Chapter 5-Measures of Variability

5.1 Variability of NoPassage group:
Range = 57 – 34 = 23
St. Dev. = 6.83
Variance = 46.62

5.3

```
### Exercise 5.3
 noread <- c(54, 52, 51, 50, 36, 55, 44, 46, 57, 44, 43, 52,
        38, 46, 55, 34, 44, 39, 43, 36, 55, 57, 36, 46,
        49, 46, 49, 47)
 read <- c(66, 75, 72, 71, 55, 56, 72, 93, 73, 72, 72, 73,
        91, 66, 71, 56, 59)
# 5.1
 cat("range",range(noread))
range 34 57
 cat("variance", var(noread))
variance 46.62434
 cat("st. dev. ",sd(noread))
st. dev.  6.828202
 cat("\n")


# 5.2
cat("range",range(read))
range 55 93
cat("variance", var(read))
variance 112.5294
cat("st. dev. ",sd(read))
st. dev.  10.60799
```

5.5 Percentages within two standard deviations in Exercise 5.2
$s = 10.61$
$\overline{X} \pm 2(10.61) = 70.18 \pm 21.22 = 48.96 — 91.4$
16 scores (or 94%) lie within 2 standard deviations of the mean

5.7 Multiplying or dividing by a constant:

Original	2	3	4	4	5	5	9	$\overline{X}_1 = 4.57$	$s_1 = 2.23$
X * 2	4	6	8	8	10	10	18	$\overline{X}_2 = 9.14$	$s_2 = 4.45$
X / 2	1	1.5	2	2	2.5	2.5	4.5	$\overline{X}_3 = 2.29$	$s_3 = 1.11$

5.9 Convert revised data to mean = 0

Since adding or subtracting a constant will not change the standard deviation, but will change the mean, I can subtract 3.27 from every score for X_2 in Exercise 5.8, making the mean = 0, and keeping $s_2 = 1.0$. the new values are

X_3 -0.889 0.539 -1.842   0.539 -0.413   1.016   1.016 $\overline{X}_1 = 0$ $s_1 = 1$

5.11 Boxplot for Exercise 5.1:

Median location = $(N + 1)/2 = 29/2 = 14.5$
Median = 46
Hinge location = (median location +1)/2 = 15/2 = 7.5
Hinge = 43 and 52
H-spread = 52 – 43 = 9
Inner fences = hinges $\pm$ 1.5*H-spread = hinges $\pm$ 1.5*9 = hinges $\pm$ 13.5 = 29.5 and 65.5
Adjacent values = 34 and 57
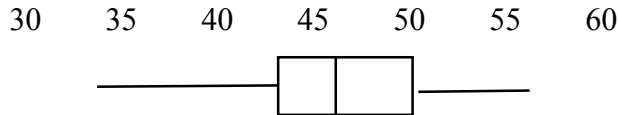
5.13 Boxplot for ADDSC:

Median location = $(N + 1)/2 = 89/2 = 44.5$Median = 50
Hinge location = (median location +1)/2 = 45/2 = 22.5
Hinge = 44.5 and 60.5
H-spread = 60.5 – 44.5 = 16
Inner fences = hinges $\pm$ 1.5*H-spread = hinges $\pm$ 1.5*16 = hinges $\pm$ 24 = 20.5 and 85.5
Adjacent values = 26 and 78
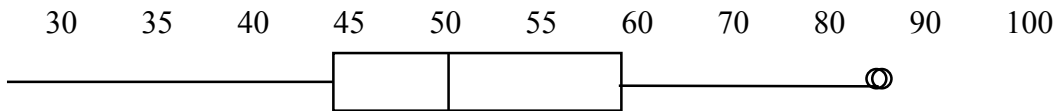
5.15 Variance when you add a score equal to the mean.

$$s^2 = \frac{\Sigma(X - \overline{X})^2}{N-1} = \frac{\Sigma(X - \overline{X})^2}{27} = 46.6257$$

$$\Sigma(X - \overline{X})^2 = (N-1)s^2 = 27(46.6257) = 1258.8933$$

Adding a score equal to the mean will not change the sum of the deviations but will increase the denominator to 28.

$$s^2_{new} = \frac{\Sigma(X - \overline{X})^2}{N-1} = \frac{1258.8933}{28} = 44.96$$

Note that the new variance is (1-1/N) times the old variance.

