**Topics: Descriptive Statistics and Probability**

1. Look at the data given below. Plot the data, find the outliers and find out (answer in Descriptive Statistics Probability q1.ipynb)

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

24.14%

24.23%

24.39%

25.41%

25.53%

25.81%

25.99%

26.71%

28.25%

29.62%

32.95%

35.00%

39.42%

40.26%

91.36%

N=number of data point

Lower Quartile (Q1) = (N+1) \* 1 / 4

(15+1)\*1/4=4thposition = 25.41%

Upper Quartile (Q3 )= (N+1) \* 3 / 4

(15+1)\*3/4=12thposition= 35.00%

IQR = Q3 – Q1

=35.00%-25.41%

Lower outlier=9.59(any value is lower than this is outlier)

IQR = Q3 + Q1

=35.00%+25.41%

=60.41(any value is higher than this is outlier)

so the conclusion is Morgan Stanley 91.36%

population mean=

*Standard deviation=*

*population variance.=*

(done in jupyter notebook)

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.

Ans iqr is approx. is 12-5=7 and iqr is range between (5)q1+(6)q2+(12)q3=iqr (pink box)

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

Ans this its not normal distribution and it is Right-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?

Ans no Outliers because of the outlier the data will have positive skewness it will normal distributed



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

Ans lie between 5 to 10 approximately

1. Comment on the skewness of the dataset.

Ans Right-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.

Ans : both will have right-skewed 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 :- Suppose that one in 200 long-distance telephone calls is misdirected

probability of call misdirecting = 1/200 Probability of call not Misdirecting = 1-1/200 = 199/200 The probability for at least one in five attempted telephone calls reaches the wrong number Number of Calls = 5 person = 1/200 = 199/200 P(x) = at least one in five attempted telephone calls reaches the wrong number 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 P(1) = 0.0245037

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 monetary outcome of the business venture is 2000 $

as it has maximum probability = 0.3

1. Is the venture likely to be successful? Explain

Expected value = ∑E(X)P(X) = 800

venture is likely to be successful as Expected value is + ve = 800 $

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

long-term average earning of business ventures = 800 $ which means on an average the returns

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 the Variability in the distribution. Higher Variance means more chances of risk Var (X) = E(X^2) –(E(X))^2 = 2800000 – 800^2 = 2160000