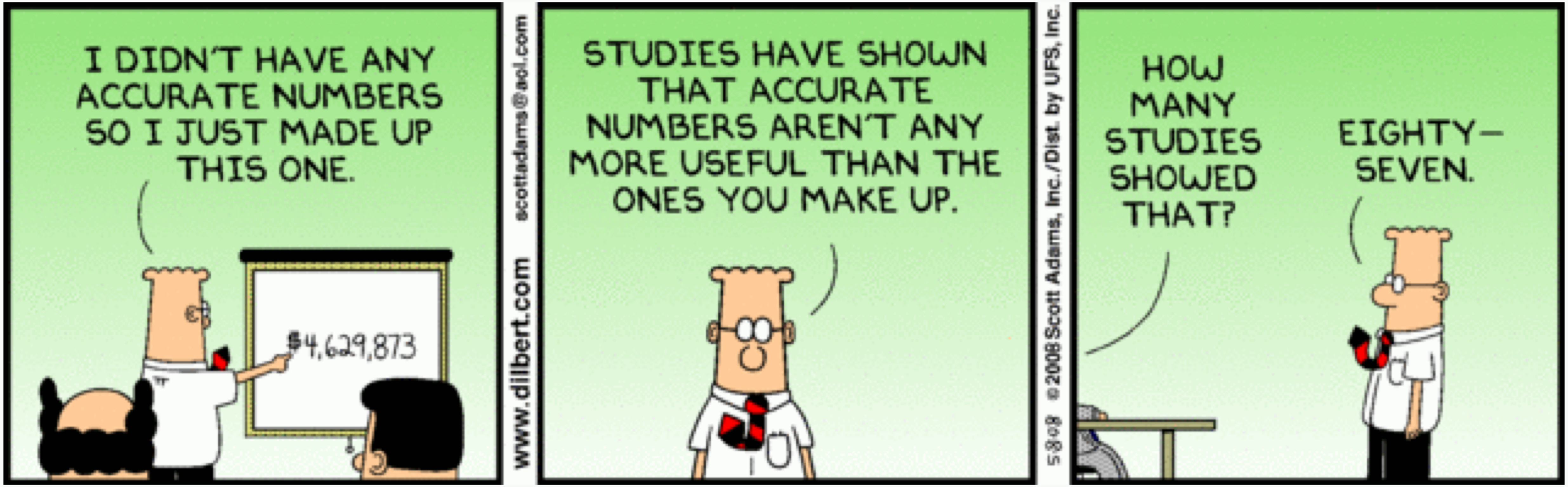
Statistics!


I DIDN'T HAVE ANY ACCURATE NUMBERS SO I JUST MADE UP THIS ONE.
$4,629,873
scottadams@aol.com
www.dilbert.com
STUDIES HAVE SHOWN THAT ACCURATE NUMBERS AREN'T ANY MORE USEFUL THAN THE ONES YOU MAKE UP.
©2008 Scott Adams, Inc./Dist. by UFS, Inc.
HOW MANY STUDIES SHOWED THAT?
EIGHTY-SEVEN.

