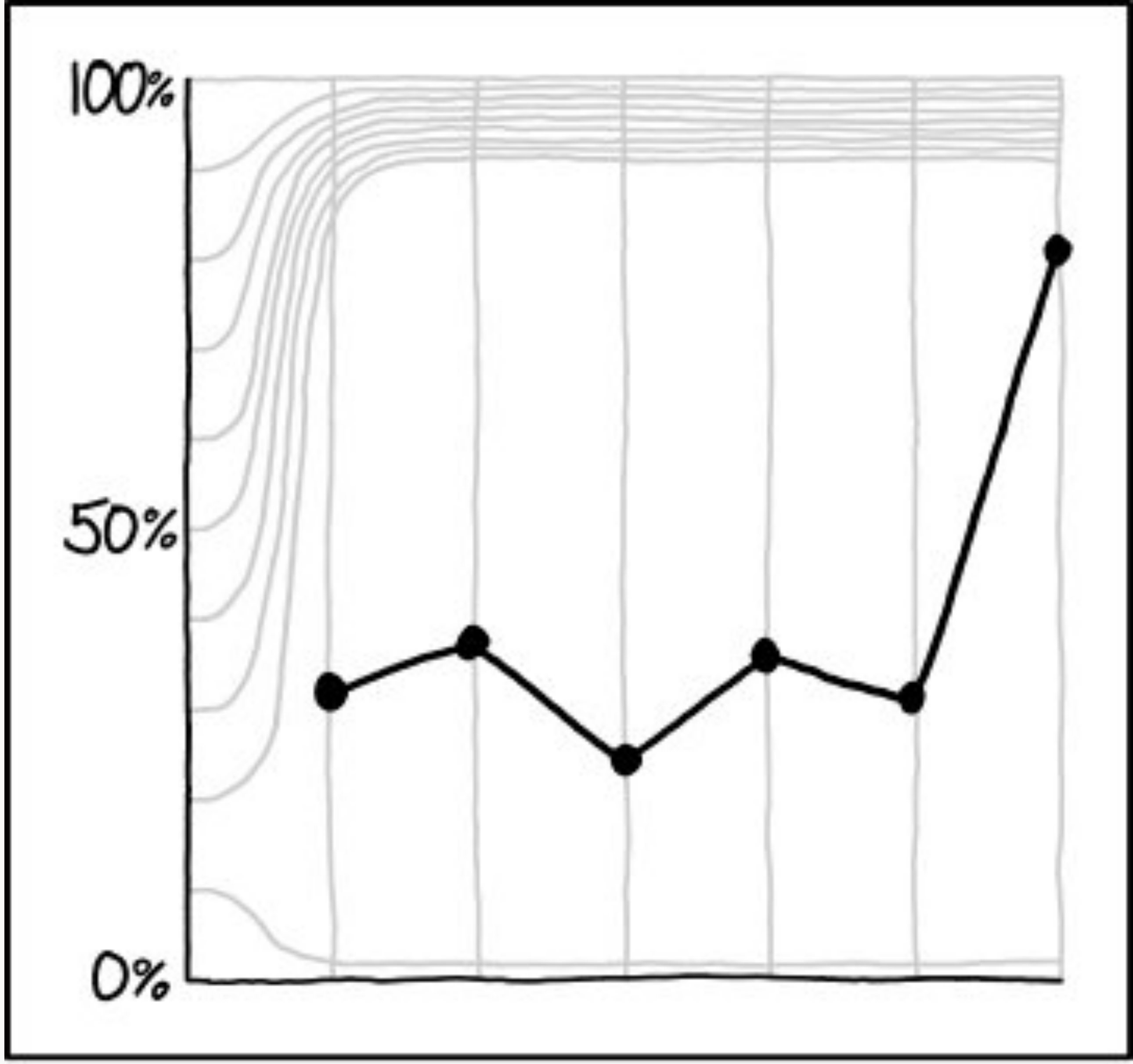
Don’t fall for all their tricks!



PEOPLE HAVE WISED UP TO THE "CAREFULLY CHOSEN Y-AXIS RANGE" TRICK, SO WE MISLEADING GRAPH MAKERS HAVE HAD TO GET CREATIVE.

EDUC 7610

Chapter 13 and 14

Interactions (Moderation)

