**Assignment No .2 SET 1**

**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%** |

**Solution**:

Import pandas as pd

Import numpy as np

Import matplotlib.pyplot as plt

data=(24,23,25.53,25.41,24.14,29.62,28.25,25.1,24.39,40.26,32.95,91.36,25.99,39.42,26.71,35.00)

df=pd.DataFrame(data)

df.mean()

df.var()

df.std()

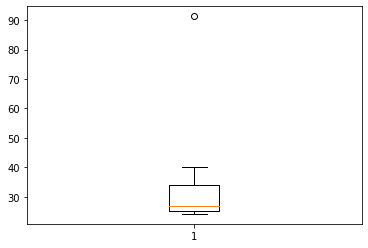
box\_plot=plt.boxplot(df[0]

µ = mean= 33.224

σ =Standard deviation= 16.968707

σ2 = Variance =287.937011

Boxplot:



Here we can observe that there is one outlier in the dada.

Q.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.**
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. Inter- quartile range=Q3-Q1

=12-5

=7

Interquartile renge suggest us how spread out the middle 50% of our data is.

1. The Q2 shifted towards left side (Q1).So we can say that is positively skewed.
2. As we see that the range of data is 0 to 20 . If 2.5 points will be added in the data then there will be no outlier in the data

Q.3



**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. Mode lies between 4 and 7
2. It is positively skewed.
3. Box plot show outlier in the dataset.

Boxplot gives us medium and Histogram gives us mode.

From both we get information about skewness.

**Q.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.)**

**SOLUTION:**

No of attempts n=5

Probability of call is misdirected=1/200

Rcode:

>n=5

>p=1/200

> pro=dbinom(0,n,p);pro #p(x=0)

[1] 0.9752488

P(x>=1)=1-p(x=0)

=0.02475125

Therfore probability that at least one in five attempted telephone calls reaches the

Wrong Number is 0.02475

**Q.5 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?**

**ANS:** The most likely monetary oytcome of the business venture is $2000

1. **Is the venture likely to be successful? Explain**

**ANS:** Yes, Because there are higher chances of positive returns. As we see in table 0.3+02+0.1 = 0.6\*100 = 60%

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

**ANS:** long-term average earning of business ventures = 800$

By, x\*p(x) = (-2000\*0.1)+(-1000\*0.1)+(0\*0.2)+(1000\*0.2)+(2000

\*0.3)+(3000\*0.1)

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

**ANS:**  large value in the standard deviation of the variable x show that there is highriskinvolved in the venture. Var = 3.500000

Sd = 1870.83