HW 1 - Solutions

Problem 1

Here are the descriptive statistics, as computed by Minitab:

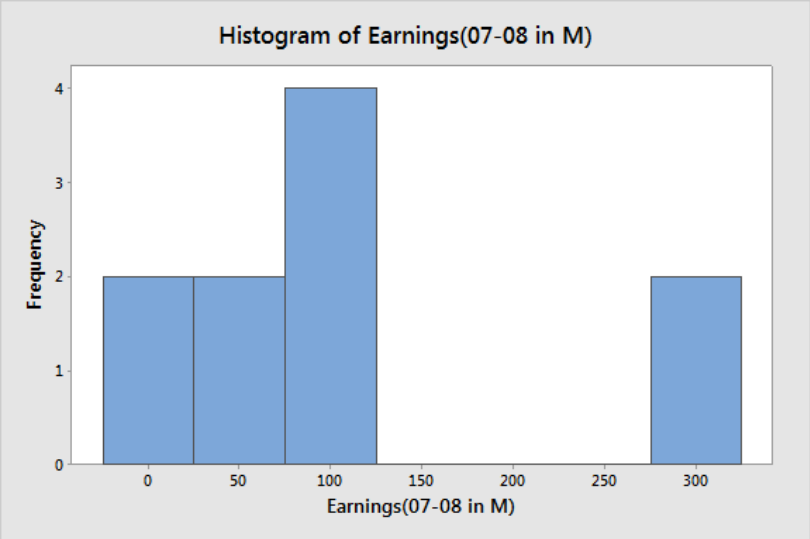
Descriptive Statistics: Earnings(07-08 in M)

Variable	N	N*	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3	Maximum
Earnings(07-08 in M)	10	0	112.3	31.2	98.5	14.0	42.5	81.0	155.0	300.0

The mean, median, and inter-quartile range are:

mean	112.3
median	81.0
IQR	155.0 - 42.5 = 112.5

The mean is greater than the median because it is heavily influenced by the extreme observations (in this case, the two large earnings for J.K. Rowling and Oprah Winfrey). We can see that these observations are extreme by looking at the histogram:



After removing these observations, the descriptive statistics are as follow:

Descriptive Statistics: Earnings(07-08 in M)

Variable	N	N*	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3	Maximum
Earnings(07-08 in M)	8	2	68.5	13.6	38.4	14.0	27.5	76.0	106.8	115.0