Fall 2018

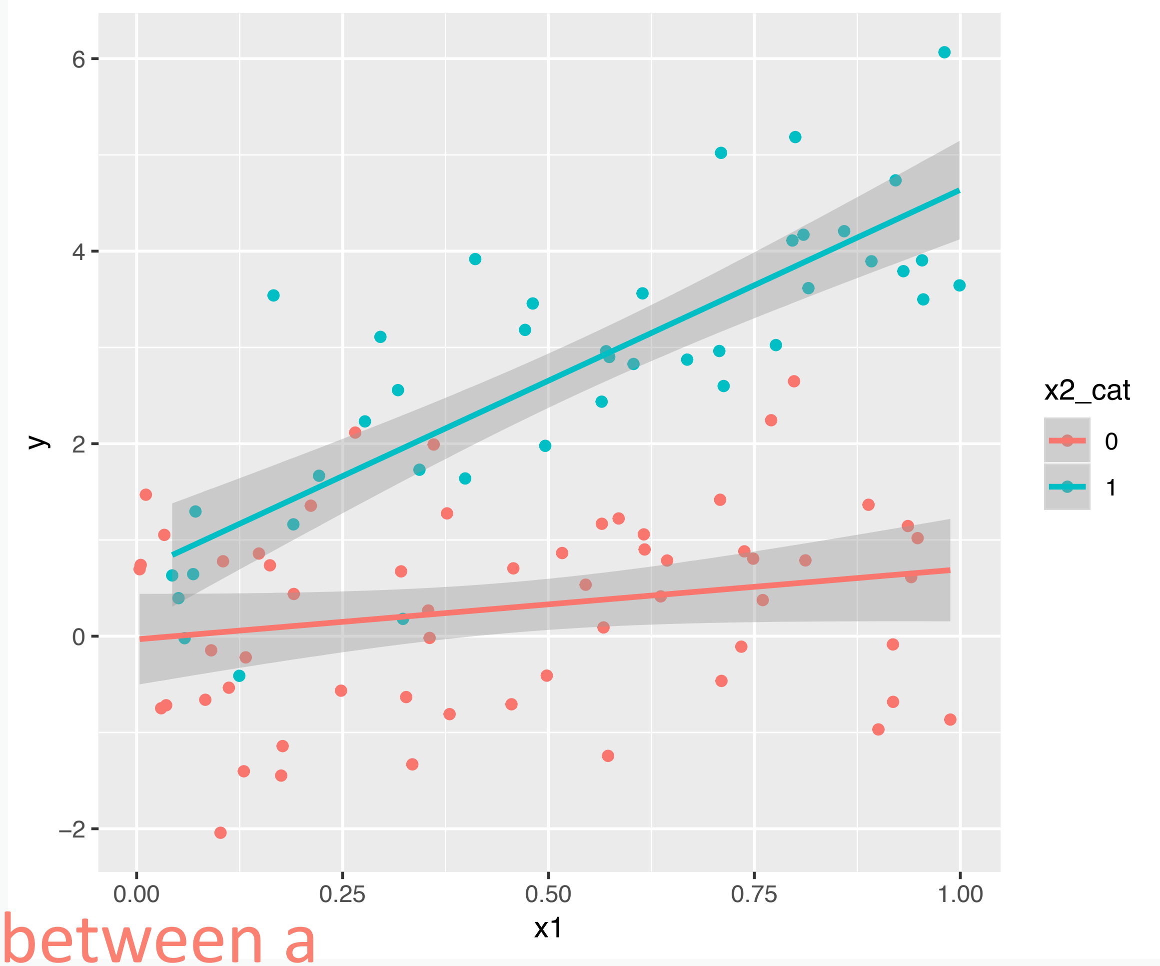
Tyson S. Barrett, PhD

What is meant by "interaction"?

When the relationship between a predictor and the outcome *depends* on another variable

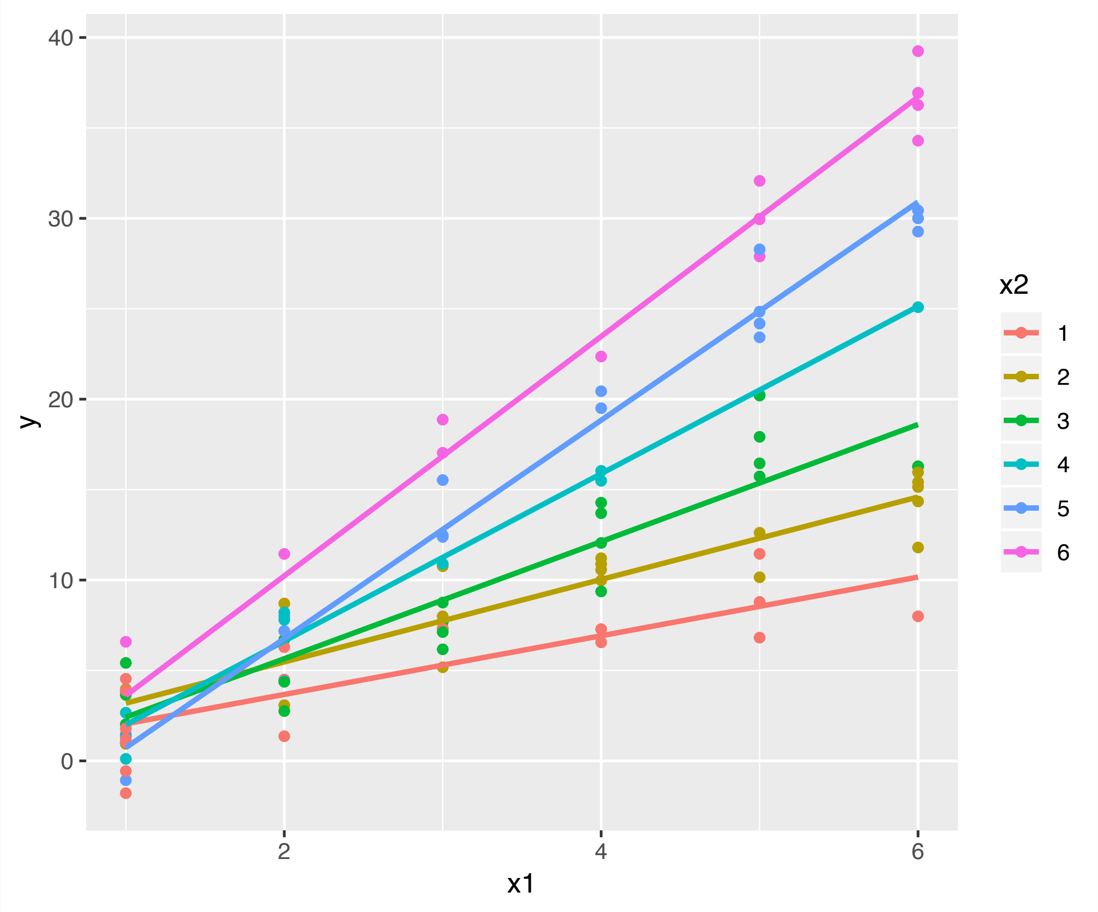
Also called "moderation"

- When called moderation, one of the variables is called the moderator



This is an example of an interaction between a continuous predictor and a dichotomous moderator

