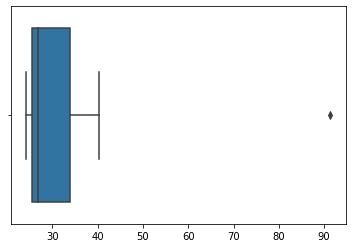
**Topics: Descriptive Statistics and Probability**

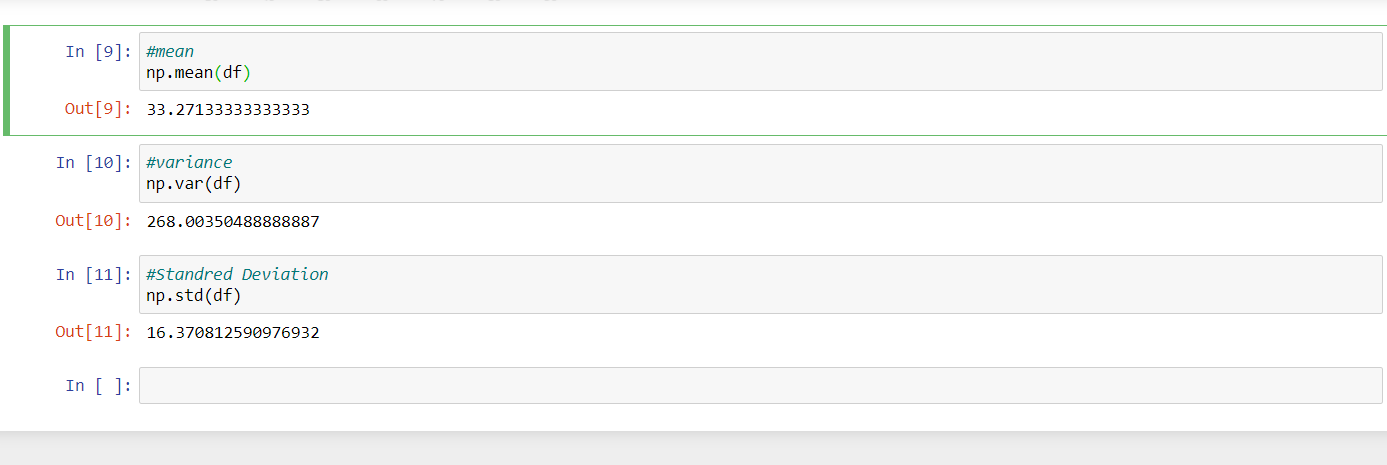
1. Look at the data given below. Plot the data, find the outliers and find out

|  |  |
| --- | --- |
| **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% |

df1=pd.Series([24.23,25.53,25.41,24.14,29.62,28.25,25.81,24.39,40.26,32.95,91.36,25.99,39.42,26.71,35.00])

sns.boxplot(df1)







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.

In boxplot IQR is in between upper quartile and lower quartile .

In this boxplot approximately lower quartile value is 5 and lowerquartile value is 12.

Therefore,IQR = upper quartile – Lower quartile

= 12 – 5

= 7

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

This is positive skewness or Right side skewness.

Therefore,No normal distribution seen

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?

Approximately the range of upper extream is 0 and lower Extream is 19.

Therefore, at a value 25 that point found an outlier but its actual value is 2.5

Is in range then no outlier are seen.



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

Mode is a most frequent number.

In this histogram mode is lie approximately 4 to 8.

1. Comment on the skewness of the dataset.

For above histogram Right skewness ie positive skewness are found.

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.

Above both graph Boxplot and Histogram are Right Skewed ie positive skewed.

Outlier are seen in both graph at a point of value 25.

In Boxplot median are visualized easily and in Histrogram mode are easily visualized.

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 commerciaal 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.)

If one in 200 long-distance telephone calls is misdirected

Probability of misdirected call = 1/200

Probability of not misdirected call = 1-1/200 = 199

Number of call (n)=5

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

P(x) = ⁿCₓpˣqⁿ⁻ˣ

P(1) = 1- P(0)

= 1- ⁵C₀(1/200)⁰(199/200)⁵⁻⁰

= 1- (199/200)⁵

= 1-0.97524875

=0.02475125

Therefore, the probability that at least one in five attempted telephone calls reaches the wrong number is 0.2475125

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?

The most likely monetary outcome of the business venture is 2000$ of a maximum probability 0.3.

1. Is the venture likely to be successful? Explain

Probability of success = 0.6[0.2+0.3+0.1]

So ,60% chance that the venture would be successful

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

The long-term average earning of business ventures of this kind is

(-2000\*0.1) + (-1000\*0.1) + (0\*0.2) + (1000\*0.2) + (2000\*0.3) + (3000\*0.1)

= 800$

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

A good measure to evaluate the risk would be variance and standard deviation of the variable x

Variance =

= 3500000

Standred deviation =)

= 1870.83

The value of standred deviation is 1870$

is considered along with the average returns of $ 800 indicates that this venture is highly risky