Homework №5

Author: David Oniani
Instructor: Dr. Eric Westlund

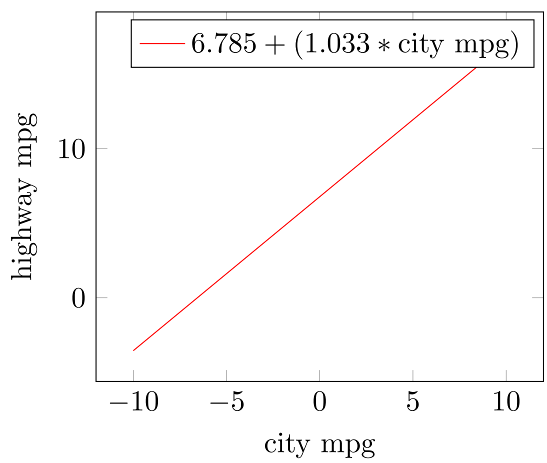
February 20, 2019

5.1 (a) The slope is 1.033. Slope is the rate of change. Particularly, it tells us that highway mileage increases by 1.016 for every 1 mpg change in the city mileage.

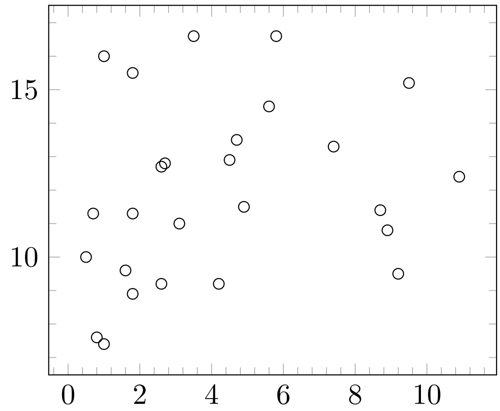
(b) The intercept is 6.785. This would be a mileage for a car with 0 mpg and such car would probably not be present.

(c) Let's plug the numbers. For city mpg $= 16$, we get highway mpg $= 6.785 + (1.033 \times 16) = 23.313$. And for city mpg $= 28$, we get highway mpg $= 6.785 + (1.033 \times 28) = 35.709$.

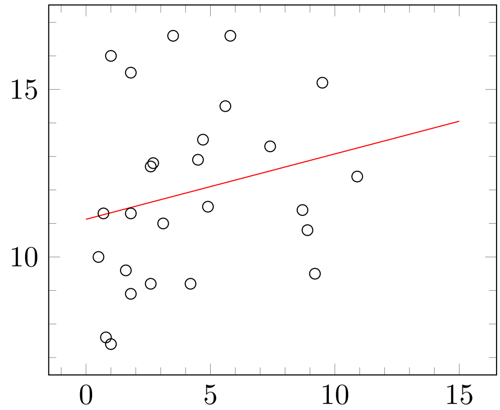
(d)



5.5 (a) Below is the scatterplot for the data.



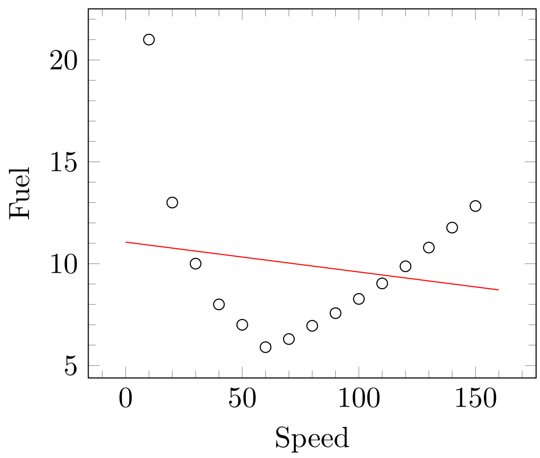
(b) Using the `Haskell` library `Statistics.LinearRegression` (link: `https://bit.ly/2BLQ7Qb`), we get the regression line $y = 11.125 + 0.195x$. Below is the plot with the line added.



(c) It tells us the expected increase. Particularly, it tells us that for every suicide, there are approximately 0.195 homicides.

(d) We can just plug the value in the regression equation. We get: $\hat{y} = 11.125 + 0.195 \times 8.0 = 12.685$. Therefore, the rate is 12.685 per 100, 000 people.

