3-2.

a.

$$\mu = \frac{\sum_{i=1}^{N} x_i}{N} = 168.81/10 = \$16.881$$

To compute the median, rank the observations and compute the average of the middle two.

9.95 11.22 14.52 14.98 16.65 17.87 18.74 19.95 21.98 22.95

Median = (16.65 + 17.87)/2 = $17.26

No values are repeated. Therefore, there is no mode.

3-8.

a. $\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = 7.09$

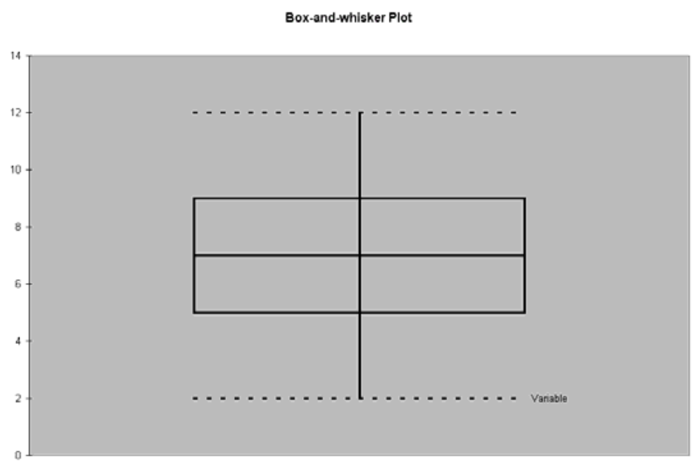
Median is the center value after the data have been arranged in numerical order.
Median = 7.0
Mode is the value that appears most frequently. The following frequency distribution can be developed:

1	0
2	1
3	3
4	2
5	4
6	5
7	6
8	1
9	6
10	3
11	3
12	1

Two Modes:
7 and 9 occur 6 times

b. Box and Whisker can be developed using PHStat Add-in for Excel.

Box-and-whisker Plot	
Five-number Summary	
Minimum	2
First Quartile	5
Median	7
Third Quartile	9
Maximum	12



3-18

a. Sorting the data and determine what position 45,000 is in you can solve the percentile equation for the percentile.

$i = \frac{p}{100}(n+1)$ = (75/100)*(200+1) = 150.75 ; This 75th percentile can be approximated as a value somewhere between the 150th and 151st value in the data. You can use PHStat's Stack feature under Data Preparation to reorganize the data. Then sorting the data you get the following:

148	44,879
149	44,879
150	44,904
151	44,980
152	45,052
153	45,148
154	45,153
155	45,227
156	45,228
157	45,276

Thus, the 75th percentile is a value between 44,904 and 44,980. It would be .75 of the distance between these two values up from 44,904 which gives 44,961. Note, Excel gives a slightly different value of 44,923 since it uses a slightly different methodology for interpolating the difference between the 150th and 151st data values.

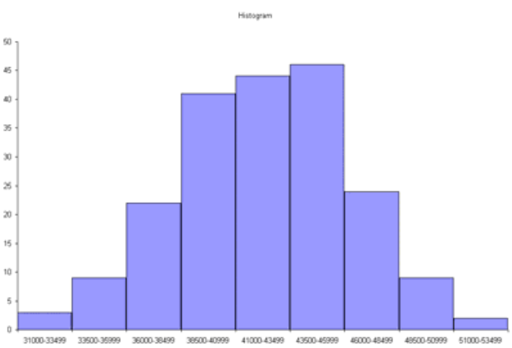
b. Using Excel's Average and Median functions you can find that
Mean = 42,261
Median = 42,326

