Homework Solutions – Chapters 1 & 2

Homework 1.1

1. answers will vary
2. answers will vary
3. answers will vary
4. When an explicit reference is made to a sample, a study, a survey group, *etc*. when reporting statistics, then this is an indicator that these are descriptive statistics. When a statistic is reported in a more general way, as if it applies to a population, not a sample, survey, or study, then we say that it is an instance of statistical inference. When such a statement is accompanied by a margin of error, then it definitely is statistical inference.

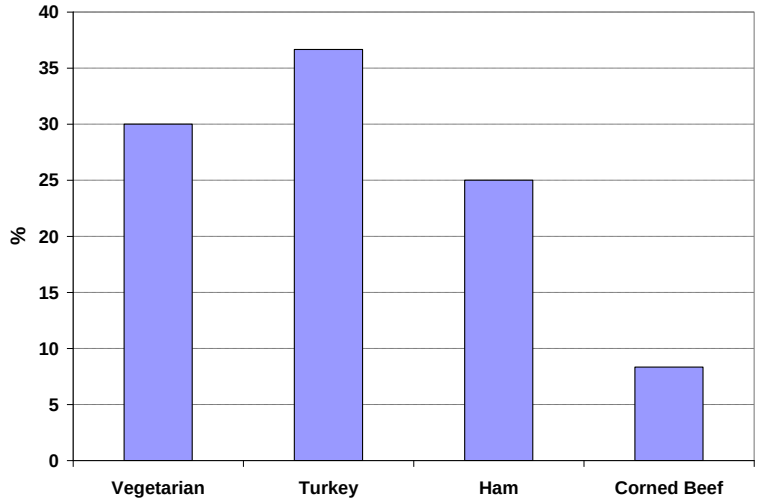
Homework 2.1

1.

(a)

Sandwich	Frequency	Relative Frequency
Vegetarian	18	30%
Turkey	22	37%
Ham	15	25%
Corned Beef	5	8%

(b)



(c) The Mode is *Turkey* sandwich.

Homework 2.2

1.

	Sample Average	Median
Men	39.15	36.30
Women	45.42	45.75

In this study it was found that men studied on average, 39.15 hours per week while women studied 45.42 hours per week. Thus, in this study, women studied about 16% longer than men on average. It was also found that the median for men was 36.30 hours per week as compared to 45.75 hours per week for the women. Thus, the women's median was 26% larger than the men's.

Homework 2.3

1. (a)

	Q_1	Q_3	63^{rd} Percentile	80^{th} Percentile
Men	31.0	44.8	41.4	47.7
Women	41.4	50.0	47.9	52.5

(b)

- 63% of men study less than 41.4 hours per week while only 25% of women study less than this amount.
- 20% of men study more than 47.7 hours per week while more than 37% of women study more than this amount.

2. As shown in the table below, the minimums for men and women, as well as the maximums, are very close in value.

	Min(hr)	***Max(hr)***
Men	28.2	60.0
Women	27.5	61.2
Difference	0.7	-1.23
% Difference	2.5%	-2.0%

3. About half the data in sample 2 is less than all the data in sample 1. About 25% of the data in sample 1 is larger than all the data in sample 2.

Homework 2.4

1. (a)

	R	*s*	*IQR*
Men	31.8	9.2	13.7
Women	33.7	7.7	8.6

(b)
The range shows that the women's data has slightly more variability than the men's data while the standard deviation and the inter-quartile range show that men's data have more variability. The standard deviation for the men, 9.2 *hrs* is almost 20% larger than the value for the women, 7.7 *hrs*.

Homework 2.5

1.

