Motivation: How Useful Is A Model?

We have a data set with observations of four variables measuring advertising budgets and sales for a product in each of 200 markets:

- `sales` is a measure of sales volume in thousands of units
- `TV` is TV advertising budget
- `radio` is radio advertising budget
- `newspaper` is newspaper advertising budget

Below is R code for making plots displaying three separate simple linear regression fits to the data (the actual plots are on the other side of the page). In all three plots/models, `sales` is the response variable; the explanatory variable is different for each model.

```r
library(readr)
library(ggplot2)
library(gridExtra) # for grid.arrange, puts plots next to each other

Advertising <- read_csv("http://www.evanlray.com/data/islr/Advertising.csv")

p1 <- ggplot(data = Advertising, mapping = aes(x = TV, y = sales)) +
  geom_point() +
  geom_smooth(method = "lm", se = FALSE) +
  theme_bw(base_size = 14)

p2 <- ggplot(data = Advertising, mapping = aes(x = radio, y = sales)) +
  geom_point() +
  geom_smooth(method = "lm", se = FALSE) +
  theme_bw(base_size = 14)

p3 <- ggplot(data = Advertising, mapping = aes(x = newspaper, y = sales)) +
  geom_point() +
  geom_smooth(method = "lm", se = FALSE) +
  theme_bw(base_size = 14)

grid.arrange(p1, p2, p3, nrow = 1)
```
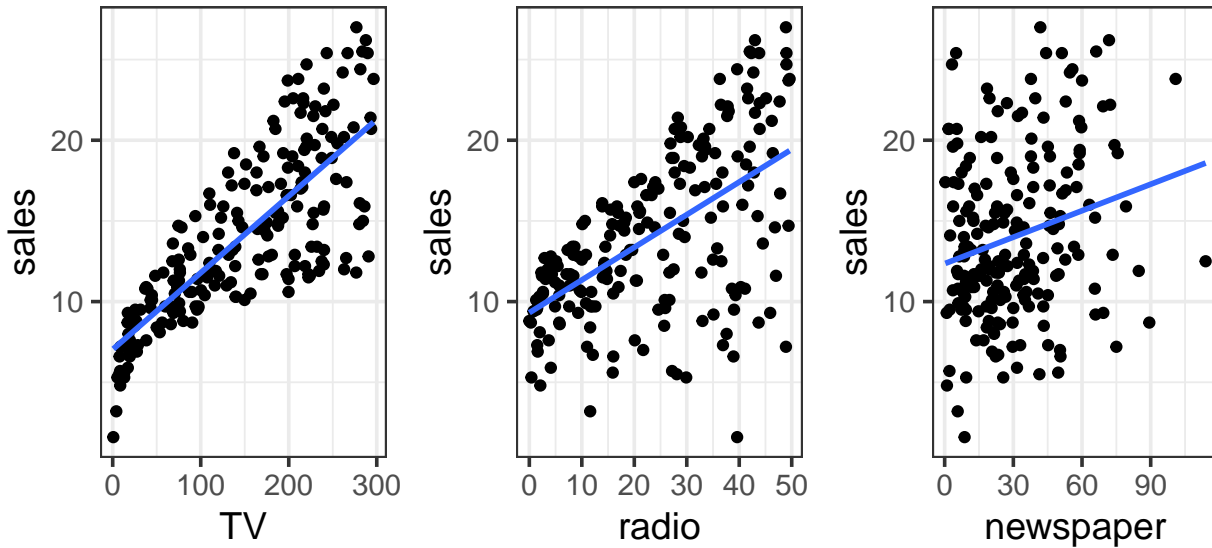
With your neighbors, discuss which of these models would you prefer to use for predicting sales and why. Then answer the questions below:

Being as specific and concrete as possible, write down a rule for selecting your preferred model based only on *visual* characteristics of the plot. (That is, your rule should not involve any calculations of numeric quantities).

Being as specific and concrete as possible, write down a rule for selecting your preferred model based only on a *quantitative* summary of the data. You can describe how you would calculate your numeric summary of the data in a general sense; if you'd like you can write down a formula.