EDUC 7610

Chapter 3

The Multiple Regression Model

$$Y_i = \beta_0 + \beta_1 X_{1i}\, \beta_2 X_{2i} + \epsilon_i$$

Fall 2018

Tyson S. Barrett, PhD

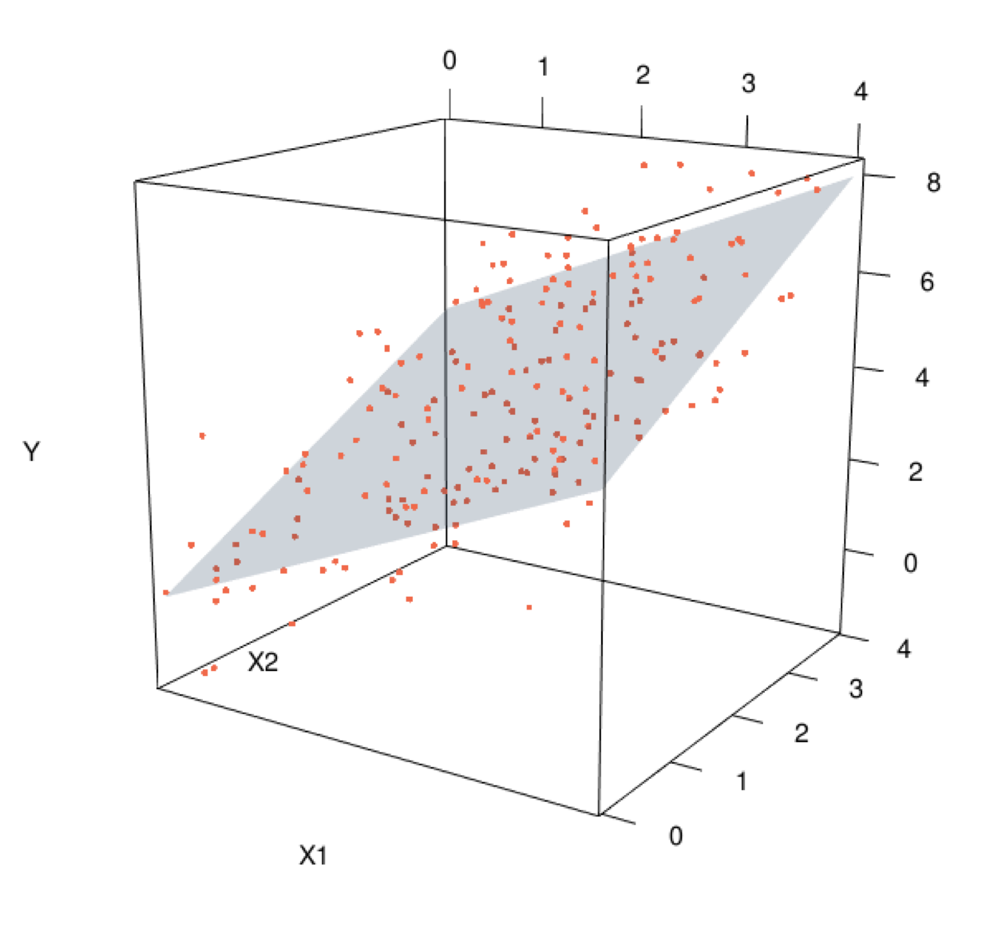
Why Multiple Regression?

2+ predictors in the same model

Allows us to "control for" the effects of other variables

- This can clarify weird results (e.g., Simpson's Paradox)
- Causal relationships without experiment

Can look at nonlinear relationships too (later in the class)