a. Relative Frequencies

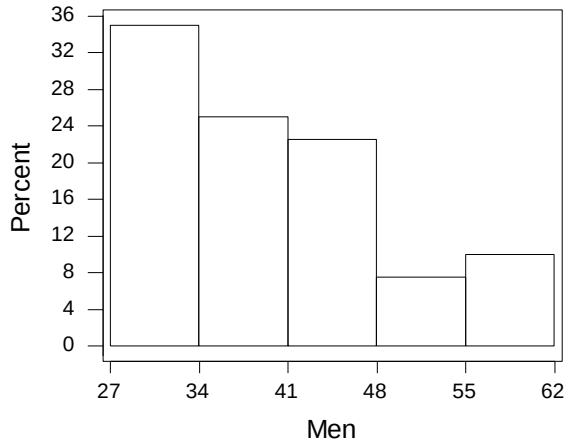
5 Classes – Men

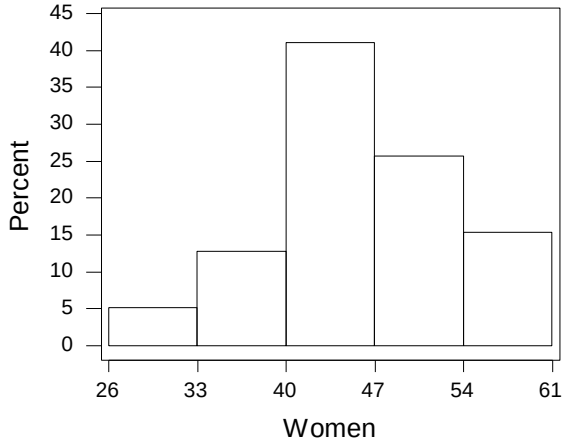
Class	Relative Freq
(27, 34)	0.35
[34, 41)	0.25
[41, 48)	0.225
[48, 55)	0.075
[55, 62)	0.1

5 Classes - Women

Class	Relative Freq
(27, 34)	0.1
[34, 41)	0.125
[41, 48)	0.425
[48, 55)	0.25
[55, 62)	0.1

b. Histograms





a. Relative Frequencies

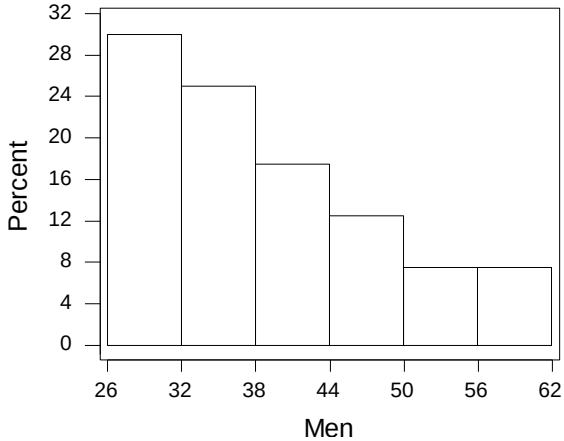
6 Classes - Men

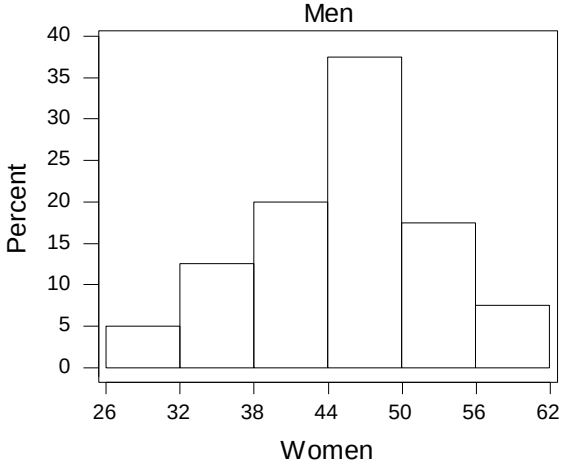
Class	Relative Freq.
(26, 32)	0.30
[32, 38)	0.25
[38, 44)	0.175
[44, 50)	0.125
[50, 56)	0.075
[56, 62)	0.075

6 Classes - Women

Class	Relative Freq.
(26, 32)	0.05
[32, 38)	0.125
[38, 44)	0.2
[44, 50)	0.375
[50, 56)	0.175
[56, 62)	0.075