> The point that I was trying to make here is that adding scores that don't deviate from the mean actually decreases the variance because they decrease the average deviation from the mean.

5.17 Angle of rotation:
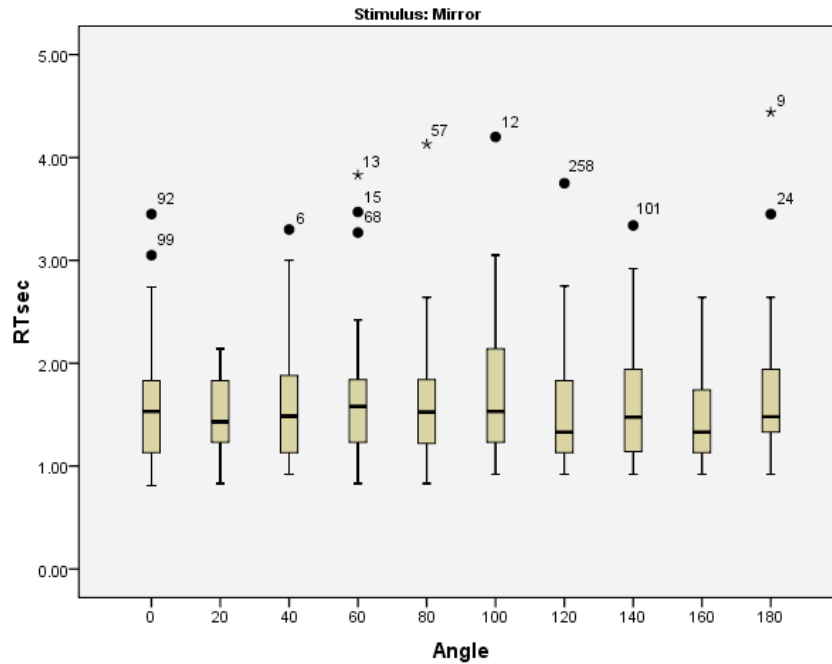
5.19   The vertical bars lie at those points that cut off the minimum, the lowest 10%, the lowest 25%, the 50% point, the lowest 75%, the lowest 90%, and the maximum score. The diamond delineates the mean and a region around that mean that we will later identify as the 95% confidence interval. The mean is at the tallest point of the diamond. That is a lot of information for one simple graphic.

5.21 Treatment of anorexia:

I hypothesize that the two treatment groups will show more of a weight gain that the

control group, but I have no reason to predict which treatment group would do better.

	Cognitive Behavior Therapy	Control	Family Therapy
Mean	3.01	-0.45	7.26
Median	1.40	-0.35	9.00
St. Dev.	7.31	7.99	7.16

If we look at the changes from Before to After it appears that the Control group stayed about the same, but the two experimental groups increased their weight. This is true whether we look at means or medians. Notice that the standard deviation in the two experimental groups was noticeably higher after treatment, whereas the standard deviation of the Control actually decreased slightly. This suggests that some participants were helped more than others by the therapies.

> We could look at weight gain, by subtracting Before from After, as was the case in this question, or we could look at percentage gain. This is too sophisticated a question for most students at this point, but it would be interesting to see if they could handle it. It would get them started thinking about how we measure anything, which in this case is change. They would see that there are different ways of going about it, and they could see that which one you choose depends on your view of what is happening. If we think that people should gain more if they are lighter, then a percentage measure would be appropriate. If we think they will each gain a fixed amount, then a difference score is relevant. I know, I'm being optimistic here, but if it can be done without looking like "statistics," then it should be useful.

5.23 The descriptive statistics from SPSS are given below. The variable labels should be clear.

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation	Variance
COGBEHAV	29	-9.10	20.90	3.0069	7.30850	53.414
TRIM	19	-1.40	11.70	1.8000	3.04211	9.254
WINSOR	29	-1.40	11.70	2.9552	4.88851	23.898
Valid N (listwise)	19					

Notice that the Winsorized variance is considerably greater than the trimmed variance, as it should be. However, it is lower than the variance of the original data, reflecting the fact

that the extreme values have been replaced. Cognitive behavior scores were positively skewed, with several quite high values and one or two low values. Trimming and Winsorizing reduced the influence of those values. This causes the Winsorized variance to be considerably smaller than the original variance. The trimmed mean is considerably smaller than the original mean, but the Winsorized mean is only slightly smaller.