Confounders and Colliders

Alex Alekseev

```r
library(tidyverse)
```

In this exercise we will explore conditioning on variables. In the first case, we will condition on a so-called *confounder*. This case will illustrate a well-known Simpson's paradox in statistics. In the second case, we will condition on a so-called *collider*, which will lead to a so-called Berkson's paradox.

Confounders

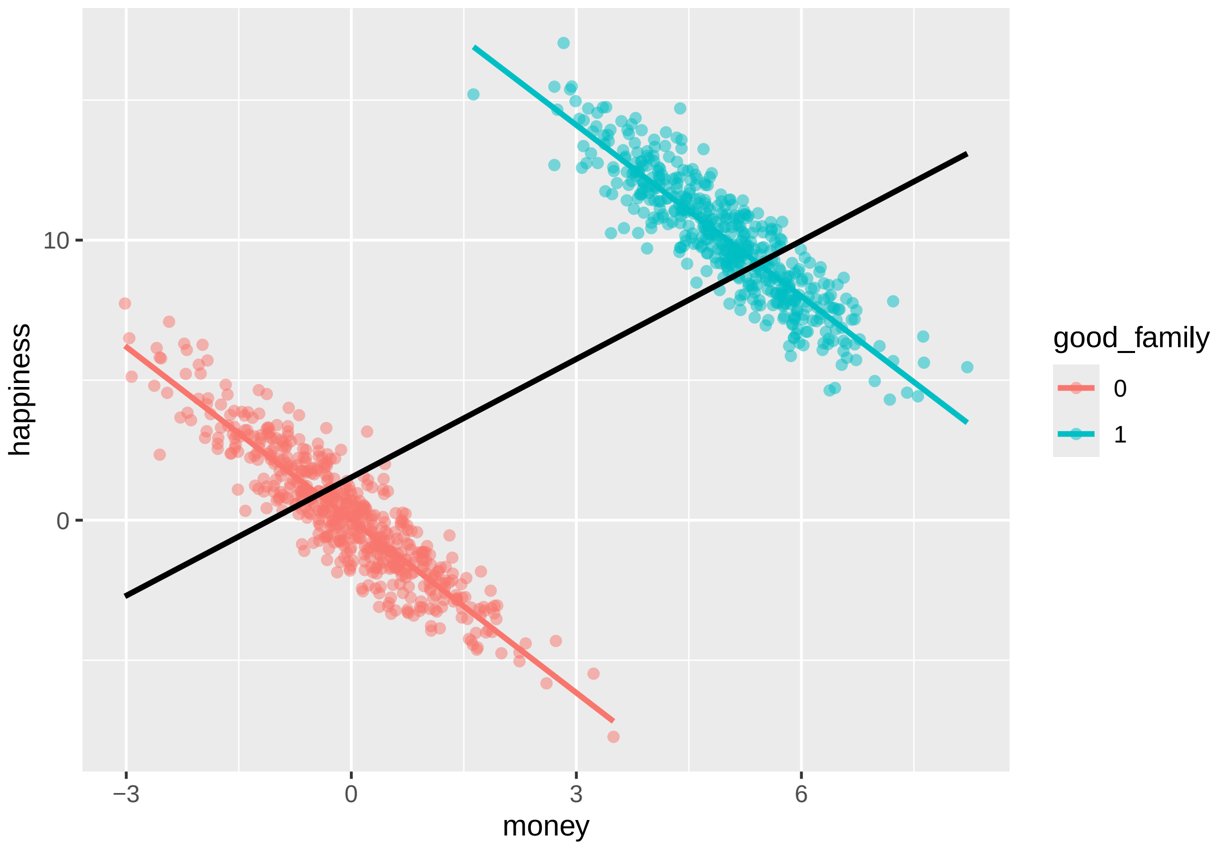
Suppose we are studying the causal effect of money on happiness. We will generate some simulated data. In this dataset, money will have a truly negative effect on happiness. Let's assume that we also have a third variable: whether a person is from a "good" family. This variable will affect both the amount of money a person has (people from good families have more money than people from not so good families) and a person's happiness (people from good families are happier on average). For example, we can generate the following data.

```r
nobs <- 1000

set.seed(42)
df <- tibble(
  good_family = 1*(runif(nobs) >= 0.5)
  , money = rnorm(nobs) + 5*good_family
  , happiness = 20*good_family + (-2)*money + rnorm(nobs)
) %>%
  mutate(good_family = factor(good_family))
```

Now we will illustrate graphically what happens when you estimate the "naive," unconditional effect of money on happiness and when you condition on the family background.

```r
df %>%
  ggplot(aes(money, happiness)) +
  geom_point(aes(color = good_family), alpha = 0.5) +
  geom_smooth(aes(color = good_family), method = "lm", se = F) +
  geom_smooth(method = "lm", se = F, color = "black")
```



As you can see, the black line (no conditioning) shows that if we simply regress happiness on money, we will observe a positive relationship. However, once we condition on family background, we see that the true effect is indeed negative. This is an example of Simpson's paradox: your conclusion about the effect of one variable on another can be reversed after conditioning on the confounder.