b. Histograms





a. Relative Frequencies

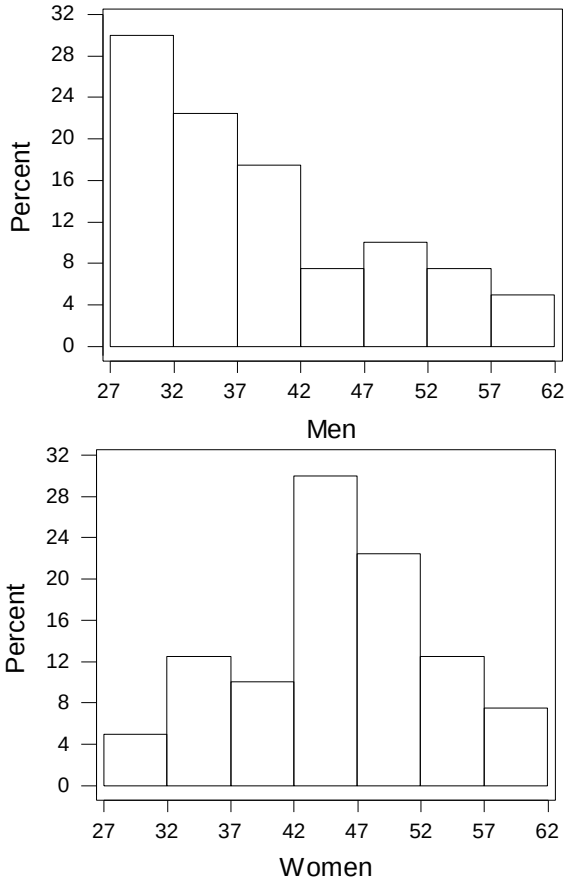
7 Classes – Men

Class	Relative Freq.
[27, 32)	0.3
[32, 37)	0.225
[37, 42)	0.175
[42, 47)	0.075
[47, 52)	0.1
[52, 57)	0.075
[57, 62)	0.05

7 Classes – Women

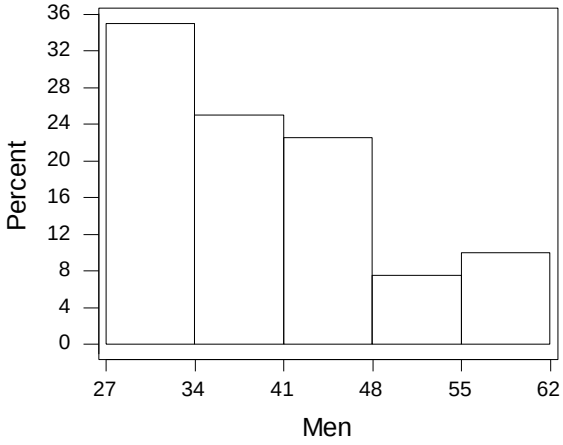
Class	Relative Freq.
[27, 32)	0.05
[32, 37)	0.125
[37, 42)	0.1
[42, 47)	0.3
[47, 52)	0.225
[52, 57)	0.125
[57, 62)	0.075

