Homework due March 2

Assigned exercises: Ch.12, OpenStax book, pg. 719-721, ex. 4, 7, 57, 60, 65, 66.
And from linked supplement: 6, 10, 12, 25, 27, 29, 31, 35, 37. (15 exercises total)
Graded exercises: From Ch.12, OpenStax book: 57, 65, 66.
From linked supplement: 6, 29.
Total (maximum) possible points = 20.
3 pt for each of 5 graded problems, plus 5 for completion of the rest.
-0.5 pt for each (ungraded) missing problem; if a graded problem is missing, student loses the 3 points allotted to it.

Exercises from Ch.12, OpenStax

(57) (a) Independent variable = age of driver. Dependent = # of fatalities per 100,000 drivers.

(b) Independent variable = number of family members. Dependent = weekly grocery bill.

(c) Independent variable = age of applicant. Dependent = life insurance premium.

(d) and (e) are not graded. But, reasonable answers would be

(d) Independent variable = power consumption. Dependent = utility bill.

(e) Independent variable = higher education level, quantified in some way. Dependent = crime rate.

Grade: (a)=(b)=(c)=1 pt. (d), (e) not graded.

(65) An r^2 value of 0.72 implies that 72% of the variability in the response is explained by the linear model (OR: ... is explained by the variability in the independent variable.)

Grade: There are so many different correct ways to say this that I don't know how to write precise grading instructions! Any answer that captures the key ideas is sufficient. Words like "variability" can be expressed in other ways (e.g., variance, variation). "Response variable" is the same as "dependent variable," etc.

(66) A coefficient of determination is always positive, because it is the square of the correlation.

Grade: 2 pt for correct answer; 1 pt for providing some reason or justification.

Exercises from linked supplement

(6) (a) The slope is given by $b_1 = r\frac{s_y}{s_x}$, where r = correlation, s_x = SD of x-variable, and s_y = SD of the y-variable. Plugin given numbers and get

$$b_1 = 0.994\left(\frac{151.26}{1.4469}\right) = 103.91 \text{ \$ per terabyte}$$

(b) The slope means: For each 1 terabyte increase in the capacity of a drive, the price is predicted to increase by about \$104, on average.

(c) Plug mean values into $\hat{y} = b_0 + 103.91x$ to find b_0

$$133.98 = b_0 + 103.91\ (1.110) \ \Rightarrow \ b_0 = 133.98 - 103.91\ (1.110) = 18.64\ \$$$

(d) b_0 is meaningless in this application context, because it corresponds to the predicted price of a drive with 0 capacity.

(e)

$$\widehat{\text{price}} = 18.64 + 103.91 \times \text{capacity}$$

(f) To predict the price of a 3.0 TB drive, we just plugin and get

$$\widehat{\text{price}} = 18.64 + 103.91 \times 3.0 = \$\ 330.38$$

Grade: Only (a)-(c) are graded. (a)=(b)=(c)=1 pt each.

(29) (a) Explanatory variable = Size of a home (in square feet)
Response variable = Its price (in thousands of \$)

(b) Slope units = 1000 dollars per square foot OR $\dfrac{1000 \text{ dollars}}{\text{feet}^2}$

(c) The slope will be positive, since homes with larger square footage will generally cost more than smaller homes.

Grade: (a)=2 points, (b)=1 point, (c) is not graded.
For (a): 0.5 pt each for explanatory, its units, response, its units.
For (b): 0.5 pt each for correct numerator + correct denominator.