Colliders

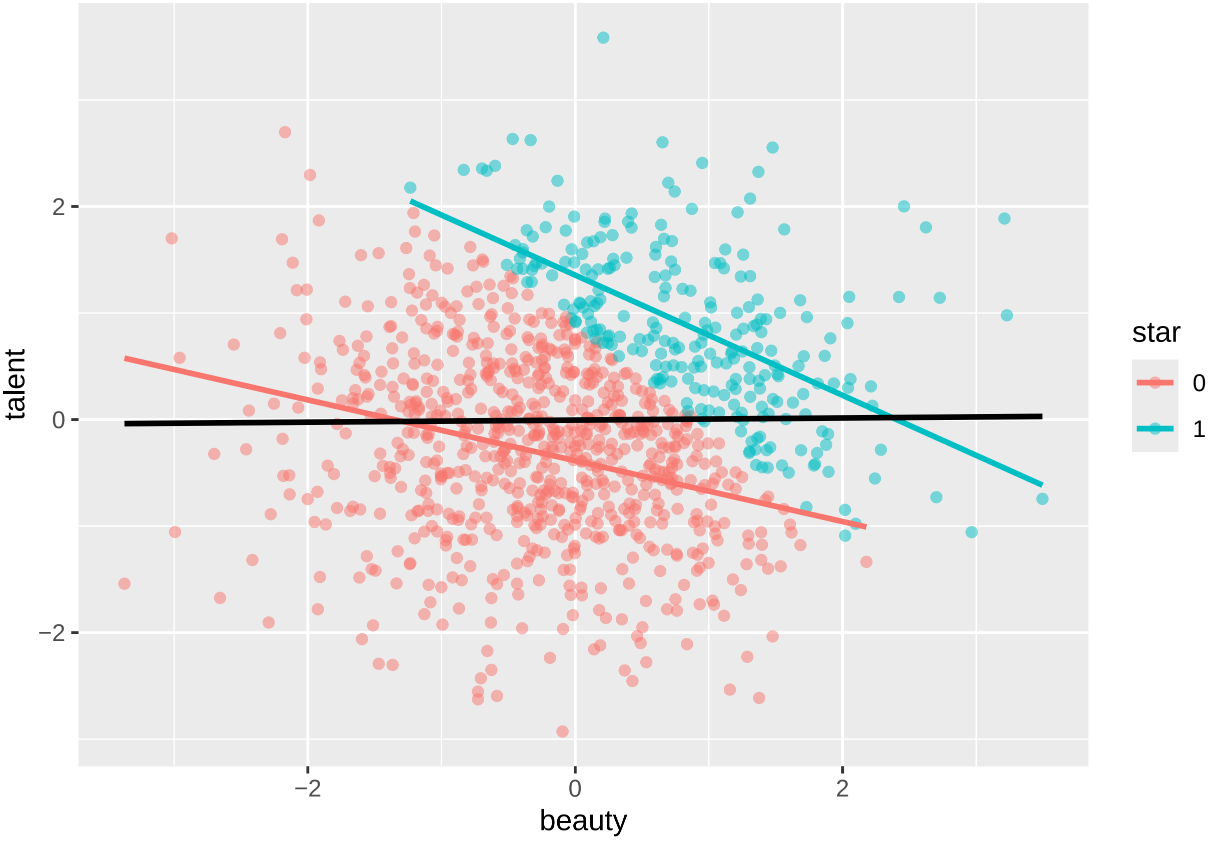
Suppose we are interested in the relationship between talent and beauty for movie actors and actresses. Let's assume the following data generating process. Beauty and talent are independent standard normal random variables. A person's "score" is an average of their beauty and talent. A person becomes a movie star if their score is in the top quartile.

```
nobs <- 1000

set.seed(42)
df <- tibble(
  beauty = rnorm(nobs)
  , talent = rnorm(nobs)
  , score = 0.5*beauty + 0.5*talent
  , star = 1*(score >= quantile(score, 0.75))
) %>%
  mutate(star = factor(star))
```

Now let's plot the relationship between beauty and talent with and without conditioning on whether a person is a movie star.

```r
df %>%
  ggplot(aes(beauty, talent)) +
  geom_point(aes(color = star), alpha = 0.5) +
  geom_smooth(aes(color = star), method = "lm", se = F) +
  geom_smooth(method = "lm", se = F, color = "black")
```



Without conditioning on being a movie star, the relationship between talent and beauty is basically zero. We know that this is the true effect, because in our DGP the two variables are unrelated. However, once we condition on being a movie star, we see a negative relationship between beauty and talent. This is an example of conditioning on a collider: both beauty and talent cause the movie star status, however, the two are actually independent.