Chriss Jordan Oboa

Professor Tysse

Math213

Mars 25th, 2020

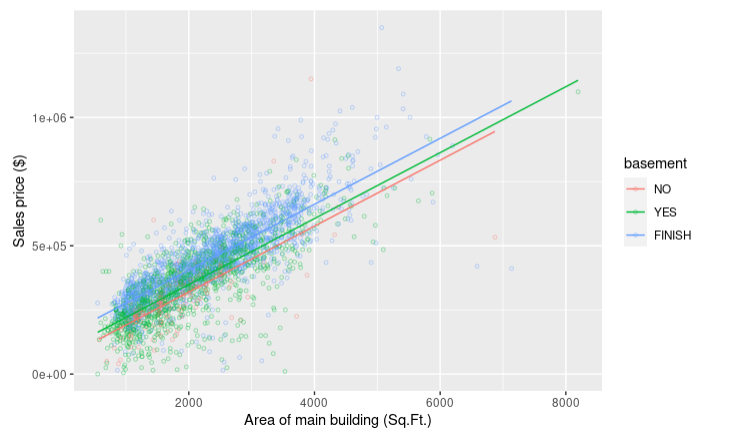
# Homework 10

1. This question pertains to a dataset (that we saw in class before the break and on HW 9) including records of sales of residential properties in Frederick County from 2019.

Here is part of the codebook for the dataset:

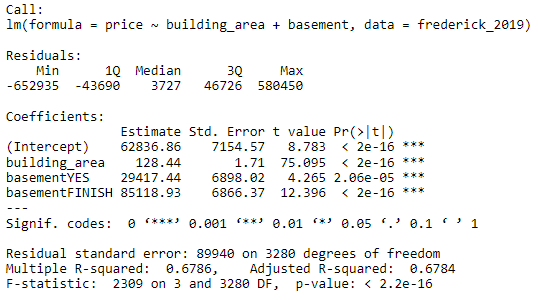
* `price`: the sales price of the property
* `building\_area`: the area of the main building on the property, in square feet, not including "below grade" space such as a basement.
* **`basement`: whether the building has a finished basement (FINISH), unfinished basement (YES), or no basement (NO).**

I made the following graph and created a parallel slopes model:



Please note that there are THREE levels for the variable basement, so there are THREE parallel lines that make up this parallel slopes model.

Here is the R model summary output:



1. Write the regression equation using proper statistical notation.

Price\_hat = 62836.86 + 128.44 \* (building\_area) + 29417 \* (basementYES) + 85118.93 \* (basementFINISH)

1. Write four separate sentences to interpret the meanings of the coefficients in the models.

*(Note: If you are not sure how R’s new variables basementYES and basementFINISH work, you should ask \*before\* turning in this assignment!)*

The coefficient 29417 is based on whether the basement is incorporated in the house.

The coefficient 85118.93 is based on whether the basement is finished when the basement is incorporated in the house.

1. Researchers at the University of Pennsylvania conducted a study of an approach to weight loss inspired by behavioral economics in 2007-8. They randomly assigned participants in their study to one of three weight-loss plans. One weight-loss plan (the control) involved only weekly weigh-ins. The other two weight-loss plans both involved financial incentives, designed following the advice of behavioral economists, to encourage weight loss.

The researchers recruited participants at the Philadelphia Veteran’s Affairs Medical Center in May through August 2007 and followed participants through June 2008. Participants were required to be healthy adults, aged 30 to 70, with BMI between 30 and 40. The authors described their sample as follows:

No significant differences were found in the baseline characteristics of any of the groups. The sample was predominantly male, with total household incomes of about $30,000, and a mean enrollment BMI of between 33.8 and 35.5 across the 3 groups. Participants in all 3 groups rated the importance of controlling their weight to be high on a 10-point scale (9.11-9.31) and had high confidence in their ability to lose weight (8.32-8.47).

Before starting on the plans, all of the participants had individual meetings with dieticians to discuss diet and exercise strategies for losing weight. The main part of the study gathered data on weight loss over 16 weeks, during which participants were given the goal of losing 1 pound per week. Participants in the control group only had weekly weight checks, but the two incentive plans provided different types of financial incentives for participants to meet their weekly weight-loss goals. We will focus on comparing the control plan with one of the incentive plans, which the authors call “Deposit Contract.”