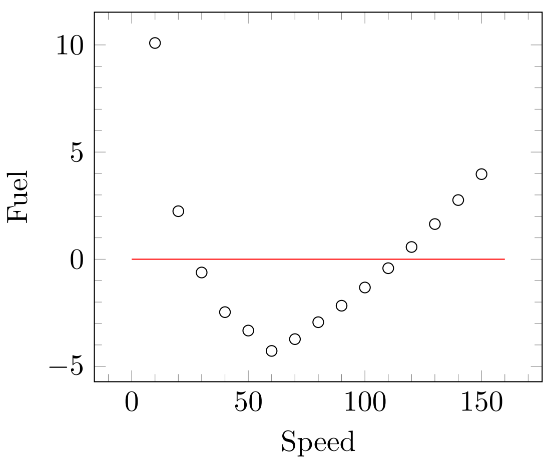
5.9 (a)



(b) It is clear that the relationship between the speed and fuel is parabolic. Therefore, predictions based on the output of the give regression line would be inaccurate. Hence, the answer is NO.

(c) We have: $11.058 - 0.01466 \times 10 = 10.9114$. Then, we know that the residual is predicted value minus the actual value which gives us $21 - 10.9114 = 10.0886$. After rounding 10.0886, we indeed get 10.09 which is the residual value on the table.

(d)



The pattern seems to be virtually the same as the pattern exhibited in the part (a) of this exercise.

5.15 (a) Once again, using the software, we get that the regression line is $\hat{y} = -44.831 + 0.132x$.

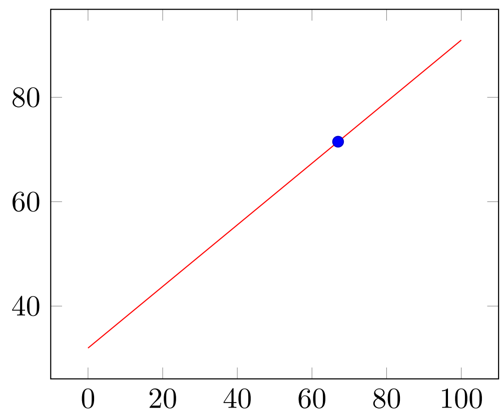
(b) The prediction would be $-44.831 + 0.132 \times 890 \approx 73$. We can trust this prediction since 890 is in-between (but not outside) the table values (observed values).

(c) That happens if $x = 0$. In this case we get that -44.831 manatees are killed which does not seem reasonable. This happens since 0 is outside the range of table values.

5.17 Income, parents' (or mother's) IQ. There might also be some other genetic reasons which did not show up in mother but did in her child(ren).

5.34 (a) The slope is $m = r \times \frac{s_y}{s_x} = 0.5 \times \frac{3.1}{2.7} \approx 0.59$.
The intercept is $b = \bar{y} - m \times \bar{x} = 69.9 - 0.59 \times 64.3 = 31.96$.

(b) The regression line equation is $\hat{y} = 31.96 + 0.59x$. For $x = 67$, we get ≈ 71.5.



(c) The regression line only explains r^2 which is 25% of the variation in the height of the husband.

5.35 (a) The slope is $m = 0.5 \times \frac{8}{40} = 0.1$.
The intercept is $b = 75 - 0.1 \times 280 = 47$.
The regression line equation is $\hat{y} = 47 + 0.1x$

(b) We just plug in the value 300 and get $47 + 0.1 \times 300 = 77$.

(c) I actually agree with Julie. The regression line explains only r^2 which is 25% variation in student final exam scores.

5.55 Predicted second-round score for a player who shot 80 is $56.47 + 0.243 \times 80 = 75.91$.
Predicted second-round score for a player who shot 70 is $56.47 + 0.243 \times 70 = 73.48$.

The player who shot 80 in the first round is predicted to score below the average, but overall, a better score.
The player who shot 80 in the first round is predicted to score above the average, but overall, a worse score.

5.56 We know that the point $(\overline{x}, \overline{y})$ lies on the least squares regression line. Then equality $\overline{y} = 46.6 + 0.41\overline{x}$ holds. Since Octavio scored $\overline{x}$ 10 points above the mean, it means that he scored $\overline{x} + 10$. We can now plug this value in the regression equation to get $\hat{y} = 46.6 + 0.41 \times (\overline{x} + 10) = 46.6 + 0,41\overline{x} + 4.1 = \overline{y} + 4.1$. Finally, we got that $\hat{y} = \overline{y} + 4.1$ which means that Octavio's final exam score is predicted to be 4.1 points above the class mean on that final exam.