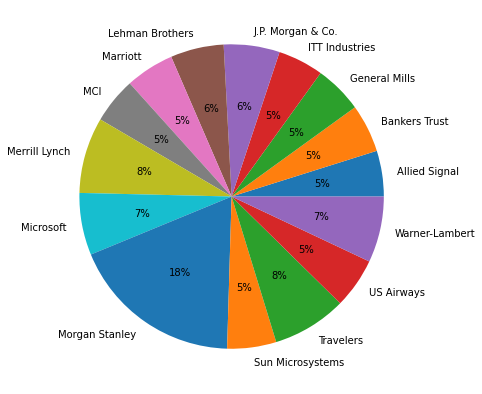
**Topics: Descriptive Statistics and Probability**

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% |

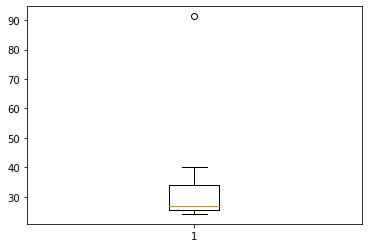
**Answer:**



Pie chart showing with name of company

* Mean(mu) = ∑( Measure X ) / Count of Measure X data = ( 499.07 / 15 ) = **33.27%**
* Variance(sigma sq.) =( (Xi−¯X)^2) / N = **0.02871**
* Standard Deviation(sigma) = **0.16945401**

**Box-Plot**



Here, boxplot shows outlier at pt. **91.36%**

* #For code section refer Assignment no.2[Set1].ipynb file.

2.



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.

* **Inter-quartile range = (5th quartile) – (3rd quartile)**
* IQR = Q5- Q3 = 12 – 5
* IQR = 7

***Inter-quartile range is 7***.

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

* From the given boxplot we can see that data is shifted towards right so the dataset is **positively 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?

* If we found that the data point with the value 25 is actually 2.5 then there will be no outlier in given plot and the data will be normal distributed and will follow **normal distribution.**

3.



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

* The mode will lie **Between 4 to 8**

1. Comment on the skewness of the dataset.

* The given dataset is **Positively 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.

* From above both graph from Q2 shows that both are positively skewed. Boxplot shows the outlier and median. Whereas, histogram shows mode of data.

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:**

Total Calls = 200

After 1 misdirected call total correct calls =199

Probability for 1 misdirected call is 1/200

Probability for correct call is 199/200

Let A ,B ,C, D, E will be 5 calls

Probability that none of the call reaches the wrong no. = P(0) = 5C0(1/200)0(199/200)5

P (1 misdirected call out of 5 calls) = 1  -  (probability that none of the call reaches the wrong no.)

= 1  - P(0)

= 1 - ⁵C0(1/200)⁰(199/200)⁵

= 1 – (199/200)5

= 0.02475

**Probability that 1 out of 5 calls are misdirected is 0.02475.**

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?

* Most likely momentary outcome of business venture is **$ 2000** because it has highest probability of **0.3.**

1. Is the venture likely to be successful? Explain

* **Yes**, because venture has positive profit i.e.,

P(x>1000)+P(x>2000)+P(x>3000) = 0.2+0.3+0.1 = **0.6.**

Which means venture has 80% chances of profit. So, business will be successful.

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

* E[x] = (-2000\*0.1) +(-1000\*0.1) +(1000\*0)+(1000\*0.2) + (2000\*0.3) + (3000\*0.1)
* E[X] = -200 + (-100) + 0 + (200) + (600) + (300)
* **E[x] = 800**

Therefore, Long-term average earnings of business ventures of this king is 800.

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

* The good measure of the risk involved in a venture of this kind depends on Measure of variance.
* Var (x) = E(X^2) –(E(X))^2 = 2800000 – 800^2 = 2160000

**Variance is 2160000 Which shows risk is high.**