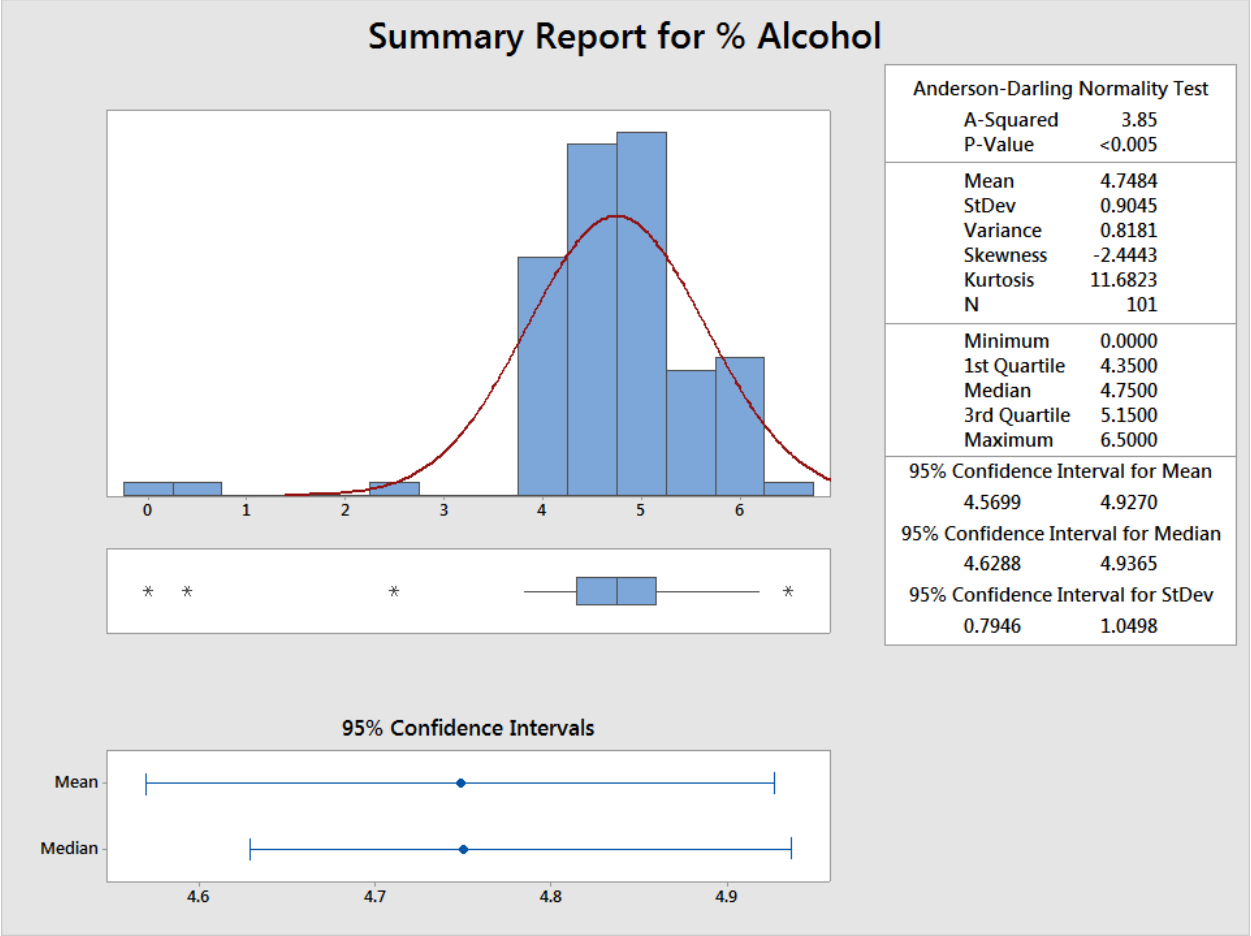
The updated statistics are:

mean	68.5
median	76.0
IQR	106.8 - 27.5 = 79.3

The mean changed dramatically, and the IQR changed, but not as much. The median was affected, but only slightly. Typically, the mean is sensitive to extreme observations, but the median and IQR are not. In this instance, since the sample size is relatively small, removing the two observations had a big affect on the IQR; if the sample size were larger, then we would not expect as big a change.

Problem 2

1. Here is the graphical summary:



This does not appear normal. The data are skewed to the right, and there are a few potential outliers in the left part of the distribution.

2. The two left-most outlying values are Sam Adams Light (0.00) and O'Doul's (0.40).

3. The z value for Sam Adams Light is (0.00 - 4.7484) / (0.9045) = -5.2; the z-value for O'Doul's is (0.40 - 4.7484) / (0.9045) = -4.8. Based on the empirical rule, it is not plausible that the data came from a normal distribution (if it did, then 99.7% of the observations should have z-values less than 3 in absolute value).

Problem 3

Here are the descriptive statistics for the two variables:

Descriptive Statistics: MarketReturn, IBMRet

Variable	N	N*	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3	Maximum
MarketReturn	11537	0	0.0178	0.00907	0.9742	-17.1600	-0.4200	0.0500	0.4700	11.5100
IBMRet	11537	0	0.0256	0.0152	1.6365	-23.5671	-0.8449	-0.0180	0.8400	13.1412