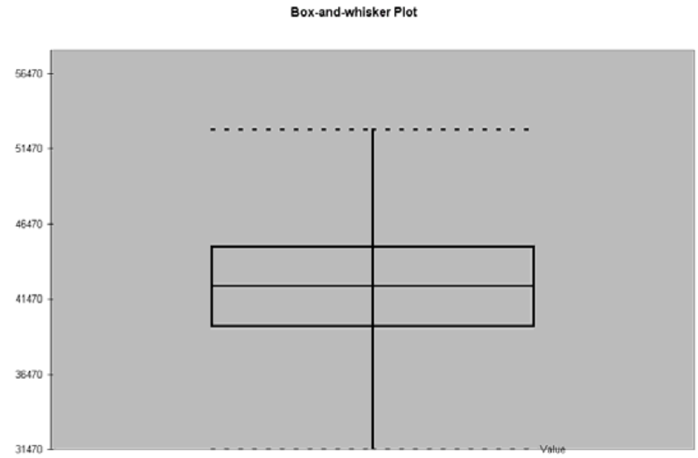
c. Using Excel's histogram tool the following histogram can be created.
According to Sturges' rule, the appropriate number of classes should be:
1 + 3.322(log(200)) = 8.6 which is rounded to 9 classes.

The class width is:

$$w = \frac{High - Low}{\#} = \frac{52{,}774 - 31{,}476}{9} = \frac{21{,}298}{9} = 2{,}366.44$$

We choose to round this to 2,500 and start at 31,000 giving the following histogram:





Histograms and box and whiskers plots have certain things in common. In both instances, we get an idea of how the data are distributed, where the center is, and what the shape of the distribution is and how spread out the data are. The histogram breaks the data down into classes and illustrates the actual number of values in each class where the box and whiskers plot shows the median and the inter-quartile range.

d. Student answers will vary.

3-22.

a. Range = 9 - 4 = 5

b.

x	x - μ	$(x - \mu)^2$
4	-1.83333	3.361111
6	0.166667	0.027778
9	3.166667	10.02778
4	-1.83333	3.361111
5	-0.83333	0.694444
7	1.166667	1.361111
		18.83333

$$\mu = \frac{\sum_{i=1}^{N} x_i}{N} = 35/6 = 5.8333, \quad \sigma^2 = \frac{\sum_{i=1}^{N} (x-\mu)^2}{N} = 18.83333/6 = 3.1389$$

$$\sigma = \sqrt{\sigma^2} = \sqrt{3.1389} = 1.7717$$

c.

$$s^2 = \frac{\sum_{i=1}^{n} (x-\bar{x})^2}{n-1} = 18.83333/(6\text{-}1) = 3.7667$$

$$s = \sqrt{S^2} = \sqrt{3.37667} = 1.9408$$

The statistics (assuming the data were a sample) will be larger than the parameters (assuming the data were a population) since the division for s and s^2 are computed using a divisor on n-1 rather than N.

3-24.

First find the sample standard deviation

X	X - $\bar{x}$	(X - $\bar{x}$)^2
51	-0.1875	0.035156
43	-8.1875	67.03516
58	6.8125	46.41016
67	15.8125	250.0352
67	15.8125	250.0352
69	17.8125	317.2852
40	-11.1875	125.1602
52	0.8125	0.660156
66	14.8125	219.4102
44	-7.1875	51.66016
47	-4.1875	17.53516
41	-10.1875	103.7852
41	-10.1875	103.7852
45	-6.1875	38.28516
47	-4.1875	17.53516
41	-10.1875	103.7852
819		1712.438

$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$ = 819/16 = 51.1875, $S^2 = \frac{\sum_{i=1}^{n}(x-\bar{x})^2}{n-1}$ = 1712.438/(16-1) = 114.1625

$S = \sqrt{S^2} = \sqrt{114.1625}$ = 10.6847

a. one standard deviation of the mean is
51.1875 $\pm$ 1(10.6847)
40.5028 – 61.8722
11 of 16 observations are within this range so the proportion = 11/16 = .6875 or 68.75%

b. two standard deviations of the mean is
51.1875 $\pm$ 2(10.6847)
29.8181 – 72.5569
16 of 16 observations are within this range so the proportion = 16/16 = 1.0000 or 100%

c. three standard deviations of the mean is
51.1875 $\pm$ 3(10.6847)
19.1334 – 83.2416
16 of 16 observations are within this range so the proportion = 16/16 = 1.0000 or 100%