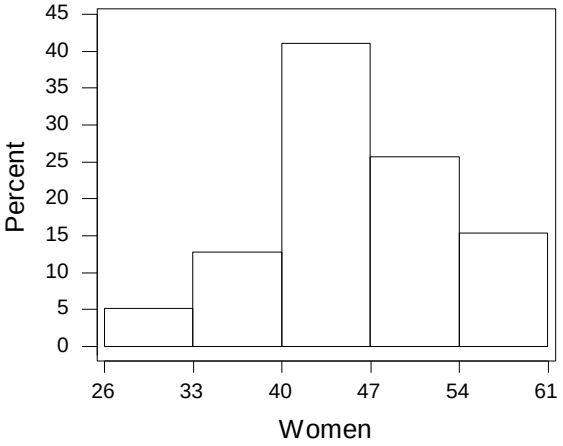
b. Histograms

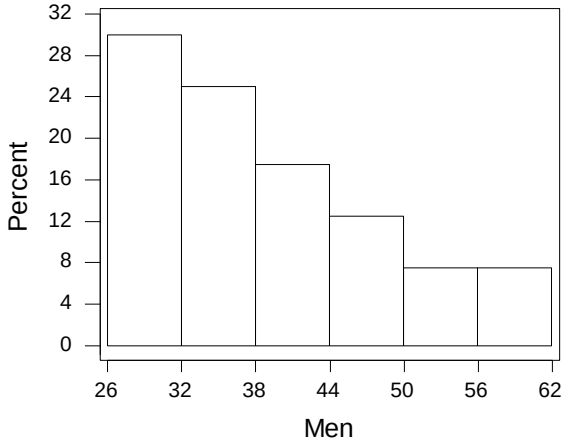


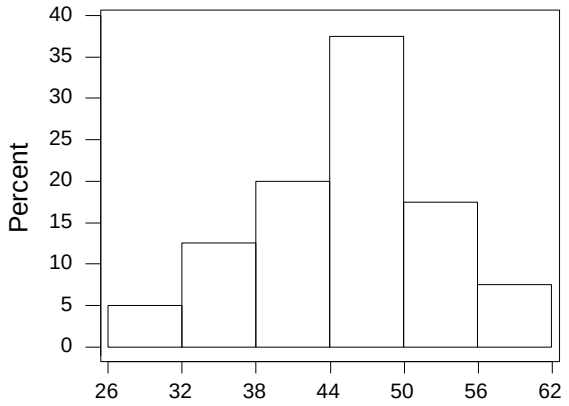
Homework 2.6

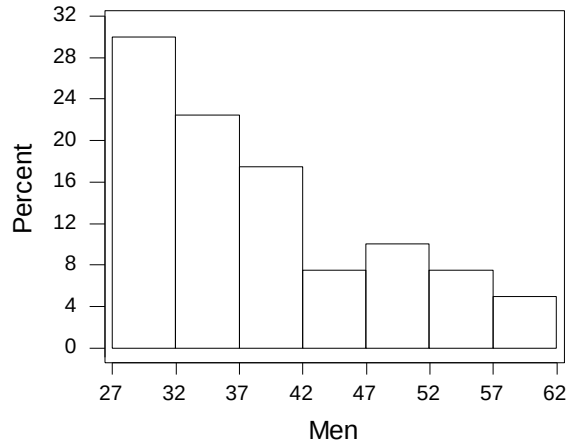
1.

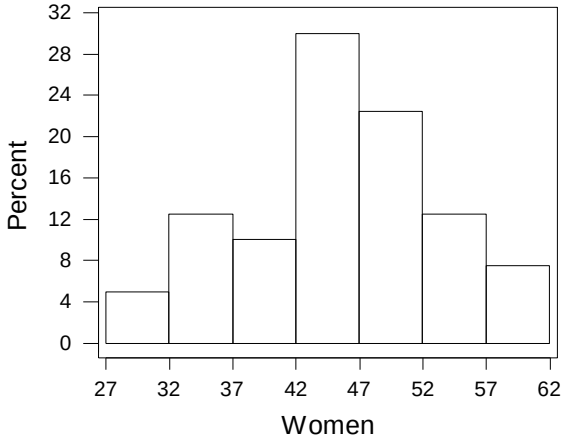












Homework 2.7

1. All the histograms for the men's data indicate that it is right skewed. The histograms for the women's data indicate that the data is symmetric and bell-shaped. The histograms with 6 bars indicates this best for both men's and women's data.

Homework 2.8

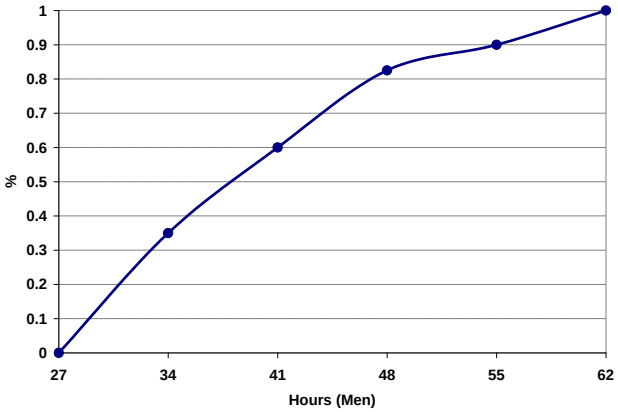
1. Ogives.

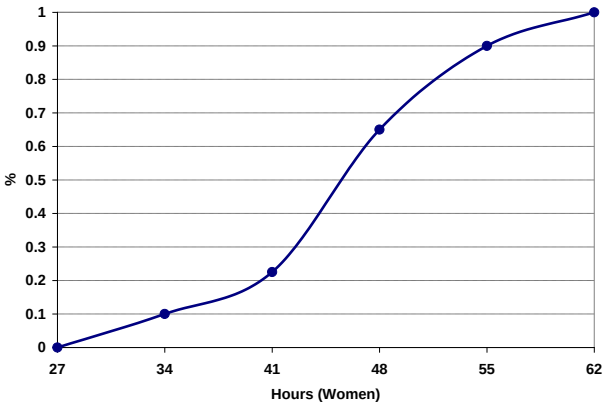
5 Classes – Men

Class	Relative Freq	Cumulative Rel. Freq.
(27, 34)	0.35	0.35
[34, 41)	0.25	0.6
[41, 48)	0.225	0.825
[48, 55)	0.075	0.9
[55, 62)	0.1	1.0

5 Classes - Women

Class	Relative Freq	Cumulative Rel. Freq.
(27, 34)	0.1	0.1
[34, 41)	0.125	0.225
[41, 48)	0.425	0.65
[48, 55)	0.25	0.9
[55, 62)	0.1	1.0