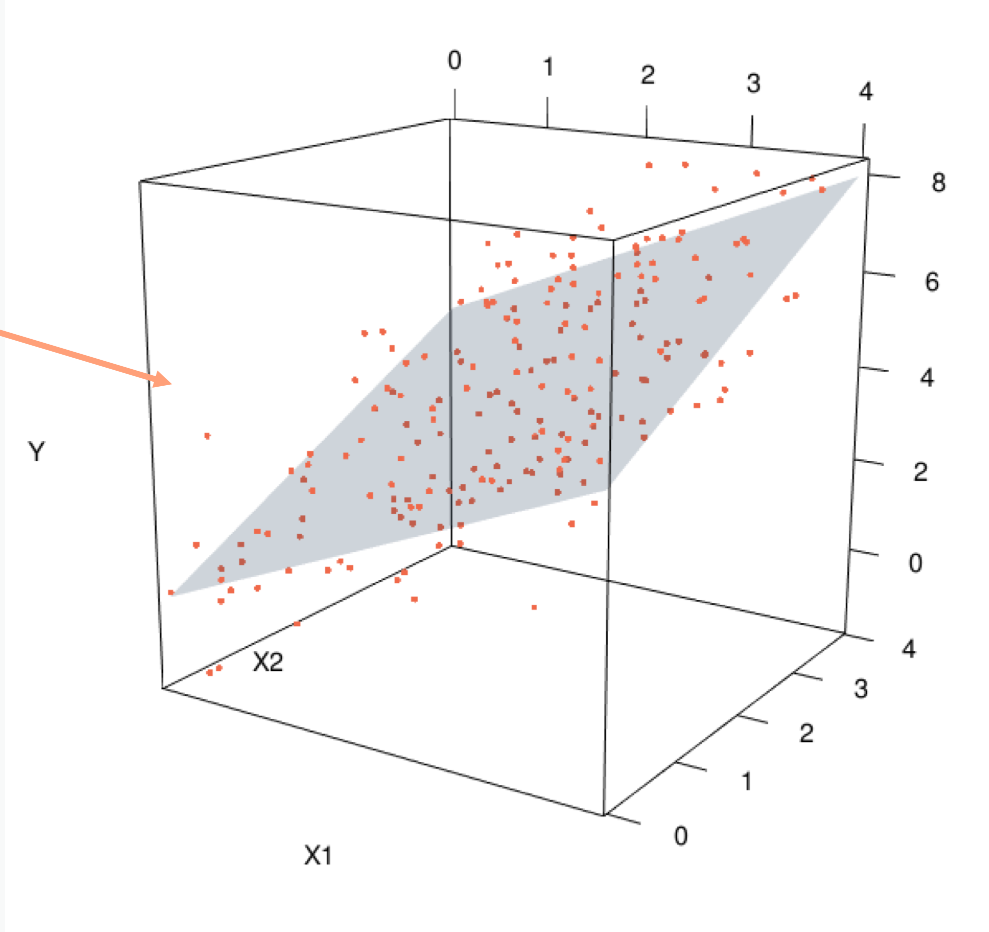


Multiple Regression

It no longer is looking for the best line but now is the best fitting plane (2 predictors) or hyperplane (3+ predictors)

- Much harder to visualize (hyperplane is essentially impossible to visualize)
- But the regression estimates are still very interpretable

The math behind the model is more complex



Multiple Regression

It no longer is looking for the best line but now is the best fitting plane (2 predictors) or hyperplane (3+ predictors)

- Much harder to visualize (hyperplane is essentially impossible to visualize)
- But the regression estimates are still very interpretable

The math behind the model is more complex

The tilted plane idea

Some vocabulary

Regressors, predictors, covariates, independent variables are all essentially synonyms

Beta Coefficients

- the estimates for each predictor, the associated change in the outcome when we increase the predictor by one unit holding all the other predictors (covariates) constant

Model

- A representation of Y as a linear function of the predictors

How do we get $\widehat{Y}_i$ in multiple regression?

Same as with simple regression, just with more +'s

$$\hat{Y}_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i}$$

How do we get $\widehat{Y}_i$ in multiple regression?

Same as with simple regression, just with more +'s

$$\hat{Y}_i = 3 + 2.5X_{1i} + 5X_{2i}$$

ID	X_1	X_2	$\hat{Y}$
1	2	0	?
2	5	4	?
3	3	2	?

Residuals

Residuals work the same way here as they did with simple regression (i.e., they are the difference between the predicted value and the observed value of Y)

Smaller errors generally means a better model

$$SS_{residual} = \sum_{i=1}^{N} (Y_i - \widehat{Y}_i)^2 = \sum_{i=1}^{N} e_i^2$$

OLS and Computation

OLS regression is a ”closed form” method

- Math can solve the minimization (using linear algebra)
- Other approaches (maximum likelihood) aren’t closed form and require a step-by-step (i.e., iterative) approach

So, if we wanted we could solve everything by hand :)

But we won’t

OLS and Computation - Example

```
gss %>%
  lm(income06 ~ educ + hompop,
     data = .)
```

