Regression Diagnostics Applications

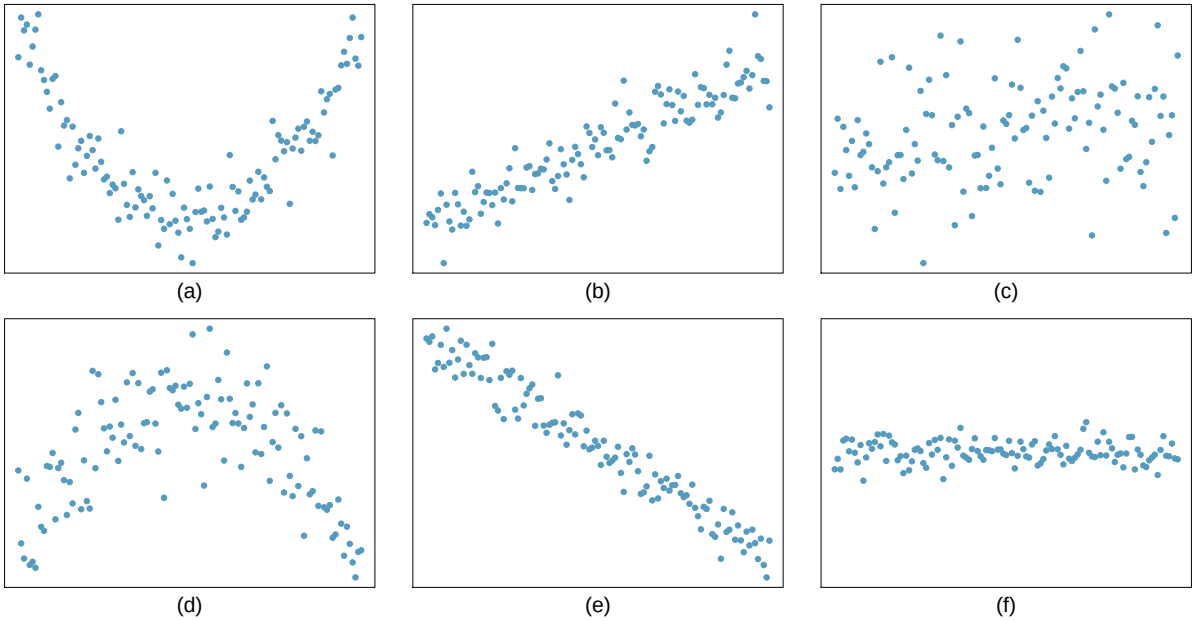
YOUR NAME

09 November, 2020

Exercises

1. Identify relationships

For each of the six plots, identify the strength of the relationship (e.g. weak, moderate, or strong) in the data and whether fitting a linear model would be reasonable.



2. Beer and blood alcohol content

We will use the blood alcohol content data again. As a reminder his is the description: *Many people believe that gender, weight, drinking habits, and many other factors are much more important in predicting blood alcohol content (BAC) than simply considering the number of drinks a person consumed. Here we examine data from sixteen student volunteers at Ohio State University who each drank a randomly assigned number of cans of beer. These students were evenly divided between men and women, and they differed in weight and drinking habits. Thirty minutes later, a police officer measured their blood alcohol content (BAC) in grams of alcohol per deciliter of blood.*

The data is in the `bac.csv` file under the `data` folder.

a. Obtain and interpret R-squared for this model.

b. Evaluate the assumptions of this model. Do we have anything to be concerned about?

3. Outliers

Identify the outliers in the scatterplots shown below and determine what type of outliers they are. Explain your reasoning.