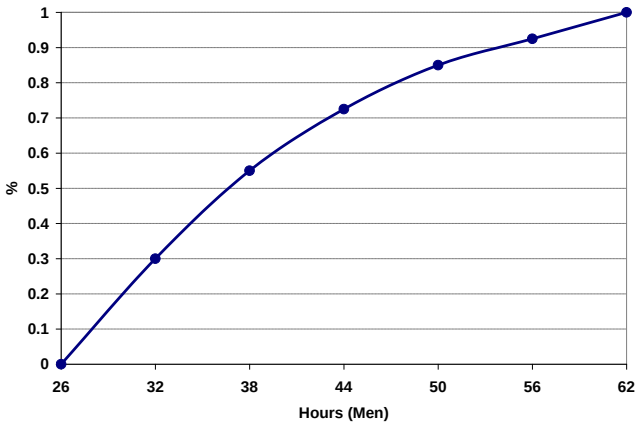


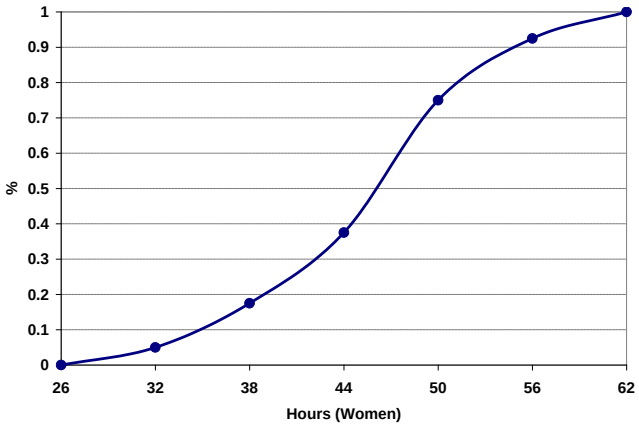
6 Classes - Men

Class	Relative Freq.	Cumulative Rel. Freq.
(26, 32)	0.30	0.30
[32, 38)	0.25	0.55
[38, 44)	0.175	0.725
[44, 50)	0.125	0.85
[50, 56)	0.075	0.925
[56, 62)	0.075	1.0

6 Classes - Women

Class	Relative Freq.	Cumulative Rel. Freq.
(26, 32)	0.05	0.05
[32, 38)	0.125	0.175
[38, 44)	0.2	0.375
[44, 50)	0.375	0.75
[50, 56)	0.175	0.925
[56, 62)	0.075	1.0



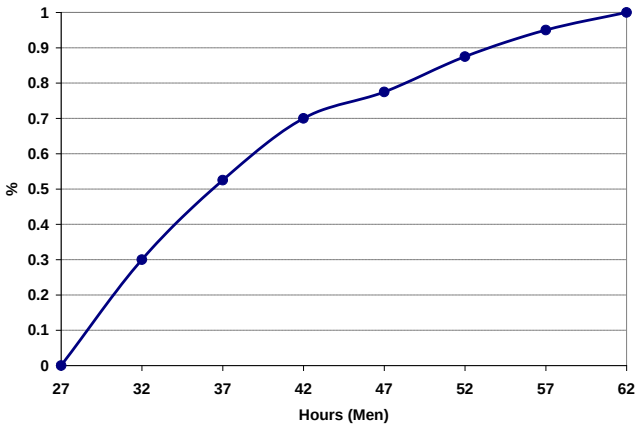


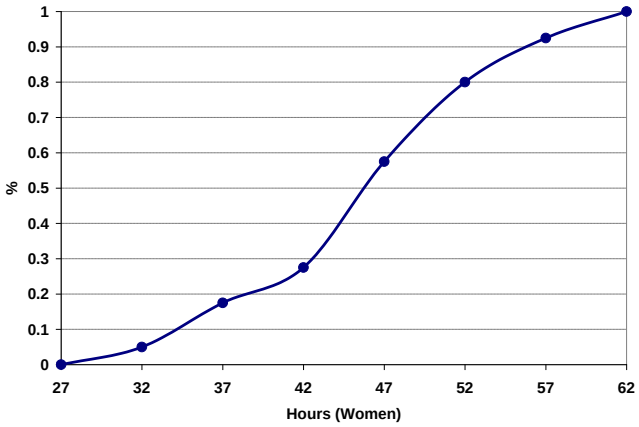
7 Classes – Men

Class	Relative Freq.	Cumulative Rel. Freq.
[27, 32)	0.3	0.3
[32, 37)	0.225	0.525
[37, 42)	0.175	0.7
[42, 47)	0.075	0.775
[47, 52)	0.1	0.875
[52, 57)	0.075	0.95
[57, 62)	0.05	1.0