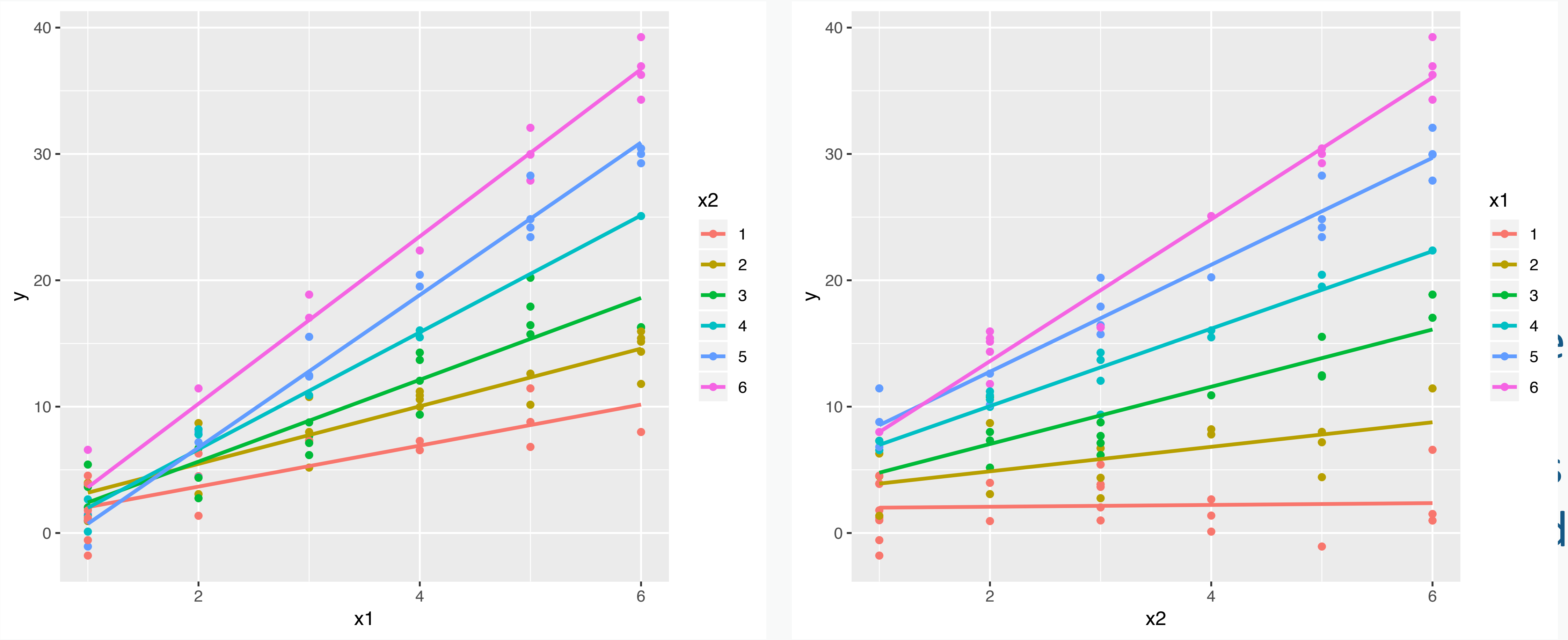
The *Symmetry* of Interaction



Either variable can be considered the "predictor" and the "moderator"

- Usually depends on the research question/story to be told
- The interpretation still means the same thing but can sound very different

The *Symmetry* of Interaction



Same interaction, different predictor/moderator distinction

Types of Interactions

Interactions can be between:

- Two numeric variables
- A dichotomous variable and a numeric variable
- Two dichotomous variables
- A multi-categorical variable and a numeric variable
- Two multi-categorical variables

Types of Interactions

Interactions can be b

- Two numeric variabl
- A dichotomous varia
- Two dichotomous va
- A multi-categorical v
- Two multi-categorica

Examples

- The effect of the treatment on the outcome may depend on the sex of the participant
- The effect of smoking on heart condition may depend on age
- The effect of education on income may depend on the economic climate

Could we reverse the way that we said each of these?

Yes

Interaction between a numeric and dichotomous variable

$$\hat{Y} = 2 + 0.5\,X_1 + 2\,D_1 + 1.5\,X_1 D_1$$

The intercept (the average Y value when all variables are zero)

The conditional effect of X_1 when D_1 equals zero

The conditional effect of D_1 when X_1 equals zero

quantifies how much the conditional effect of X_1 changes as D_1 changes by one unit

Cross-product (interaction)

Interaction between a numeric and dichotom[ous variable]

```r
hosp %>%
  lm(safety ~ exhaust * sex, data = .) %>%
  summary()
```

```
##
## Call:
## lm(formula = safety ~ exhaust * sex, data = .)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.23439 -0.61719  0.00054  0.74033  2.61092
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)      3.32847    0.27543  12.085  < 2e-16 ***
## exhaust          0.13674    0.07956   1.719  0.08673 .
## sexmale         -0.45969    0.40175  -1.144  0.25345
## exhaust:sexmale  0.31860    0.11456   2.781  0.00576 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.9816 on 296 degrees of freedom
## Multiple R-squared:  0.1812, Adjusted R-squared:  0.1729
## F-statistic: 21.83 on 3 and 296 DF,  p-value: 8.46e-13
```

The intercept (the average Y value when all variables are zero)

The conditional effect of X_1 when D_1 equals zero

Cross-product (interaction)

The conditional effect of D_1 when X_1 equals zero

quantifies how much the conditional effect of X_1 changes as D_1 changes by one unit

Interaction between two numeric variables

$$\hat{Y} = 2 + 0.5\,X_1 + 2\,X_2 + 1.5\,X_1 X_2$$

The intercept (the average Y value when all variables are zero)

The conditional effect of X_1 when X_2 equals zero

The conditional effect of X_2 when X_1 equals zero

quantifies how much the conditional effect of X_1 changes as X_2 changes by one unit

Cross-product (interaction)

Interaction between two dichotomous variables

$$\hat{Y} = 2 + 0.5\, D_1 + 2\, D_2 + 1.5\, D_1 D_2$$

The intercept (the average Y value when all variables are zero)

The conditional effect of D_1 when D_2 equals zero

The conditional effect of D_2 when D_1 equals zero

quantifies how much the conditional effect of D_1 changes as D_2 changes by one unit

Cross-product (interaction)

Interaction between numeric and multi-categorical variables

Dummy codes for the multi-categorical var

$$\hat{Y} = 2 + 0.5\, X_1 + 2\, D_{1.1} + .5\, D_{1.2} + 1.5\, X_1 D_{1.1} + 2\, X_1 D_{1.2}$$

The intercept (the average Y value when all variables are zero)

The conditional effect of X_1 when D_1 and D_2 equals zero

The conditional effect of D_1 and D_2 when X_1 equals zero

Cross-product (interaction)

quantifies how much the conditional effect of X_1 changes as D_1 and D_2 changes by one unit

General interpretation of coefficients

Doesn't matter if they are numeric or categorical

$$\hat{Y} = 2 + 0.5\,X_1 + 2\,X_2 + 1.5\,X_1X_2$$

General interpretation of coefficients

Doesn't matter if they are numeric or categorical

When X_2 is zero:

$$\hat{Y} = 2 + 0.5\, X_1 + \cancel{2\, X_2} + \cancel{1.5\, X_1 X_2}$$

Meaning when X_2 is zero, the effect of a one unit increase in X_1 is associated with a 0.5 unit increase in Y

