graph**
2. The data is right skewed. Because, the data is not normally distributed and if we can observe, the right side the graph have longer tile.
3. We can see that there is an outlier in the data and we can say that whether the data is normalized or not, by seeing the histogram graph. But, the data is right skewed
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)**

Probability to get 1 wrong call out of 200 = 1/200

= 0.005

Probability to get 1 wrong call in 5 phone call =1/5

=0.2

Probability to get at least one in five attempted phone calls reaches wrong number = 0.005/0.2

= 0.025

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**

**Ans)**

1. P (2000) have highest number of probability 0.3. So, the most likely monetary outcome of the business venture is x=2000, P(x)=0.3
2. Successful events are p (1000), p (2000), p (3000) => I. e; 0.2+0.3+0.1=0.6. 0.6>0.5, which is not close to 1 and not close to 0, but it close to 0.5, we can say it is moderate data.
3. (-2000) (0.1) +(-1000) (0.1) +(0) (0.2) +(1000) (0.2) +(2000) (0.3) +(3000) (0.1) = 800
4. Standard deviation of x=1870.82869

Standard deviation of P(x)=0.08165