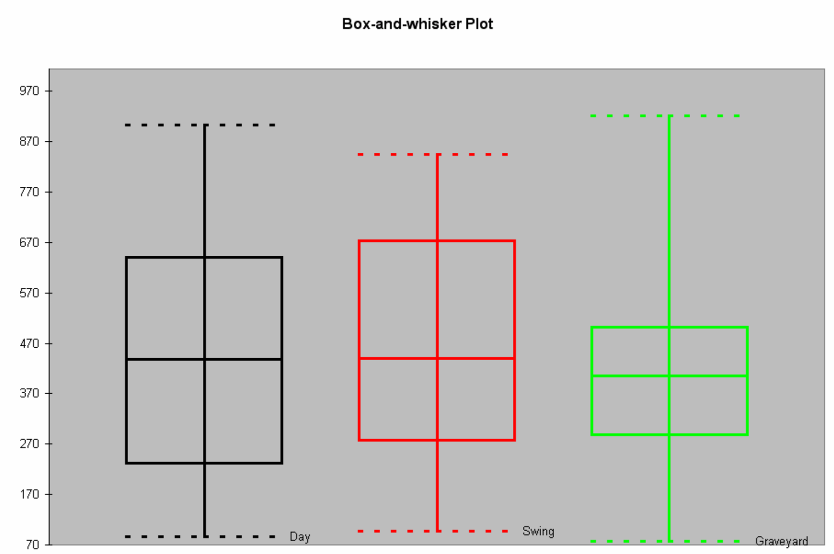
3-32.

Assume that the data represents a sample from the population.

a. The Descriptive Statistics procedure under the data analysis tool in Excel can be used to generate the desired statistics for each shift as follows:

Day		*Swing*		*Graveyard*	
Mean	426.05	Mean	445.65	Mean	415.5
Standard Error	55.6685	Standard Error	47.52286	Standard Error	45.10599
Median	438.5	Median	440.5	Median	405.5
Mode	456	Mode	#N/A	Mode	467
Standard Deviation	248.9571	Standard Deviation	212.5287	Standard Deviation	201.7201
Sample Variance	61979.63	Sample Variance	45168.45	Sample Variance	40691
Kurtosis	-0.92573	Kurtosis	-0.78272	Kurtosis	1.001117
Skewness	0.282289	Skewness	0.245261	Skewness	0.662917
Range	815	Range	747	Range	843
Minimum	87	Minimum	98	Minimum	78
Maximum	902	Maximum	845	Maximum	921
Sum	8521	Sum	8913	Sum	8310
Count	20	Count	20	Count	20

b. PHStat's Box and Whisker's procedure using multiple groups unstacked is a good way to generate the required box and whisker plots.



c. Student reports should highlight differences between the three shifts in terms of spread in the data with special attention to how the Graveyard shift differs from the other two. The box and whisker plots and the descriptive statistics should be referred to in the discussion.

3-36.

The standardized value is computed using:

$$z = \frac{x - \bar{x}}{s}$$

So for the data in this case we get: $z = \dfrac{615 - 500}{75} = 1.53$

3-38.

a. Since the shape of the distribution is unknown, we can use Tchebysheff's theorem to determine the solution. Since σ=200 and μ=3,000, then the range 2,600 to 3,400 is $\mu \pm 2(\sigma)$. According to Tchebysheff's theorem, at least 75% of the data in a distribution will fall within $\mu \pm 2(\sigma)$.

b. The range $\mu \pm 3(\sigma)$ should include at least 8/9 (89%) of the data values. That means that the range 2,400 to 3,600 should contain at least 89 percent of the data values. However, since we don't know the shape of the population, we can't say for sure what percentage will be greater than 3,600. We do know that the percentage will be less than 11% and most likely it will be considerably less.

c. The same issue is present here as in part b. We can't say with any certainty what the percentage will be. We do know that it will be less than 11 percent but we don't know how much less.

3-48.

Use Excel's average and standard deviation functions to determine the mean and standard deviation of each type of bread. The results are shown below.

a. White Bread has the highest average daily demand.