General interpretation of coefficients

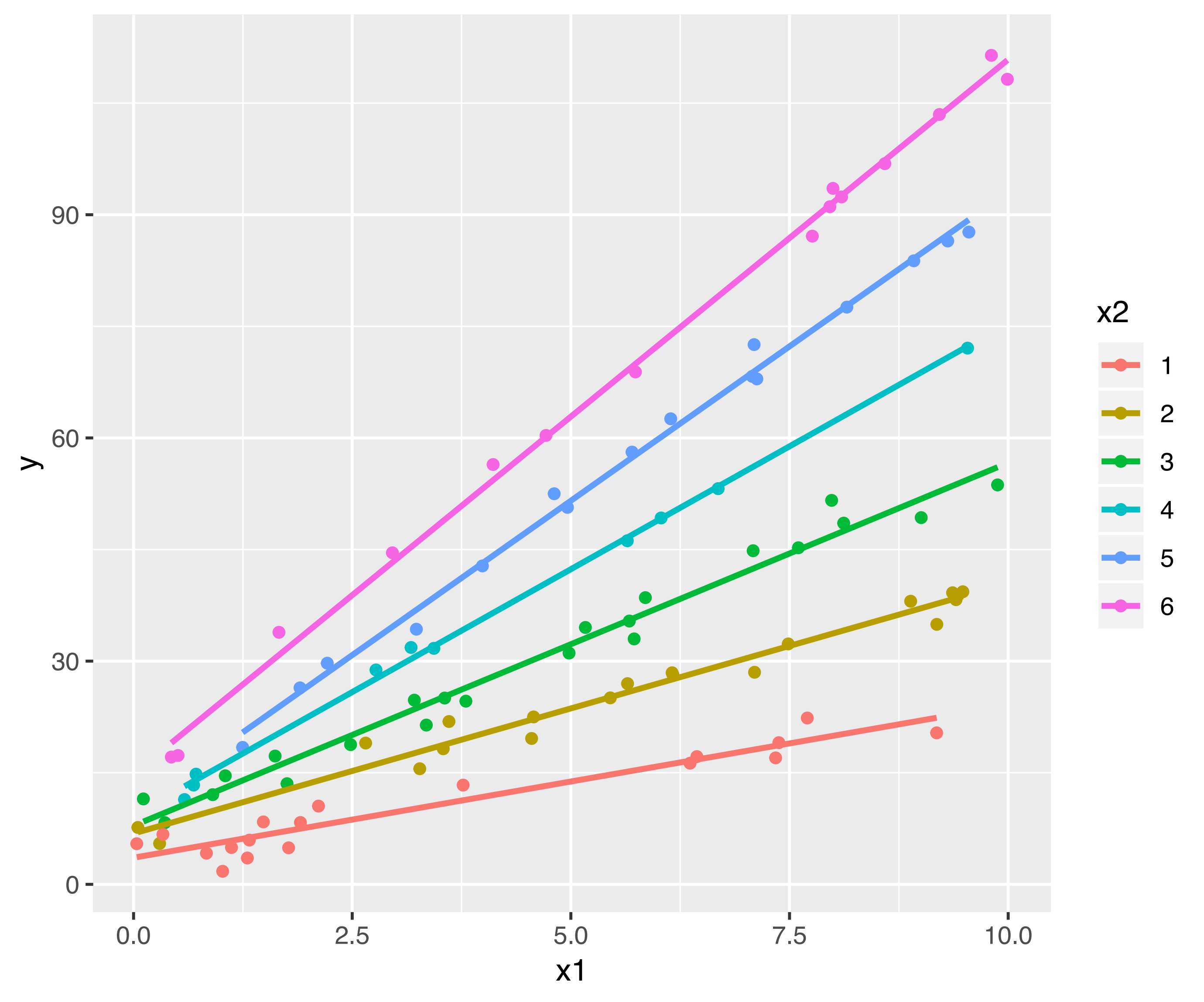
Doesn't matter if they are numeric or categorical

When X_1 is zero:

$$\hat{Y} = 2 + \cancel{0.5\ X_1} + 2\ X_2 + \cancel{1.5\ X_1 X_2}$$

Meaning when X_1 is zero, the effect of a one unit increase in X_2 is associated with a 2 unit increase in Y

General interpretation of coefficients



The difference in the slopes is 1.5 on average

$$+ 1.5\, X_1 X_2$$

All of these interpretations can be with additional covariates, "holding the covariates constant"

General interpretation of coefficients

Doesn't matter if they are numeric or categorical

The total effect of X_1 is:

$$\hat{Y} = \cancel{2 +} 0.5\, X_1 + \cancel{2\, X_2} + 1.5\, X_1 X_2$$

$$\hat{Y} = (0.5\ + 1.5\, X_2) X_1$$

$$\text{Effect of } X_1 = 0.5\ + 1.5\, X_2$$

General interpretation of coefficients

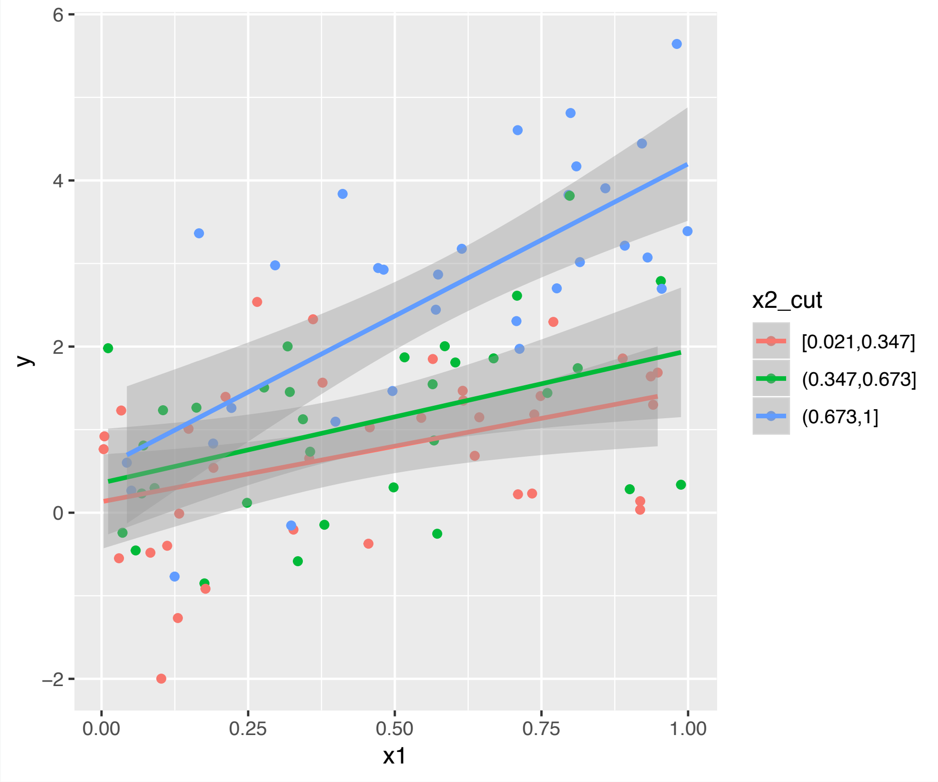
Doesn't matter if they are numeric or categorical

