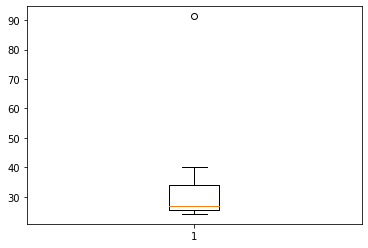
**Topics: Descriptive Statistics and Probability**

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% |

import matplotlib.pyplot as plt

plt.boxplot(data)

#outlier is >90 i.e Morgan stanly=91.36%

Mean: *)* np.round(data.mean(),2): 33.27

Variance: ()np.round(data.var(),2): 287.15

Standard deviation: ()np.round(data.std(),2): 16.95



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.

Q1(First Quartile Range)=5,Q3(Third Quartile Range)=12

Q2(Second quartile range)=7 (median)

IQR(Inner Quartile Range)=Q3-Q1=7, 7 is also Q2

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

The dataset is right skewed as the whisker range of maximum value is more than lower whisker range. and it is not normal distribution.

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 data point 25 is 2.5 then there will be no outlier and skewness will reduce and resulting distribution will be normal.



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

The mode of the dataset lies between 4 to 8.

1. Comment on the skewness of the dataset.

The above data set 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.

The above histogram and que2 box plot both are right skewed (positive skew)

Both have outlier at value=25.

the median can be easily visualized in box plot where as in histogram it is more visible.

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.)

probability of call misdirecting = 1/200

probability of call not misdirecting = 1 - 1/200 =199/200

if at least one in five attempted telephone calls reaches the wrong number Number of Calls then other are rightly placed so,

p(x)= at least one in five attempted telephone calls(where x is >=1)

P(x) = ⁿCₓ pˣ qⁿ⁻ˣ

P(x) = (nCx) (p^x) (q^n-x) # nCr = n! / r! \* (n - r)!

P(1) = (5C1) (1/200)^1 (199/200)^5-1 = 0.0245037

P(2) = (5C2) (1/200)^2 (199/200)^5-2 = 0.0002463

P(3) = (5C3) (1/200)^3 (199/200)^5-3 =1.23753125e-06

P(4) = (5C4) (1/200)^4 (199/200)^5-4 = 3.109375e-09

P(5) = (5C5) (1/200)^5 (199/200)^0 = 3.1250000000000005e-12

P(x)=p(1)+p(2)+p(3)+p(4)+p(5)= 0.0248

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 outcome is 2000 because p(2000) is maximum among all monetary outcome.

1. Is the venture likely to be successful? Explain

Yes venture will be successful because p(x>0)=0.1+0.2+0.3=0.6 i.e 60%

There is 60% chance of venture getting yield more than $1000

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

Weighted average=x\*p(x) = 800 , so in long-run this business would be earning around 800 (including loss and profit).

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

Good measure of loss =p(x=-2000) +p(x=-1000)

=0.1+0.1=0.2 which is 20 %