7 Classes – Women

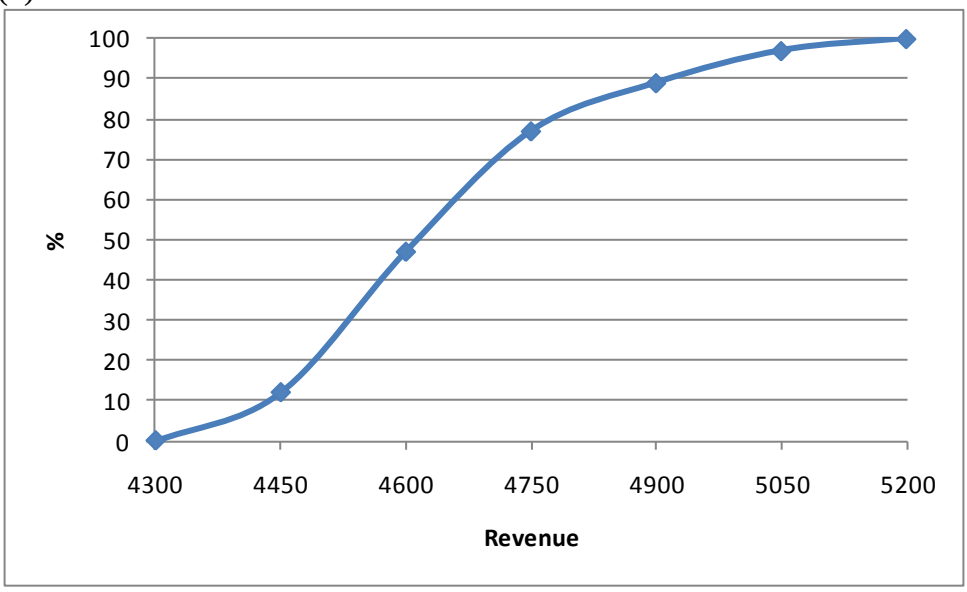
Class	Relative Freq.	Cumulative Rel. Freq.
[27, 32)	0.05	0.05
[32, 37)	0.125	0.175
[37, 42)	0.1	0.275
[42, 47)	0.3	0.575
[47, 52)	0.225	0.8
[52, 57)	0.125	0.925
[57, 62)	0.075	1.0


1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
%
27
32
37
42
47
52
57
62
Hours (Men)

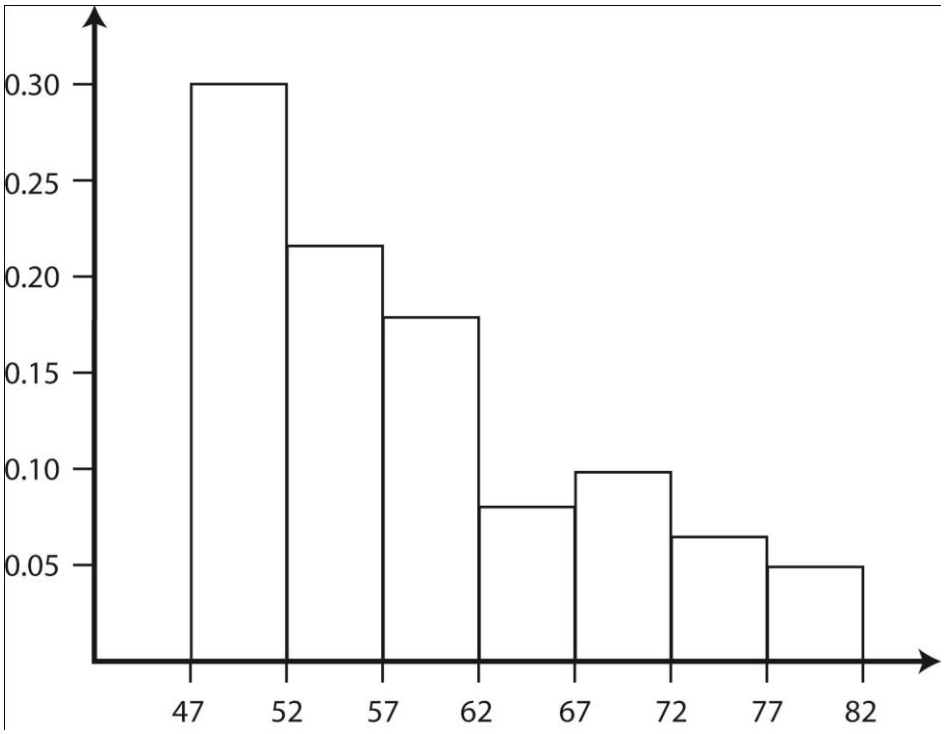

1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
%
27
32
37
42
47
52
57
62
Hours (Women)

Homework 2.9

1. (a) about 8% (b) about 35%+13% = 48%, (c) about 13%+8%+2% = 23% (d)



