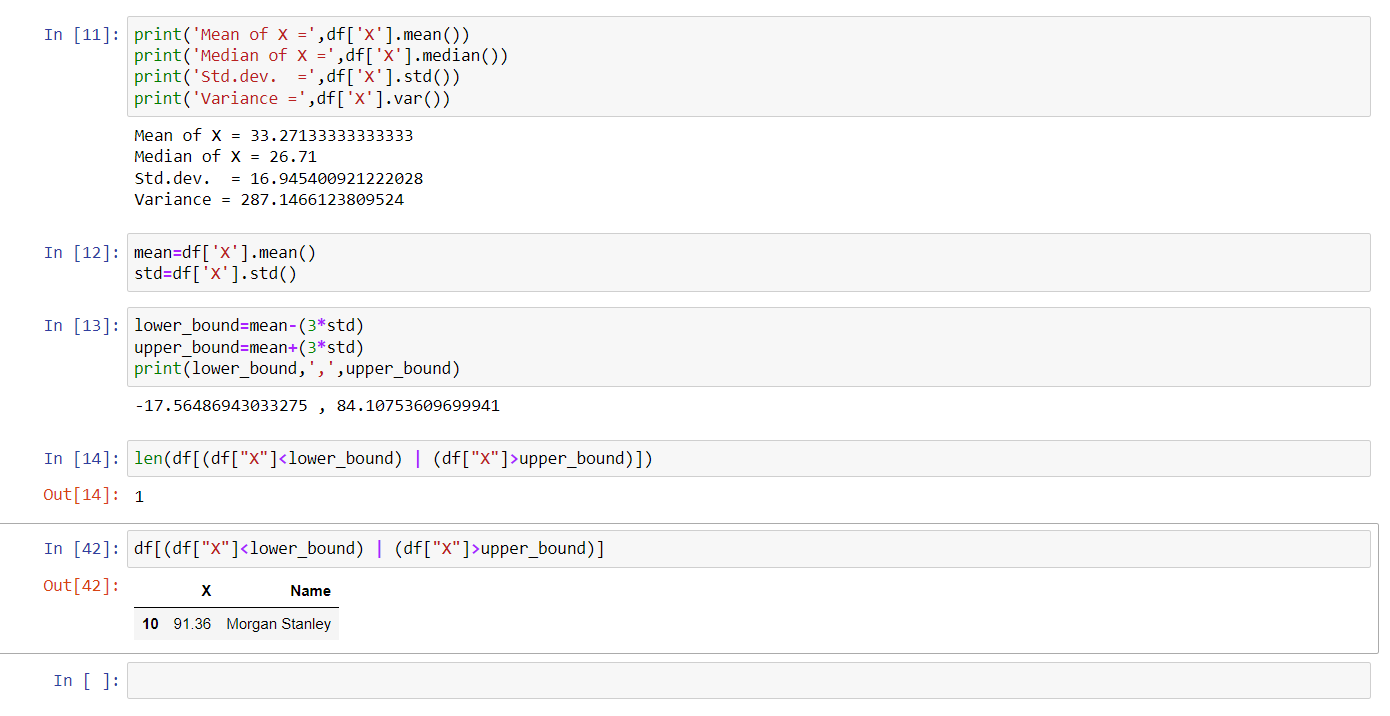
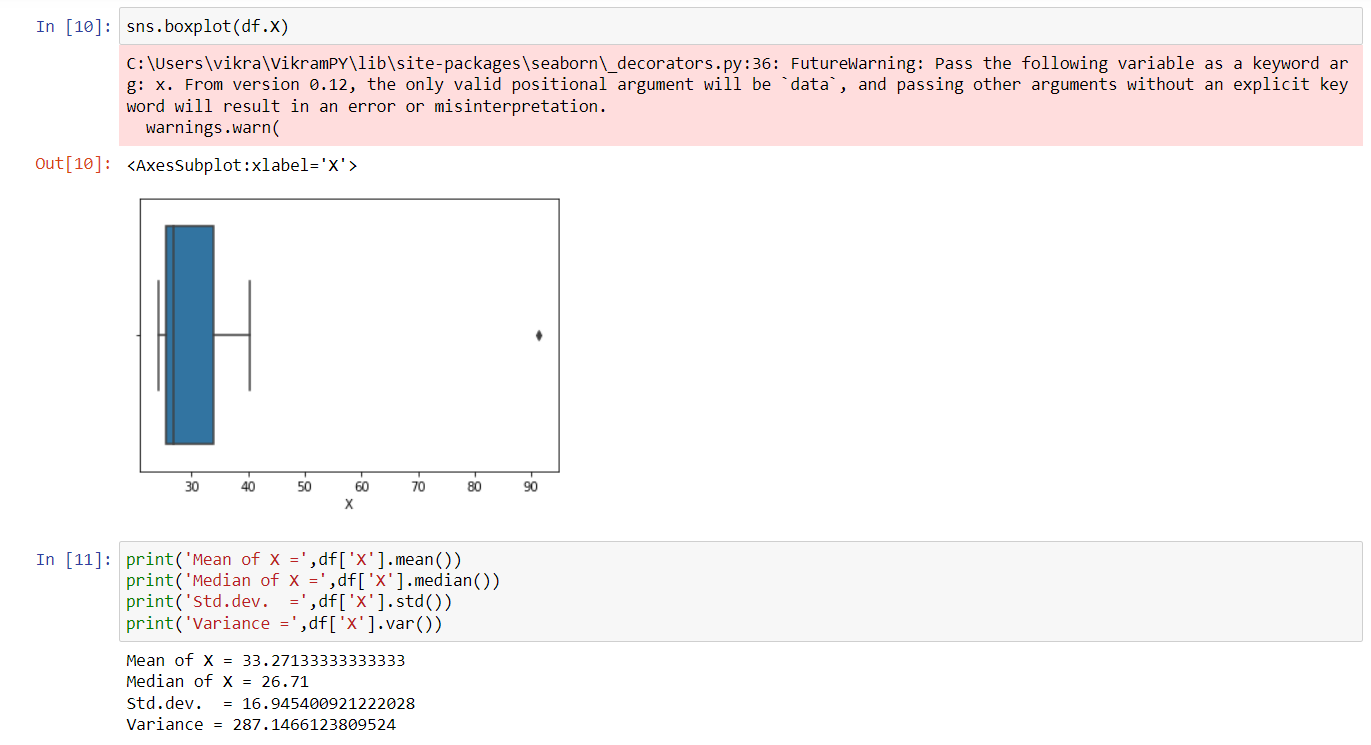
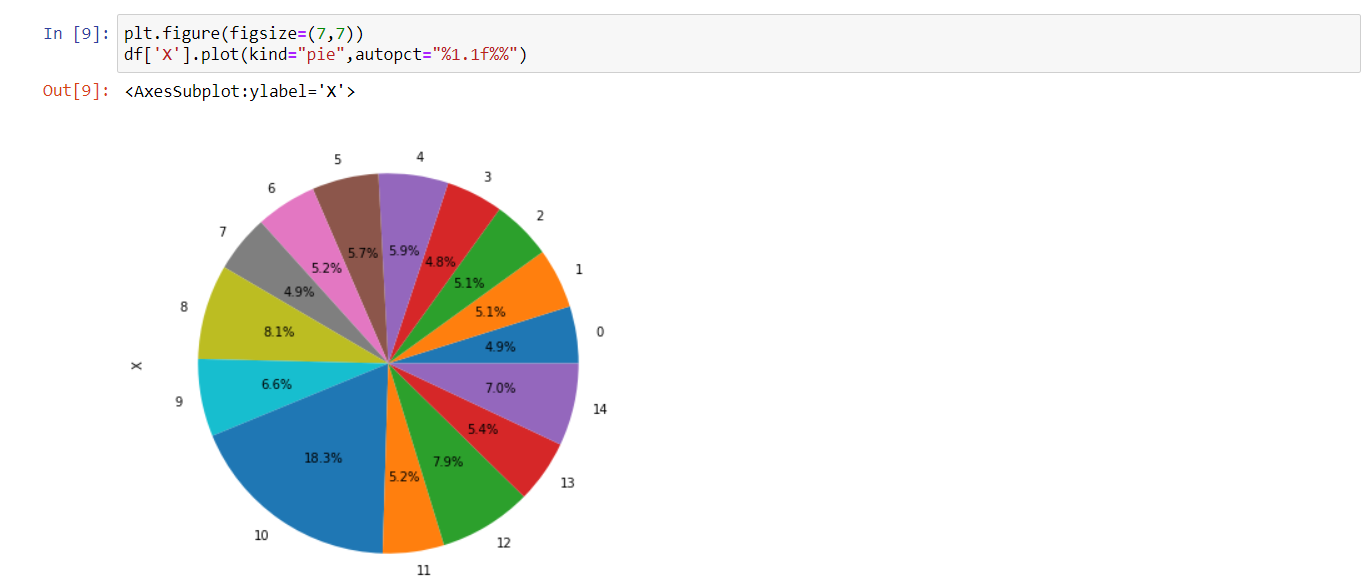
