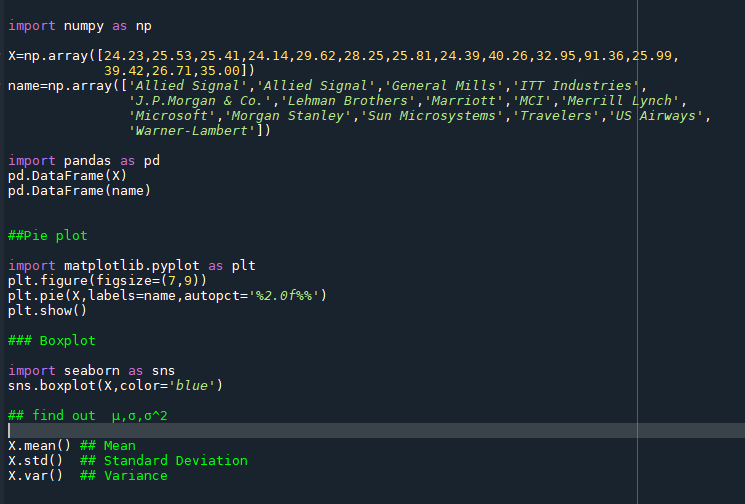
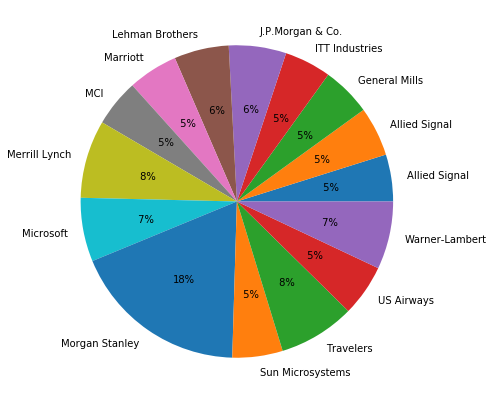
**Topics: Descriptive Statistics and Probability**

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% |

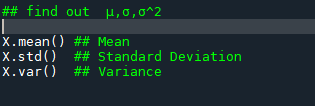
**Ans.**

****



Chart, box and whisker chart

Description automatically generated





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. Approximately(First Quantile Range) Q1=5

(Third Quantile Range) Q3=12,

Median(Second Quartile Range) Q2=7

(Inter-Quartile Range) IQR= Q3-Q1= 12 -5=7

Second Quartile Range is the Median Value

1. Right-Skewed median is towards left side it is not normal distribution.
2. In that case there will be no Outliers on the given dataset because the outliers of the data have a positive skewness as it will reduce, and the data will be normally distributed.



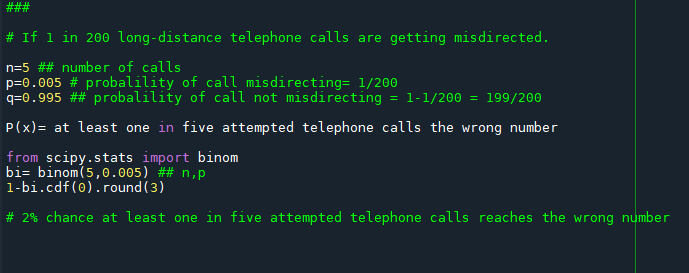
Answer the following three questions based on the histogram above.

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. The mode of this data set lies in between 5 to 10 and approximately between 4 to 8.
2. Right Skewed. Mean>Mean>Mode.
3. They both are right-skewed and both of them contains outliers the median can be easily visualized in the box plot where the histogram mode is more visible.
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.)**



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$. As far 2000$ the probability is 0.3 which is maximum as compared to others .
2. Yes, the probability that the venture will make more than 0 or a profit p(X>0)+p(X>1000)+p(X>2000)+p(X=3000) =0.2+0.2+0.3+0.1=0.8. It states that there is an 80% chance for the venture to be making a profit.
3. The long-term average is expected value= Sum(X\*P(X))=800$ which means that on an average the returns will be +800$.
4. The good measure of risk is involved in a venture of this kind will depends on the variability in the distribution. Higher variance means more chances of risk:

VAR(X)=E(X^2)-(E(X))^2

2800000-800^2

2160000