2. (a) about 52% (b) about 77%-52% = 25%, (c) about 100%-77% = 23% (d)

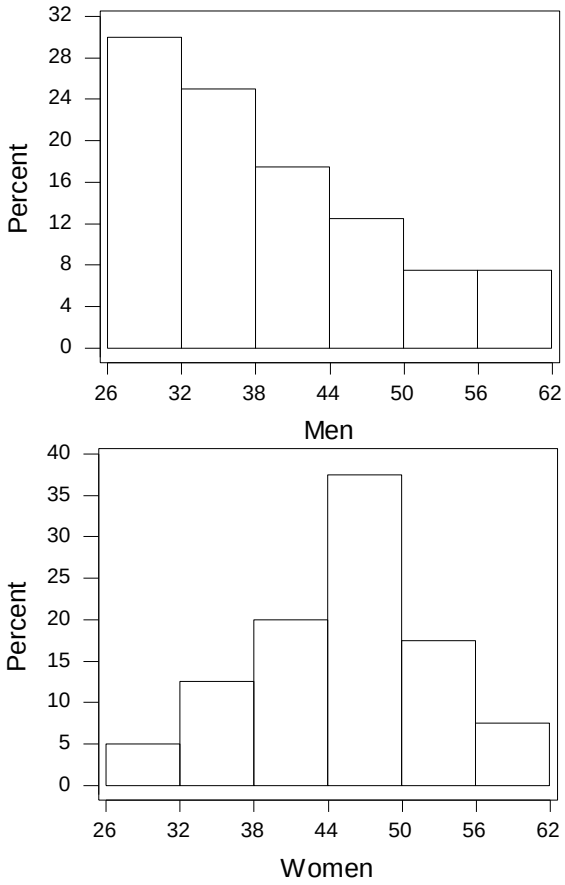


Homework 2.10

1.

Data	Mean	Median	Difference	% Difference	Judgement
Men	39.15	36.30	2.85	7.8%	Right Skew
Women	45.42	45.75	0.33	0.7%	Symmetric

The corresponding histograms are also shown:



2. Since the average is larger than the median, there is indication of right skew.
3. Since the average is less than the median, there is indication of left skew.
4. Compute the differences to get: 2.1, -1.6, 4.23, -8.4. Thus, (a) sample 2 has the least skew (-1.6), (b) sample 4 has the most skew (-8.4), (c) sample 4 has the most left skew (-8.4), (d) sample 3 has the most right skew (4.23), (e) sample 2 has the least left skew (-1.6), (f) sample 1 has the least right skew (2.10), (g) sample 2 has the most symmetry, same as (a), (h) sample 4 has the least symmetry, same as (b)

Homework 2.11

1. The men's data is right-skewed so the Empirical Rule does not apply. The women's data is symmetric so the Empirical Rule does apply. The chart below shows the results of using the Empirical Rule on the women's data at 68% and 95%

Interval Size	Interval	Num. in Interval	% in Interval	Judgement
68%	37.7, 53.1	30	75%	Reasonably Accurate
95%	30.1, 60.8	38	95%	Very Accurate
99%	22.4, 68.5	40	100%	Very Accurate

2. (a) 16% (b) 84% (c) 50% (d) 84% (e) 16% (f) 2.5% (g) 97.5% (h) 97.5% (i) 2.5% (j) 95% (k) 68% (l) 99.7% or *almost all*.

3. (a) 46 (b) 30 (c) 34, 42 (d) 42 (e) 30 (f) 38 (g) 34 (h) 30, 46 (i) 26, 50

Homework 2.12

1.

Men, 6 bar histogram:
Median – First two bars add to about 54, so M is less than 38. I will estimate 36.
Range – 62-26 = 36

Women, 6 bar histogram:
Median – First 4 bars add to 37. Fifth bar added brings it to 72. Thus, M is between 44 and 50. I will estimate 47.
Range – 62-26 = 36
Standard Deviation – Assume 3*s* equals ½ of 36. Thus, 3*s* = 18, and *s* = 6.