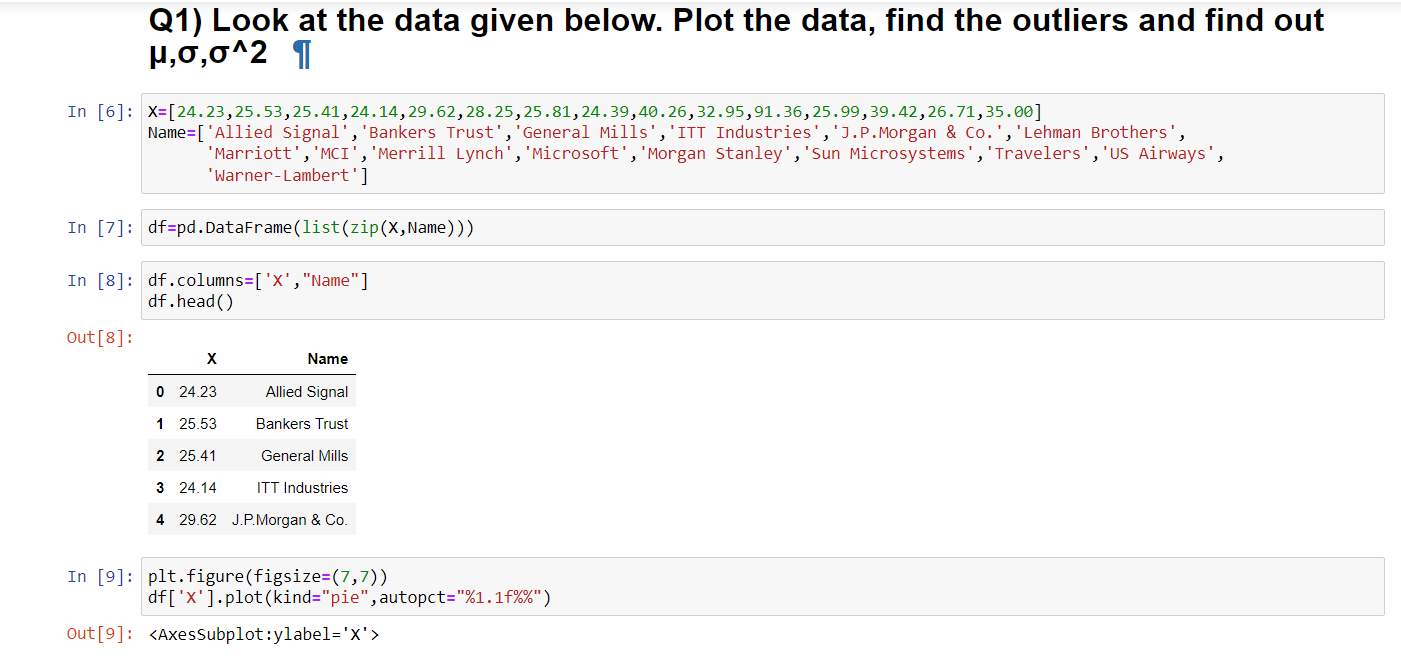
**Topics: Descriptive Statistics and Probability**