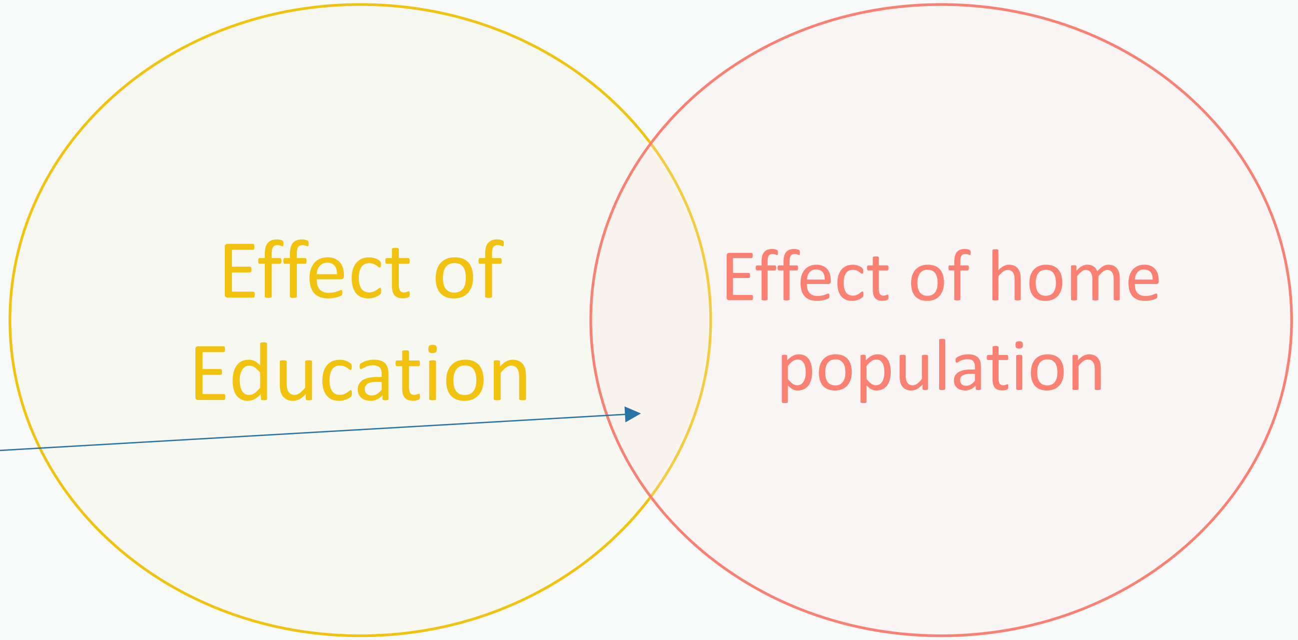
```
Coefficients:
(Intercept)         educ        hompop
     -18417         4286          7125
```

Partial **regression coefficients**

Partial Regression Coefficients

*When you see the word "**Partial**" – almost always refers to a relationship that is controlling for other factors*

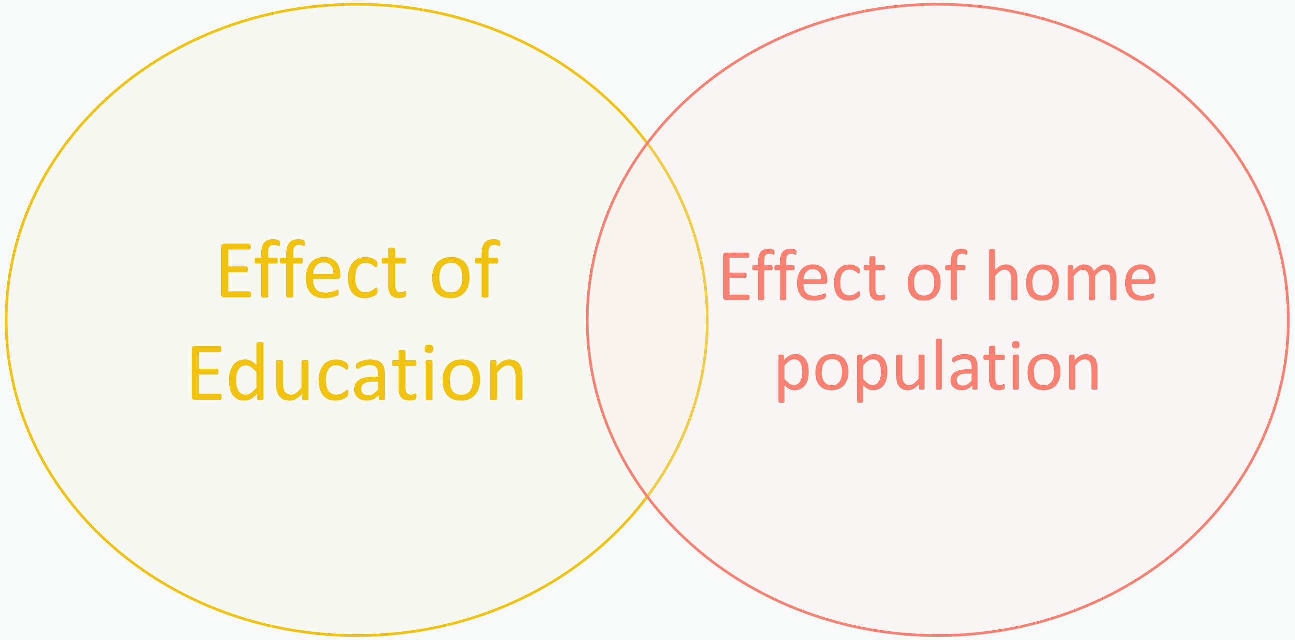
There is some amount of overlap between the effect of one and the other (when they are correlated)



