**Unit 4, Module 11: Scatterplots, Linear Relationships and Correlation**

1. Suppose you were to collect data for each of the following pairs of variables. You want to make a scatterplot. Which variable would you use as the explanatory variable and which as the response? Why? What would you expect to see in the scatterplot? Discuss the likely direction and strength.
	1. **Lightning Strikes**: distance from lightning, time delay of thunder
	2. **Streetlight:** its apparent brightness, your distance from it
2. Which of the following scatter plots show (you can use the graphs multiple times, if needed):
	1. Little or no association
	2. A negative association?
	3. A linear association?
	4. A moderately strong association?
	5. A very strong association?
3. Here are several scatterplots. The calculated correlations are: -0.977, -0.021, 0.736 and 0.951. Which correlation belongs to which graph?



1. A survey was conducted in the US and 10 countries of Western Europe to determine the percentage of teenagers who had used marijuana and other drugs. The results are summarized in the table below:

|  |  |
| --- | --- |
|  | **Percent Who Have Used** |
| **Country** | **Marijuana** | **Other Drugs** |
| Czech Rep. | 22 | 4 |
| Denmark | 17 | 3 |
| England | 40 | 21 |
| Finland | 5 | 1 |
| Ireland | 37 | 16 |
| Italy | 19 | 8 |
| No. Ireland | 23 | 14 |
| Norway | 6 | 3 |
| Portugal | 7 | 3 |
| Scotland | 53 | 31 |
| USA | 34 | 24 |

* 1. Create a scatterplot of the data



* 1. What is the correlation between the percent of teens who have used marijuana and the percent who have used other drugs? (Use technology for this!)
	2. Write a brief description of the association (direction, form, strength)
	3. Do these results confirm that marijuana is a “gateway drug”, that is, that marijuana use leads to the use of other drugs? Explain.

**Unit 4, Module 12:** Fitting a Line

5. A regression analysis of the initial drop (in feet) of a roller coaster ride and the duration of the ride (in seconds) yields the model: , where *Duration* is the response variable and *Drop* is the explanatory variable.

* 1. Explain what the slope of the line says about how long a roller coaster ride may last and the height of the coaster.
	2. Explain what the y-intercept of the line means in this situation. Does this make sense for roller coasters in “real life”?
	3. A new roller coaster advertises an initial drop of 200 feet. How long would you predict the ride lasts?