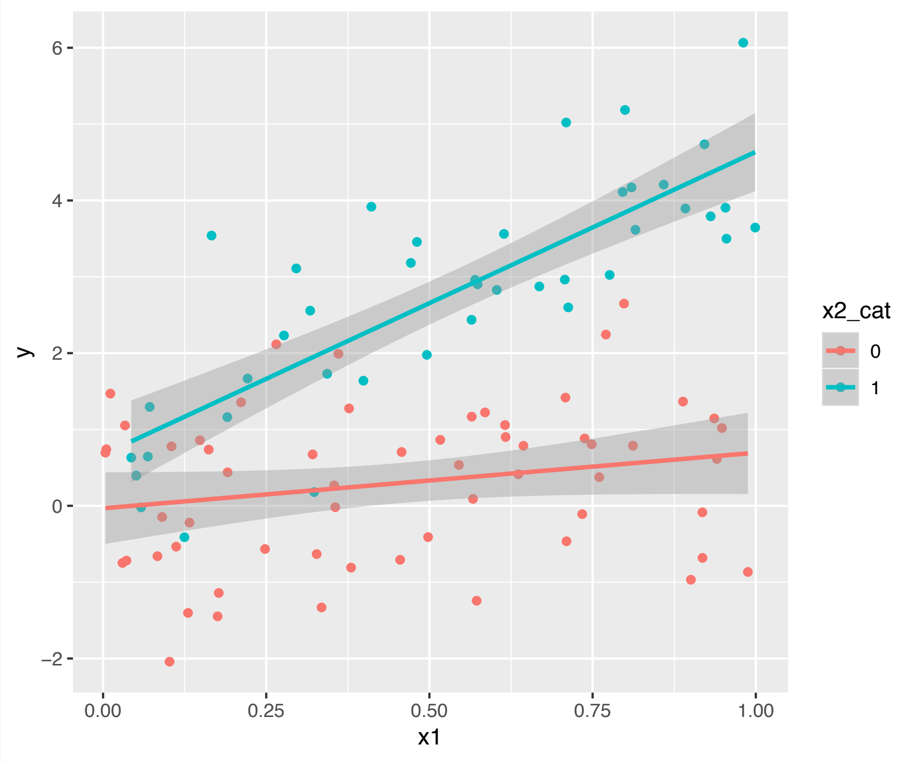
The same idea holds for the effect of X_2

$$\text{Effect of } X_2 = 2 + 1.5\, X_1$$

Recommendation: Always use a figure

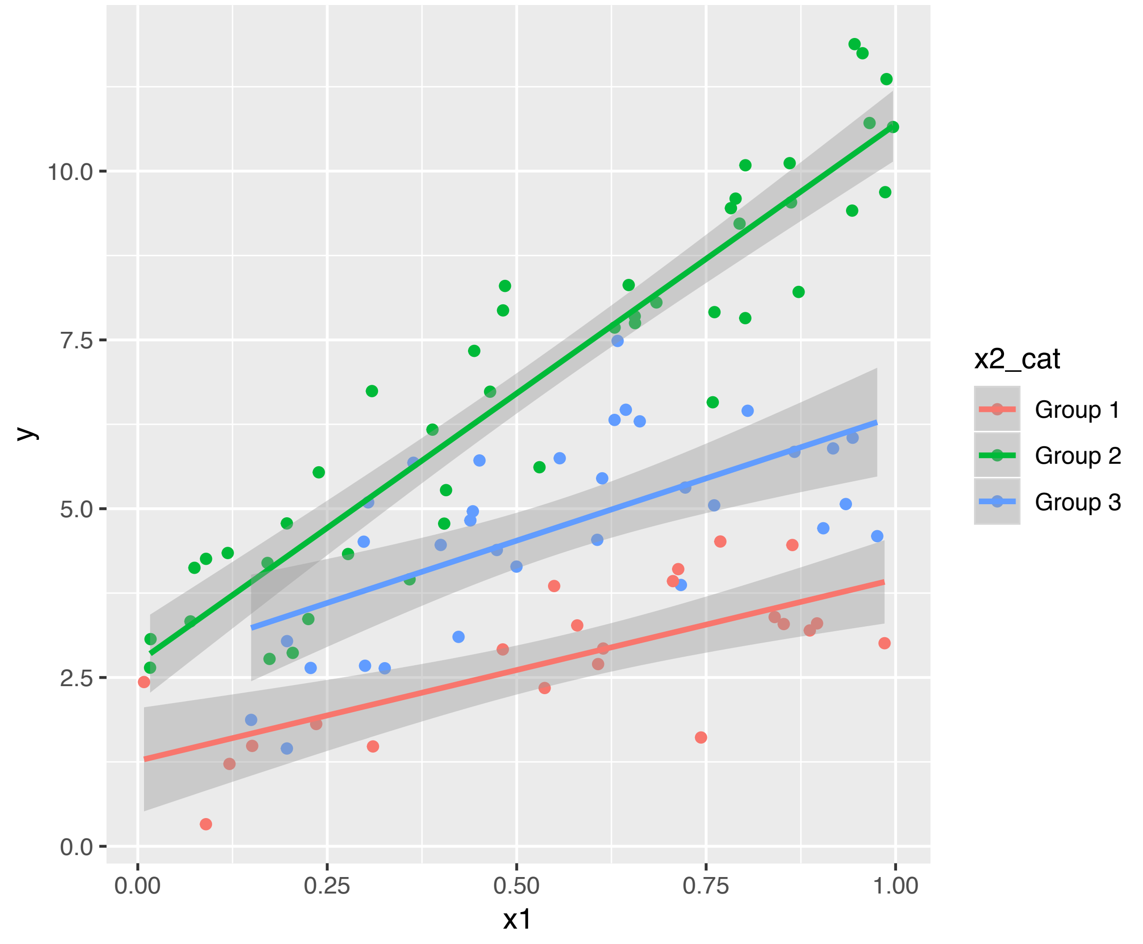
When we visualize an interaction with two numeric variable, we need to group one of them into some groups to better see it