Partial Regression Coefficients

*When you see the word "**Partial**" – almost always refers to a relationship that is controlling for other factors*

The partial effect of education is the non-overlapping parts of the total effect



Partial Regression Coefficients

*When you see the word "**Partial**" – almost always refers to a relationship that is controlling for other factors*

```
Coefficients:
(Intercept)          educ       hompop
     -18417          4286         7125
```

Interestingly, the partial effect can be bigger than the unadjusted effect (simple regression has the effect of education at 4127)

Partial Regression Coefficients

Two main ways of getting partial regression estimates:

1. *Use the residuals*
2. *Use matrix algebra (this is what R does behind the scenes)*

Residuals | Algebra

Partial Regression Coefficients

Two main ways of getting partial regression estimates:

1. *Use the residuals*
2. *Use matrix algebra (thi*

Residuals

Important!
What is a residual, again?

Partial Regression Coefficients

Two main ways of getting partial regression estimates:

1. *Use the residuals*
2. *Use matrix algebra (this is what R does behind the scenes)*

Residuals

1. Obtain the residuals of Y ~ covariates (let's call it Y_r)
2. Obtain the residuals of X ~ covariates (let's call it X_r)
3. Run the regression $Y_r \sim X_r$
4. This is the partial regression coefficient of X predicting Y when controlling for covariates

Algebra

$$\mathrm{B} = (X'X)^{-1}\,X'Y$$

where B is all of the partial regression estimates of the multiple regression model

Partial Correlation

We can also get a correlation while controlling for covariates, termed "Partial Correlation"

```
partial r = .361 (controlling for hompop)
```

How might we interpret this correlations?

- Consider what we just learned about partial coefficients

Partial Correlation

Main way of getting partial correlation estimates: Use the residuals

Residuals

1. Obtain the residuals of Y ~ covariates (let's call it Y_r)
2. Obtain the residuals of X ~ covariates (let's call it X_r)
3. Run the correlation of Y_r with X_r
4. This is the partial correlation of X and Y when controlling for covariates

Squared Partial Correlation

How did we interpret the regular partial correlations?

When we square them, we get the:

"proportion of the variance in Y explained by X and not explained by the covariates"

Or the unique amount of the variance that X accounts for in Y

*This will have a lot to do with when we talk about R and R^2 in a minute

Standardized Coefficients

We can also get standardized regression effects while controlling for covariates

```
Coefficients:
(Intercept)          educ        hompop
 -1.544e-16     3.540e-01     2.277e-01
```

$$b_{standardized} = b\,\frac{s_x}{s_y}$$

Standardized Coefficients

We can also get standardized regression effects while controlling for covariates

```
Coefficients:
(Intercept)              educ         hompop
 < -.000001              .354           .228
```

Two important considerations:

- What units would these be in?
- Are they similar to the partial correlations?

$$b_{standardized} = b\frac{s_x}{s_y}$$

R and R^2

Multiple Correlation

The correlation between the predicted values ($\hat{Y}$) and the observed values (Y)

