**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% |

Code:

import pandas as pd

importnumpy as np

df=pd.read\_csv("C:/Users/ashiq/Desktop/csv/assignment\_2/csv\_file (3).csv")

df.iloc[:,2].plot(kind='box') ## from box plot we got one outlier greater than 90

a=[df.iloc[:,2]]

b=df.describe()

c=df.var()

a1=df[df['Measure Xpercent']>90] ##taking the outlier value

Outlier: Morgan Stanley

Mean: 33.2713

std:16.9454

Variance: 287.147



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 : 12-5=7

The values are concentrated between q3 and q1

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

Positive 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?

It will act as an outlier itselfwith value at 2.5



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

6

1. Comment on the skewness of the dataset.

Positive 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.

Skewness

Median

Type of distribution

outlier

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:

Independent event probabily is the product of probability

P(misdirected)= 1/200

P(correct)=199/200

P(atleast one )=1- (P(correct)^5)

= .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 |

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

Variable with high probability: 2000

1. Is the venture likely to be successful? Explain

For that calculate the return more than 0 =(.2\*1000 + .3\*1000 + .1\*3000)=1100

the return less than or equal to 0=(.1\*2000 + 1000\*.1 +0\*.2)=-300

yes,the return more than 0 is greater

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

Mean=800

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

Std= 1469.694