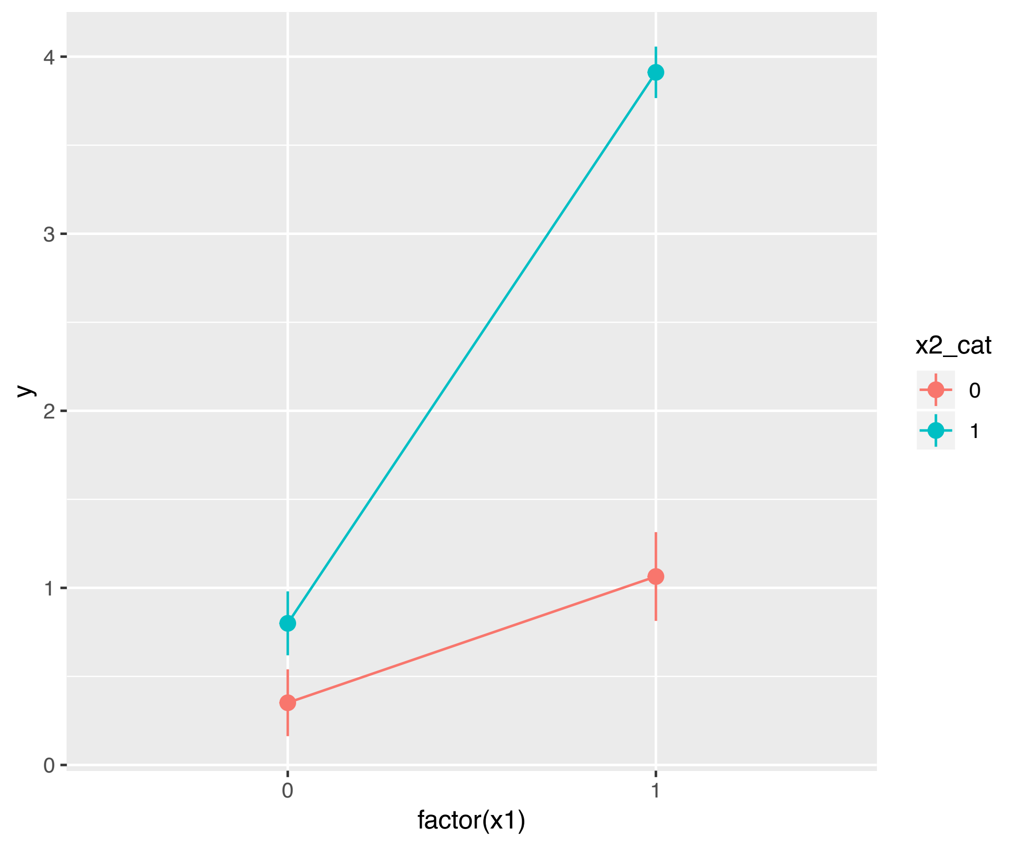




With a numeric variable and a dichotomous variable, we use the dichotomous as the grouping variable

When we visualize an interaction with a numeric variable and multi-categorical, we can use the multi-cat variable as the grouping variable





With two dichotomous variables, we can use either dummy variable as the grouping variable

More about *interpretation*

To make b_1 and b_2 more meaningful, we can shift the center of the predictor/moderator

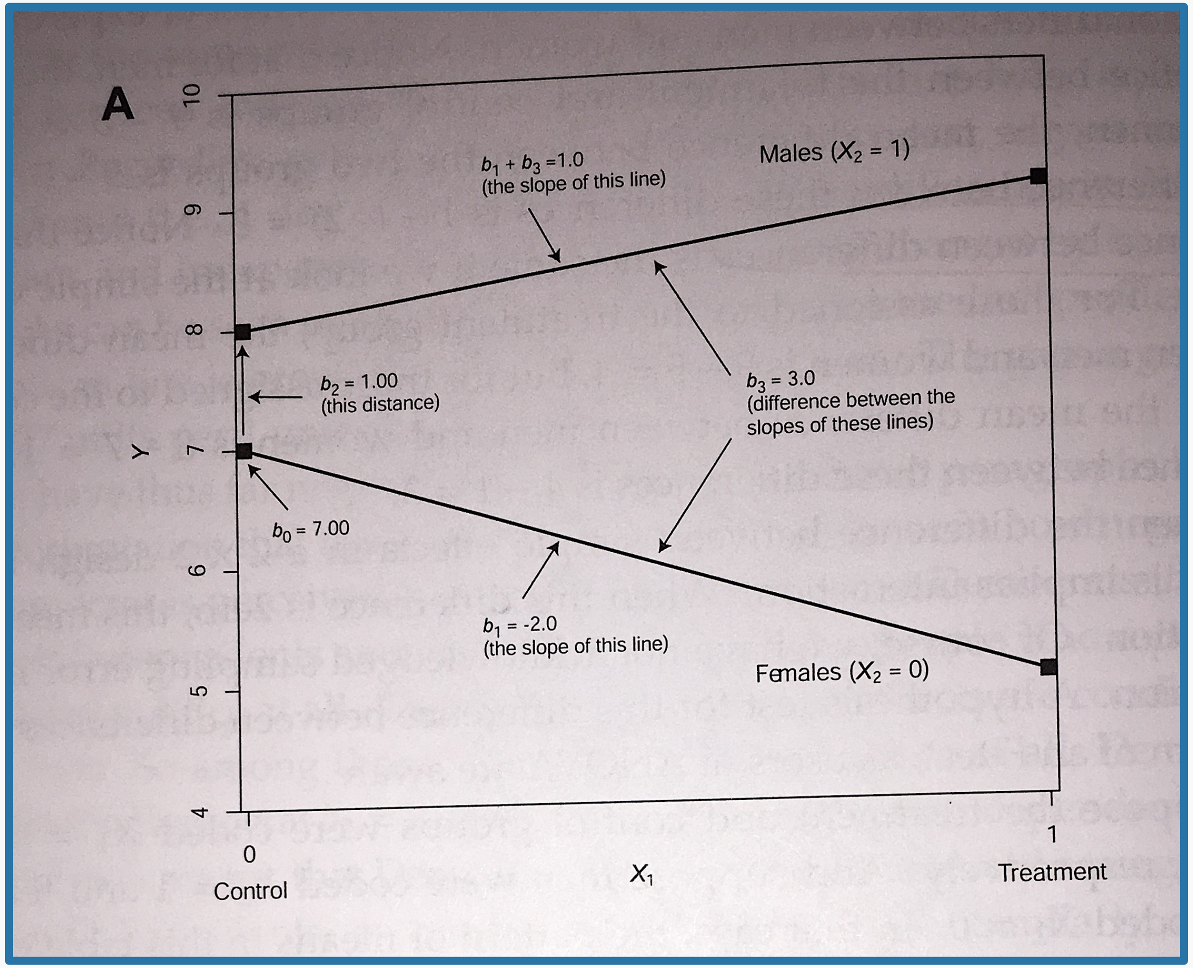
- E.g., we can mean center the variables which makes b_1 and b_2 the effect of the predictor/moderator when the other is at its mean
  - This is the *conditional effect* of the variable at the other variable's mean

The conditional effect means the effect conditional on the other variable being a certain value