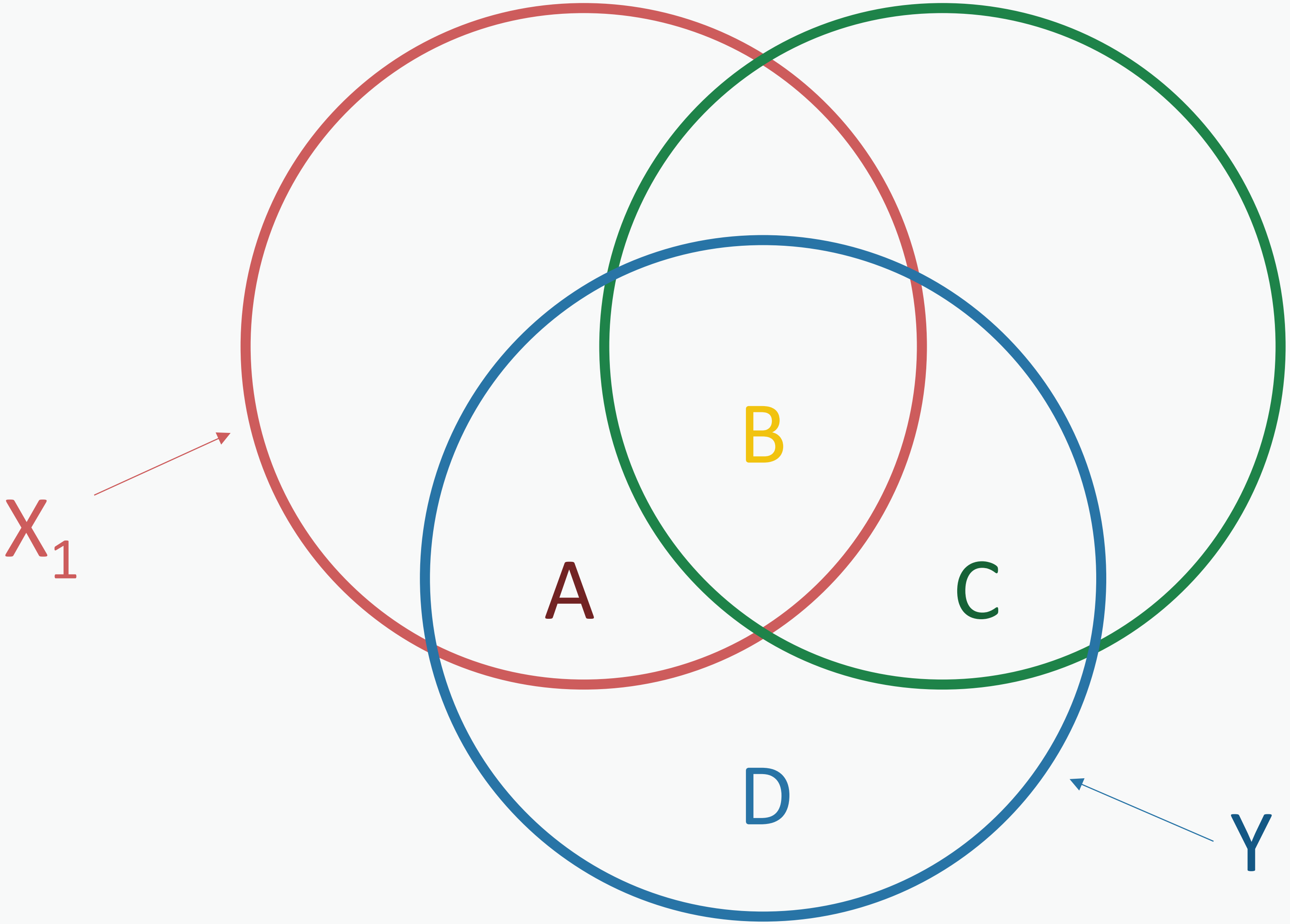
Why would this be interesting to know?