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.**



**Plz refer Data Science assignment 2 python file.**



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.**

**Q1=5 and Q3=12**

**IQR=Q3-Q1=12-5=7**

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

**ANS.**

**Data is not normally distributed , it is Right Skewed data**

**One outlier on 25**

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.**

**If it was found that the data point with the value 25 is actually 2.5 then there is no outlier on the given data ,because of outlier the data is Right skewed it will reduce and will be normally distributed**



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

**ANS.**

**Mode lies between 5 -10 (Approximately)**

1. Comment on the skewness of the dataset.

**ANS.**

**Data is Right skewed**

**Mean>Median>Mode**

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.**

**Mode is easily visible in Histogram and median is easily visible in boxplot.**

**Both the data is right skewed and both have outliers.**

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.**

**P(Call Misdirecting)=P=(1/200)**

**P(call not Misdirecting)=p=(199/200)**

**n=5**

**At least one in five attempted telephone calls reaches the wrong number**

**=1-(p(none of the call reaches the wrong no.))**

**=1-[(199/200)x(199/200)x(199/200)x(199/200)x(199/200)]**

**=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 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| x | P(x) | Profit P(X) | Loss P(X) | X\*P(x) | X^2\*x |
| -2,000 | 0.1 |  | 0.1 | -200 | 400000 |
| -1,000 | 0.1 |  | 0.1 | -100 | 100000 |
| 0 | 0.2 | 0.2 |  | 0 | 0 |
| 1000 | 0.2 | 0.2 |  | 200 | 200000 |
| 2000 | 0.3 | 0.3 |  | 600 | 1200000 |
| 3000 | 0.1 | 0.1 |  | 300 | 900000 |
|  |  | 0.8 | 0.2 | =800 | 2800000 |

1. What is the most likely monetary outcome of the business venture?

**ANS.**

**Most likely monetary outcome of the business venture is 2000 as it has maximum probability =0.3**

1. Is the venture likely to be successful? Explain

**ANS.**

**Venture is likely to be successful because expected value is +ve(800)**

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

**ANS.**

**The long-term average is Expected value=Sum(X\*P(X))=800.**

**Average will be 800.**

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

**ANS.**

**Var =( (X^2)\*X)-( X\*P(x))^2**

**=2800000-(800)^2**

**√var =2160000**

**Std =1470**

**Var is high so risk is high**