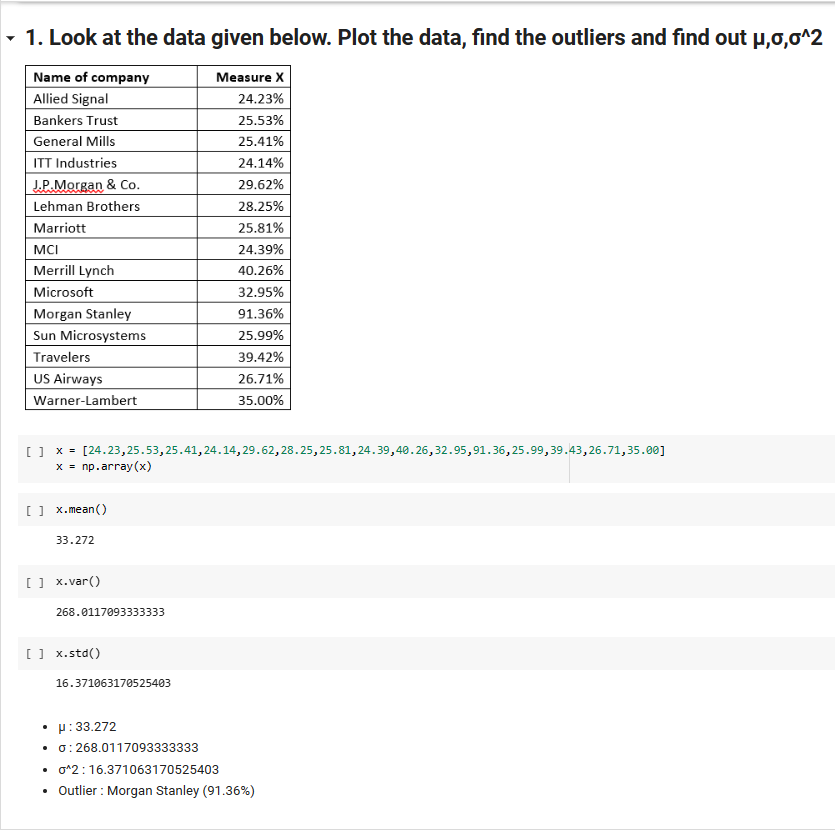
**Topics: Descriptive Statistics and Probability**

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% |





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:**

* Median : 7
* Q1 : 5
* Q3 : 12
* IQR : Q3-Q1 = 12 – 5 = 7

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

**ANS:**  Right (+Ve) 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:** Boxplot is built by the median, so it will not meagerly change anything but median may sift a bit.



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. There are two peaked points bar no 3 and 4, so we can say mode is someway around 4-8.
2. Right (+ve) skewness
3. Both graphs show that there an outlier near 25, min-max vale are same and boxplot say median is 7 which fall under histogram mode which is someway around 4-8.
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** If 1 in 200 long-distance telephone calls are getting misdirected.

Probability of call misdirecting = 1/200

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

The probability for at least one in five attempted telephone calls reaches the wrong number

Number of Calls = 5

n = 5

p = 1/200

q = 199/200

P(x) = at least one in five attempted telephone calls reaches the wrong number

P(x) = ⁿCₓ pˣ qⁿ⁻ˣ,

ⁿCₓ = !

x = 1 then after applying the values in the formula we get

P(1) = 0.0245037

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. The most likely monetary outcome of the business venture is $2000 because it have highest probability vale.
2. YES because probability of (0 to +x) : 0.2 + 0.2 + 0.3 + 0.1 = 0.8 or 80%
3. $800
4. import numpy as np

x = [-2000,-1000,0,1000,2000,3000]

prob = [0.1,0.1,0.2,0.2,0.3,0.1]

v = np.array([x,prob])

variance = np.var(v)

print('Variance: ',variance.round(3))

**Output :** 1520791.676

The variance is high, therefore the chances of risk is very high.