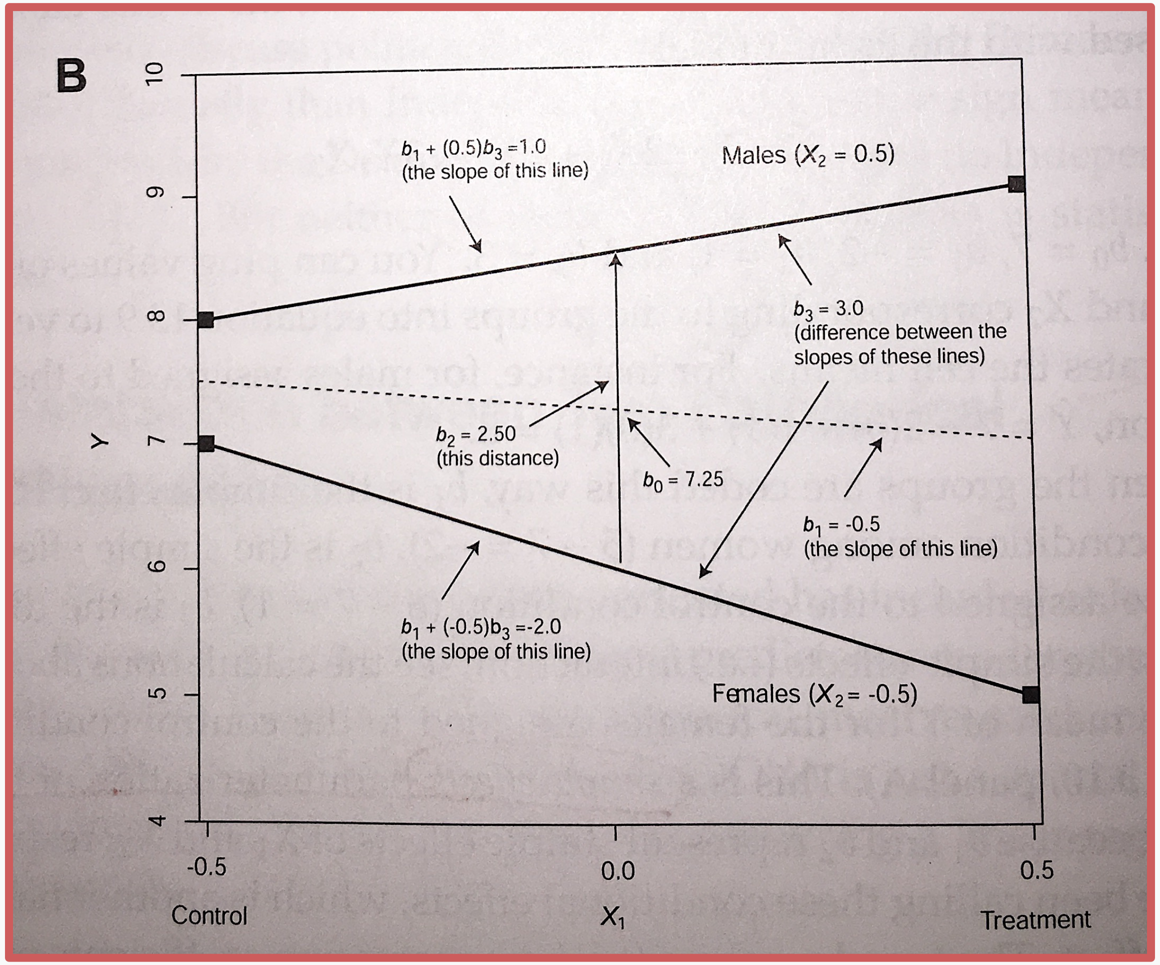
Difference between Regression and ANOVA

There are several similarities (and ultimately are the same) with some minor differences