1. Who or what were the cases in this study?

The cases in this study are Healthy adults, aged 30 to 70, with BMI between 30 and 40

1. What are the explanatory and response variables in this study? (You may need to revise what you write about the response variable after you get more details about the study below.)

Explanatory: Group assigned based on their diet and exercise strategies for losing weight

Type (categorical or numerical):

Response: Total Weight-loss goals in pounds after 16 weeks

Type (categorical or numerical):

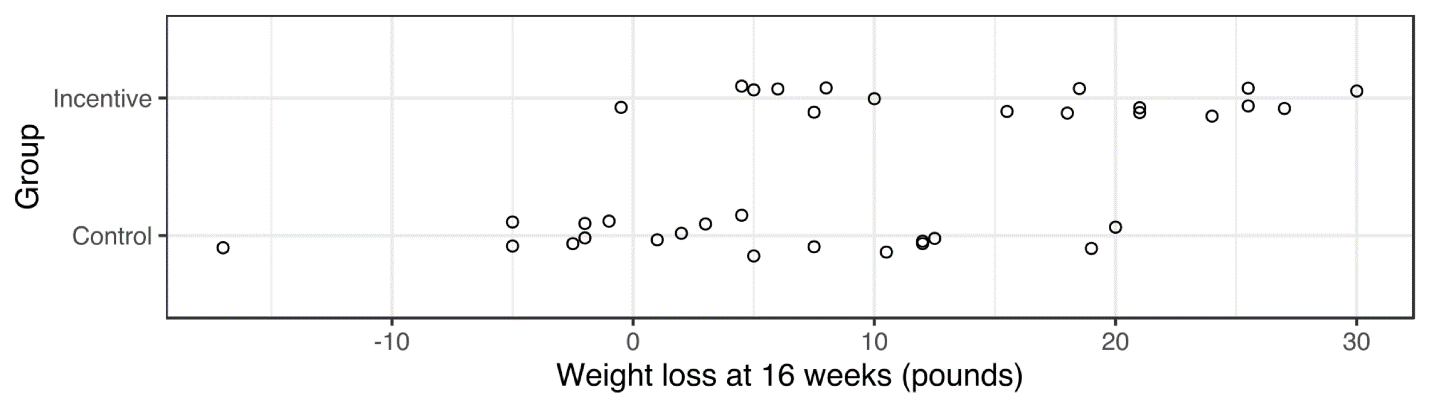
1. What makes this study an experiment? Explain briefly.

The researchers imposing the explanatory variable on the subjects, and they use random assignment

# Looking at the data

The following jittered plot and table display the data from the study. There were originally 19 participants in each group, but two of the “incentive” participants dropped out.

| Group | n | group mean weight loss at 16 weeks (pounds) | group std. dev. weight loss (pounds) |
| --- | --- | --- | --- |
| Control | 19 | 3.92 | 9.11 |
| Incentive | 17 | 15.68 | 9.41 |



1. Compare the distributions of weight loss in the two groups. Focus on the center and the variability (how spread out the dots are). What preliminary conclusions do you draw from this summary of the study?

**In the control group, the center or the mean is approximate to 3.92 pounds loss which is lower compared to the incentive group mean which is at about 15.68 pounds loss. The incentive group has also a greater standard deviation compared to the control group in which shows us how much or where the one who lost the most or less pounds is compared the average pounds loss of the initial group. The incentive standard deviation is at 9.11 and the control group standard deviation is at 9.41.**

**In the control group, about 1 to 13 pounds were lost by the majority participants. Around 25% of participants gain about 1 to 5 pounds after 16 weeks of the experiment. However, only 2 participants had exceptional weight lost compared to the rest in which one lost about 19 pounds and the other 20 pounds. On the contrary, 7 participants failed to lose weight during this experiment. The majority of participant who failed, gained about 1 to 5 pounds. However, there is one participant who gained about 17 pounds.**

**In the incentive group, about 15 to 25 pounds were lost by the majority participants. On the other side, around 30% of participants lost about 4 to 10 pounds after 16 weeks of the experiment. However, only one participant had exceptional weight lost compared to the rest in which he or she lost about 30 pounds. On the contrary, one participant exceptional failed to lose weight during this experiment and gained about 1 pounds.**