Bread Type	Average
White	599.77273
Wheat	530.40909
Multigrain	470.36364
Black	383.59091
Cinnamon Raisin	139.72727
Sour Dough French	127.09091
Light Oat	261.63636

b. Use Excel's Histogram feature to develop the frequency distribution for each bread type. Student answers will vary depending on number of classes selected and class widths used, but shown below are the results if you let Excel set up the bins.

White Bread	*Frequency*
251 - 375	1
376 - 500	5
501 - 625	7
626 - 750	5
751 - 875	4

Cinnamon Raison	*Frequency*
54.76 - 84.00	1
84.01 - 113.25	4
113.26 - 142.50	8
142.51 - 171.75	4
171.76 - 201.00	5

Wheat Bread	*Frequency*
264.26 - 352.00	1
352.01 - 439.75	3
439.76 - 527.50	8
527.51 - 615.25	4
615.26 - 703.00	6

Sour Dough French	*Frequency*
64 - 88	1
89 - 113	8
114 - 138	4
139 - 163	7
164 - 188	2

Multigrain Bread	*Frequency*
212.76 - 299.00	1
299.01 - 385.25	4
385.26 - 471.50	8
471.51 - 557.75	2
557.76 - 644.00	7

Light Oat Bread	*Frequency*
127 - 172	2
173 - 218	3
219 - 264	8
265 - 310	4
311 - 356	5

Black Bread	*Frequency*
182.76 - 256.00	1
256.01 - 329.25	5
329.26 - 402.50	7
402.51 - 475.75	6
475.76 - 549.00	3

c. White Bread has the highest standard deviation.

Bread Type	**Standard Deviation**
White	149.0550975
Wheat	107.0236552
Multigrain	108.0130263
Black	81.9510069
Cinnamon Raisin	33.01973698
Sour Dough French	28.68367561
Light Oat	57.89526295

d. Use Excel to calculate the coefficient of variation for each bread type. White bread has the greatest relative variability and wheat bread has the lowest relative variability.

Bread Type	Coefficient of Variation
White	24.85%
Wheat	20.18%
Multigrain	22.96%
Black	21.36%
Cinnamon Raisin	23.63%
Sour Dough French	22.57%
Light Oat	22.13%

e. Use Tchebysheff's Theorem to calculate the upper range of two standard deviations from the mean. You must use Tchebysheff's Theorem because you do not know if the data is bell-shaped.

Bread Type	Required Loaves
White	897.8829
Wheat	744.4564
Multigrain	686.3897
Black	547.4929
Cinnamon Raisin	205.7667
Sour Dough French	184.4583
Light Oat	377.4269

f. Use Excel's Pivot Table feature to calculate the average total loaves by day of week. The highest average is on day 6.

Average of Total Loaves Sold	
Day of Week	Total
1	2196
2	2947
3	2388
4	2335.5
5	2336.8
6	3116.5
8	1772
Grand Total	2512.590909

3-54.

The standard deviation measures how the data is spread around the mean. If the means are not the same then comparing the measure of the spread around the mean does not give any useful comparison of two data sets. The appropriate measure is the coefficient of variation. This essentially standardizes the data so that a comparison between two data sets is meaningful. The larger the coefficient of variation the more variable is the data. You would then be able to say that one data set has a larger relative variation than another data set.

3-60.

X	$X - \bar{x}$	$(X - \bar{x})^2$
62	-9.5	90.25
75	3.5	12.25
81	9.5	90.25
64	-7.5	56.25
81	9.5	90.25
66	-5.5	30.25
70	-1.5	2.25
70	-1.5	2.25
69	-2.5	6.25
73	1.5	2.25
72	0.5	0.25
75	3.5	12.25
858		395

a. $\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$ = 858/12 = 71.5

b. To compute the median, rank the observations and average the middle two values.

62 64 66 69 70 70 72 73 75 75 81 81

Median = (70+72)/2 = 71

c. The mode is 70 75 81

d. $s^2 = \frac{\sum_{i=1}^{n} (x - \bar{x})^2}{n - 1}$ = 395/(12-1) = 35.9091, $s = \sqrt{S^2} = \sqrt{35.9091}$ = 5.9924