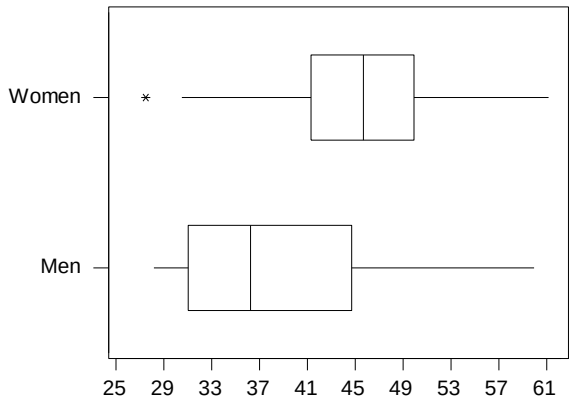
	Men's Data		*Women's Data*		
	Median	Range	Median	Range	Standard Deviation
Estimated	36.0	36.0	47.0	36.0	6.0
Actual	36.3	31.8	45.4	33.7	7.7

The median estimates are particularly close. As expected, the range estimates are a little too big. The standard deviation estimate for the women's data is in the ball park.

Homework 2.13

1.

	Men's Data	*Women's Data*
Q_1	31.0	41.4
M	36.3	45.8
Q_3	44.8	50.0
IQR	13.7	8.6
Lower Fence	10.4	28.5
Upper Fence	65.4	62.9



Homework 2.14

1. The median for the men's data is about 36 hours and 46 hours for the women's data. Thus, the women's median is about 28% larger than the men's median. The spread in the data looks to be similar as the ranges are close in value. The women's data appears to be fairly symmetric and bell-shaped. There is one outlier on the low side at approximately 27 hours. The men's data appears to have a strong left skew with no outliers indicated. We also note that about 50% of women studied longer than 45 hours, while just 25% of the men did.

2. (a) 4 (b) 2 (c) 1 (d) 4 (e) 2 (f) 3 (g) 3 (h) 1 (i) 1 & 4 (j) 2 (k) 1-highly right skewed, 2-slightly left skewed, 3-left skewed, 4-fairly symmetric, (l) 1, (m) 3

3. answer is *c*

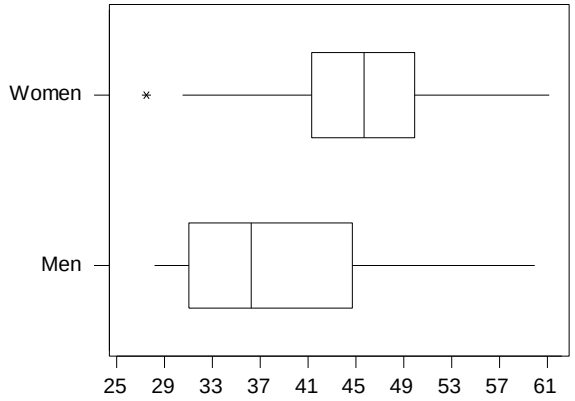
4. IQR = 60-50=10. Thus, upper fence is located at 60 + 1.5*10 = 75

5. Sample 3 has an IQR of about 77-47=30.

6. Sample 1

Homework 2.15

1.



Homework 2.16

1. As the table below shows, there are no outliers indicated by the z-scores. However, this is in contrast to the box plots which showed that the women's data indicated that 27.5 hours was an outlier.

	Data Values		Z-scores	
	Men	Women	z-Men	z-Women
Smallest	28.2	27.5	-1.2	-2.3
Largest	60.0	61.2	2.3	2.1

2. The average of this data is 9.522 and the standard deviation is 2.851. The z-score for the data value 8.2 is (8.2-9.522)/2.851 = -0.46. Thus, 8.2 is 0.46 standard deviations *below* the average.
3. Bob's z-score for Biology is (73-63)/10 = 1 and his z-score for Math is (168-148)/30 = 0.67. Thus, Bob scored higher n the Biology exam.

Homework 2.17

1. Note that we are comparing the Average to the Median.

(a)

Sample	Computation	Percent Difference
1	(27.4-25.3)/25.3*100	8.3%
2	(128.6-130.2)/130.2*100	-1.2%
3	(67.33-63.1)/63.1*100	6.7%

4	(20.2-28.6)/28.6*100	-29.4%

(b)
The Average for Sample 1, 27.4 *gal* is about 8.3% larger than the Median, 25.3 *mph*.

In Sample 2, we see that the Average, 128.6 *mph* is approximately 1.2% smaller than the Median, 130.2 *mph*.

The difference between the Average, 67.33 *m* and the Median, 63.1 *m*, is 4.23 *m* which shows that the Average is about 6.7% larger than the Median in Sample 3.

In Sample 4, we see that the Average is about 29.4% smaller than the Median ($20.2 compared to $28.6)

(c)
First, note that: 28.6/20.2 = 1.416. Thus, the Median for Sample 4 is about 1.4 times larger than the Average.

2. (a) one and a half (1.5) (b) 50% (c) two-thirds (0.667) (d) 33%