Proportion of Variance Accounted For

The proportion of the variance in Y that can be explained by the predictors

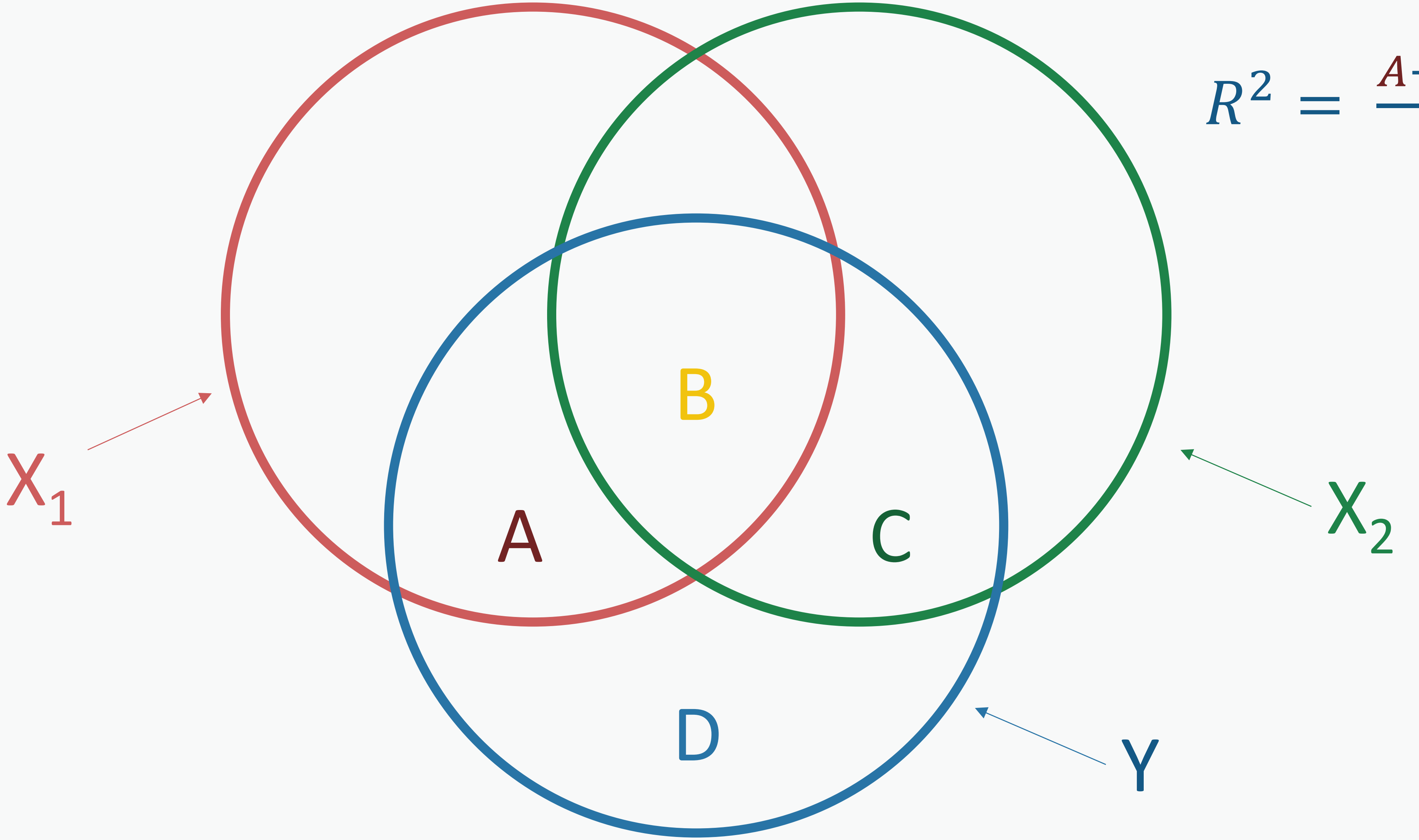
e.g., variance accounted for,
variance attributable to,
variance explained by

R^2 and Friends



Each circle represents the variables' variance

R² and Friends



$$R^2 = \frac{A+B+C}{Y} = \frac{A+B+C}{A+B+C+D}$$

$$pr_{x1}^2 = \frac{A}{A+D}$$

$$pr_{x2}^2 = \frac{C}{C+D}$$

Some important things

The simple and multiple regression coefficients can have different sizes and signs

Covariates: Can we predict the way that they'll affect a variable (e.g., b_1)?

It is based on the correlations between the covariate and X and the covariate and Y

	$Corr(x_1, x_2) > 0$	$Corr(x_1, x_2) < 0$
$b_1 > 0$	Positive bias	Negative bias
$b_1 < 0$	Negative bias	Positive bias

Some important things

The simple and multiple regression coefficients can have different sizes and signs

Covariates: Can we predict the way that they'll affect a variable (e.g., b_1)?

Next we will learn how to infer things from our model

Note: Do not memorize the formulas on page 83 – we'll get into the logic of it later