Regression



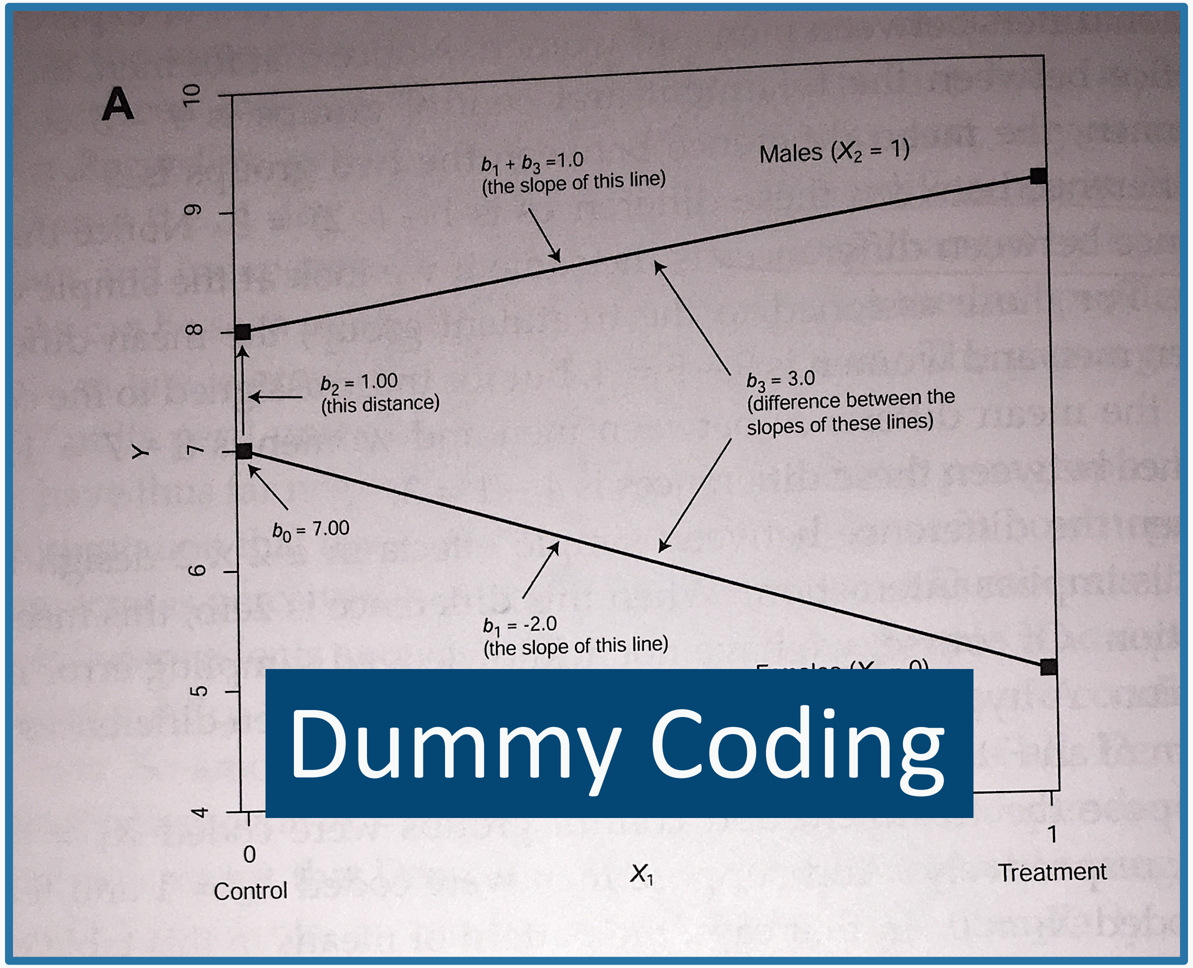
ANOVA



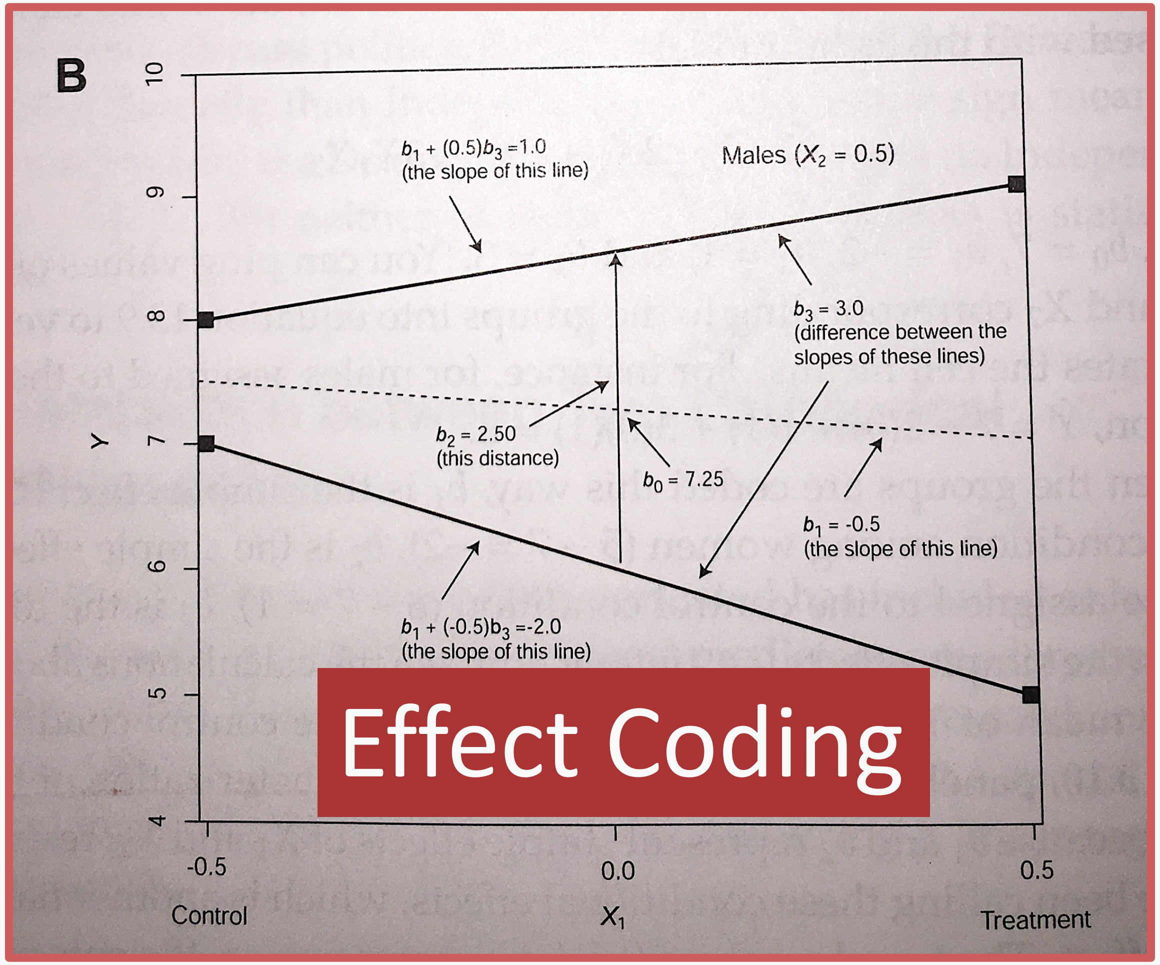
Difference between Regression and ANOVA

There are several similarities (and ultimately are the same) with some minor differences

Regression



ANOVA



Probing an Interaction

Lots of info, what to focus on?

Lots of information we can glean from a model with an interaction

- But much of it relies on p-values and can get excessive
- Almost always, a figure can highlight all the important factors more easily, simply, and clearly
    - A figure with some specific details is usually the best way to interpret it

More about *Conditional Effects*

Numeric

We can center a variable around any value and test if the conditional effect is different than zero

E.g., what is the effect of hearing loss on income when age is 20? How about 40? 60?

Dichotomous

We can switch the reference group by centering it with -1 (if coded 0 and 1)

E.g., what is the effect of age on heart disease when the person is a smoker? What about if they aren't?

Multi-Categorical

Like dichotomous, we can switch the reference group but subtracting the value of the group (for group = 2 use -2)

E.g., what is the effect of liberal views on voting behavior when the individual is black? Asian-American?

Johnson-Neyman Technique does this many times looking for areas of significant differences and non-significant ones (see page 427)

Issues with Detecting/Interpreting an Interaction

Interactions are difficult to accurately detect

- It is low powered
  - Easier to detect in experiments
- Is it curvilinear or interacting?
- Transformations of Y can greatly impact an interaction
- No “main effects” in model like in ANOVA
- Myths Get Busted (i.e., the truth is stated below):
  - Do not need to mean-center predictors in interaction
  - Don’t need to build model hierarchically
  - Don’t need to categorize variables in interaction

Organizing all the Interactions

There can be a lot of potential information from these models so how do we organize it all?

Recommendations

- Rely on theory for testing interactions
- If it is experimental, test any problematic interactions

Test all interactions in one compared to the null model	Variable-by-variable tests (test all interactions with a specific predictor together)	Test each interaction on its own (using the individual t-tests for each interaction)

Broad ⟵⟶ Narrow

*Failure to reject null does not mean there is no effect