Martin Russell

Unit 2 Lesson 3 Drill

1. The amount a person will spend on a given site in the next 24 months.
    a. A regression model would be used, the amount spent could be zero or ten billion, or somewhere in-between or even higher.
2. What color car someone is going to buy.
    a. This would be a classification model, the colors that are available for each model for any given year are finite, usually less than 20, but at least one. In theory, of car companies mixed colors, you could come up with hundreds of shades of each color, in which case this would be a regression model, but in real life, it would be classification.
3. How many children a family will have.
    a. We would always be dealing in a whole number, and the number would rarely exceed eight or nine, this would use a classification model.
4. If someone will sign up for a service.
    a. This is a “yes” or “no” question, which gives us two categories, making this use a classification model.
5. The number of times someone will get sick in a year.
    a. If we think of a year as a continuous line of days, we are really saying how many possible times can a person be sick along this continuous line, this would call for a regression model.
6. The probability someone will get sick in the next month.
    a. A shorter timespan than a year, how many times could an event occur along a line going from one to thirty? Regression could be used, but since the time period is so short, classification could also be used.
7. Which medicine will work best for a given patient
    a. Most likely a handful of drugs being used to treat a specific illness for ONE specific patient, I would use a classification model. However if we were talking about which drugs would work for the ENTIRE population, if you take into account genetics, allergies, interactions with other drugs, etc., then perhaps a classification model would work best.