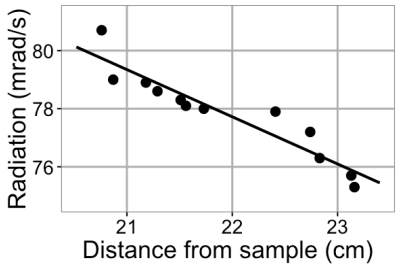
Lesson 2.1.5

2-35. a. This is a fairly strong negative linear association. It looks like the picture at right when LSRL is enabled.



b. It appears to about 2 mrad/s every cm, so the slope is approximately $-2\ \frac{\text{mrad/s}}{\text{cm}}$.

c. The value of the slope is found where "Coef" and "Distance from sample" meet. Its value is –1.62.

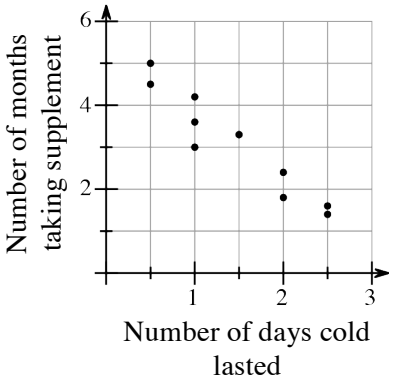
d. The slope is under "Coef" for the "distance" row, because distance is the explanatory variable and the slope is its coefficient. The *y*-intercept is found where "Coef" and "Constant" meet; its value is 113.44.

e. The line is $\hat{r} = 113.44 - 1.62d$. The slope of –1.62 means that for each centimeter further from the sample, the radiation decreases by an estimated 1.62 mrad/s. The *y*-intercept of 113.44 means if you measured the radiation 0 cm away, right on top of the sample, the expected radiation would be 113.44 mrad/s. This is an extrapolation, but could still be reasonable and is at least logically possible.

f. $S = 0.5505$ is the standard deviation of the residuals. This means that the actual radiation value is typically 0.5505 mrad/s away from the value predicted by the fit line.

g. The largest residual is 0.97 for the first point. This means that when the radiation was measured 20.76 cm away from the sample, the actual radiation was 0.97 mrad/s higher than the LSRL would predict.

2-36. a.



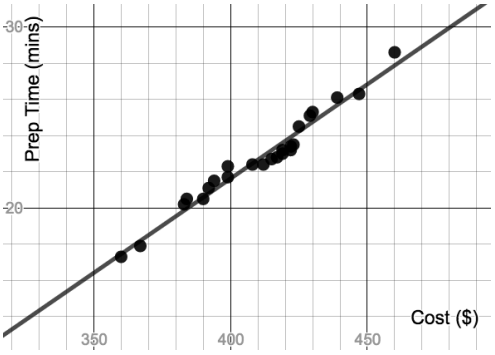
b. $\hat{c} = 5.372 - 1.581s$, where c = cold length in months and s = supplement time in days.

c. SSR = 1.34, $S = 0.409$ days. The typical difference between the actual length of a cold and the length predicted using the model is 0.409 days.

d. About 2.21 days. This seems reasonable.

e. –4 days! This is unreasonable, a classic example of extrapolation beyond the data.

2-37. b. This is a strong, positive, linear association with no obvious outlier.



c. If all has gone well, at this point you should have an image like the one at right.

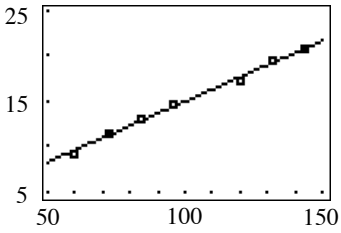
d. The residual is 0.545, meaning it took about half a minute longer than expected to prepare for this crown from the cost.

e. The equation is $\hat{p} = -20.048 + 0.10418c$. The slope of 0.10418 means each additional dollar of cost adds about 0.10 minutes, or 6 seconds, to the expected preparation time. The y-intercept means that a \$0 crown would be expected to have a –20 minute preparation time, which is absurd in both cases.

f. Just because two variables are associated does not mean one causes the other; reducing cost will not necessarily reduce preparation time and reducing preparation time will not necessarily reduce cost! They may be associated for a completely different reason.

g. $S = 0.4881$ minutes, meaning the typical preparation time is only 0.49 minutes, or about 29 seconds, from the value predicted by the cost.

2-38. a. See scatterplot at right. $y = 1.6568 + 0.1336x$. Students need to round to four decimals because if they round to fewer decimals, the LSRL gets too far away from the actual points due to the lack of precision.



b.

Distance from wall (in)	Residual (in)
144	–0.198
132	0.305
120	–0.391
96	0.316
84	0.219
72	0.123
60	–0.374

Sum of the squares is 0.5881 in^2.

c. S is $\sqrt{\frac{0.5881}{5}} = 0.343$ in. The typical difference between the actual field of view and the field of view predicted by the model is about 0.343 in.

2-39. a. Answers will vary. Some people may say negative because people like to feel like they are getting a bargain, or some may say positive because a higher price may indicate better ingredients, service, or atmosphere.

b. For each additional dollar a coffee beverage costs the satisfaction rating typically increases by 0.42 points.

c. When using price to predict customer satisfaction ratings, one can typically expect to be off by 1.30 points on a scale of 1 to 10.

d. 6.5

2-40. a. The LSRL is $\hat{m} = 23.04 - 0.024h$. The slope is much closer to 0 because the outlier that was ignored in the original line pulls the right side of the line towards it, lowering the slope. This also decreases the *y*-intercept, because when the right side gets pulled up, the left side of the line drops.

b. The typical car's mileage is 3.898 mpg away from the value estimated by the LSRL.

c. The LSRL should get much closer to the original eyeballed model, with the *y*-intercept dropping and the slope getting more negative. *S* should decrease substantially.

d. The new LSRL is $\hat{m} = 26.1 - 0.038h$. The new value of *S* is 2.5 mpg, significantly lower than before.

2-41. a. No, because 30 is the third quartile of the data. This means that only about 25% of the students walked more than 30 laps; the actual percentage of students varies with the number of data points.

b. This portion represents roughly 50% of the students; it could slightly vary depending on the number of data points.

c. You could change one of the data points between 30 and the maximum, but keep in the same range of values. You could move a data point from below the median but above the first quartile to above the median but below the third quartile.

d. No, you cannot determine this information from the data in the boxplot.