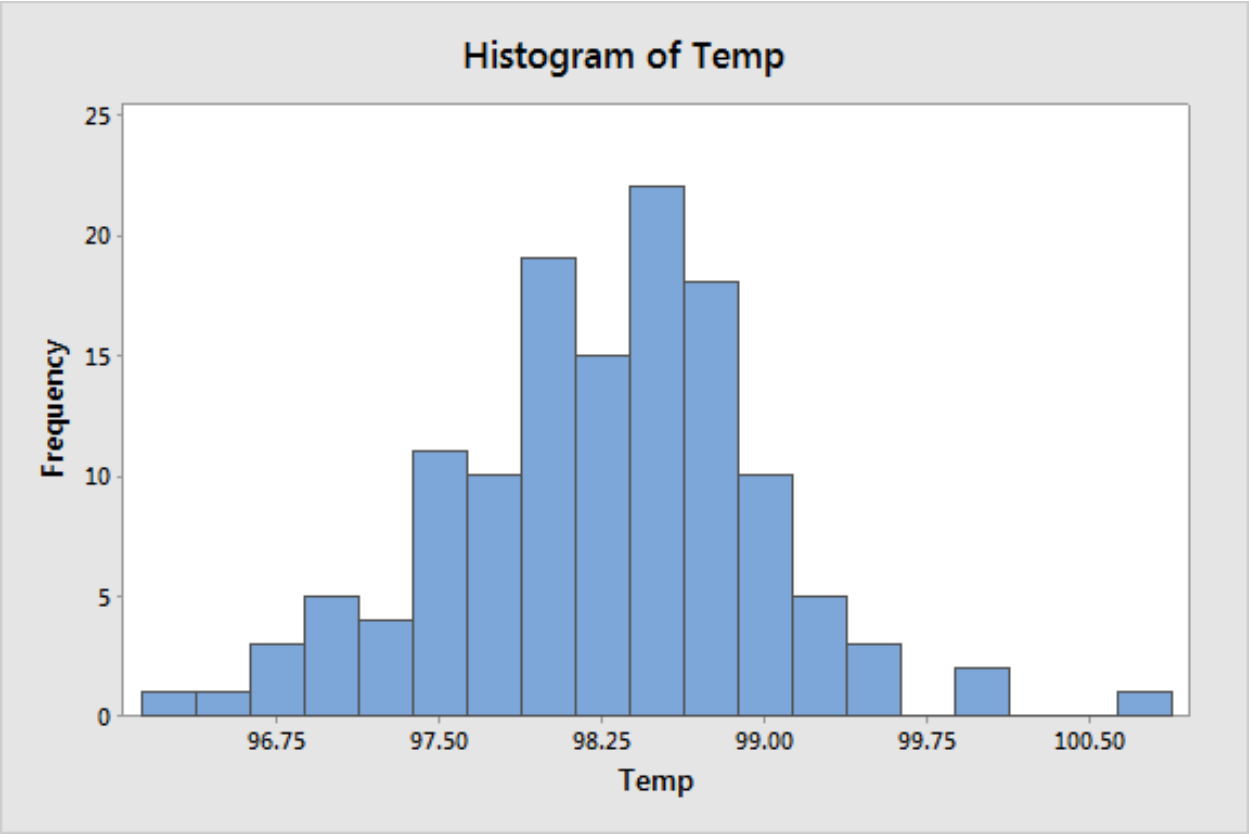
The means and Sharpe ratios for the two investments are

Market	mean = 0.0178	Sharpe ratio = 0.0178 / 0.9742 = 0.0183
IBM	mean = 0.0256	Sharpe ratio = 0.0256 / 1.6365 = 0.0156

IBM has the higher mean return, but a lower Sharpe ratio. There is more variability in the returns from IBM than there is in the returns from the Market.

Problem 4

1. Here is a histogram of *Temp:*



The data has a reasonably bell-shaped distribution, though the high observation (100.8) might be an outlier

2. If the population is “all temperatures for healthy individuals,” then the population mean is probably close to 98.6 (the “normal” body temperature for humans).

3. The histogram is roughly symmetric, so the sample mean is close to the center of symmetry. In this case, it looks like the sample mean is close to 98.6.

4. Here are the descriptive statistics for *Temp*:

Descriptive Statistics: Temp

Variable	N	N*	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3	Maximum
Temp	130	0	98.249	0.0643	0.733	96.300	97.800	98.300	98.700	100.800

The sample mean is 98.249. It seems reasonably close, but not exactly equal to 98.6. “Reasonably close” means off by only a few tenths of a degree. Later on in class, we will use learn a better way to define “reasonably close.”