2.

Disk drives have been getting larger. Their capacity in now often given in terabytes (TB) where 1 TB = 1000 gigabytes, or about a trillion bytes. A survey of prices for external disk drives found the following data:

|  |  |
| --- | --- |
| Capacity (in TB) | Price (in $) |
| .080 | 29.95 |
| .120 | 35.00 |
| .200 | 299.00 |
| .250 | 49.95 |
| .320 | 69.95 |
| 1.0 | 99.00 |
| 2.0 | 205.00 |
| 4.0 | 449.00 |

4.

A company that relies on internet-based advertising linked to key search terms wants to understand the relationship between the amount it spends on this advertising and revenue (in $).

1. Which variable is the explanatory or predictor variable?
2. Which variable is the response variable?
3. Which variables would you plot on the *x* axis?

14.

For the disk drives in problem 2 above, we want to predict *Price* from *Capacity.*

1. Find the slope estimate, *b*1.
2. What does it mean in this context?
3. Find the intercept, *b*0.
4. What does it mean in this context? Is it meaningful?
5. Write down the equation that predicts *Price* from *Capacity*.
6. What would you predict for the price of a 3.0 TB disk?
7. You have found a 3.0 TB drive for $300. Is this a good buy? How much would you save compared to what you expected to pay?
8. Does the model overestimate or underestimate the price?

54.

An online clothing retailer examined their transactional database to see if total yearly *Purchases* ($) were related to customers’ *Income*  ($). (You may assume that the assumptions and conditions for regression are met). The least squares linear regression is:

*Purchase* = -31.6 + 0.012 *Income.*

1. Interpret the intercept in the linear model.
2. Interpret the slope in the linear model.
3. If a customer has an *Income* of $20,000, what is his predicted total yearly *Purchases*?
4. This customer’s yearly Purchases were actually $100. What is the residual using this linear model? Did the model provide an underestimate or overestimate for this customer?