

▪ **ECOTS '26 · BEYOND SESSION**

Finding the Best College: *Misconceptions in Multiple Regression*

AUTHORS

Khanh Do · Joyce Gill · Shonda Kuiper
Grinnell College

THE PROBLEM

Multiple regression is foundational to statistics, yet there are common misunderstandings.

THE PROBLEM

Multiple regression is foundational to statistics, yet there are common misunderstandings.



Multicollinearity

Coefficient signs can flip
when correlated
predictors are added

THE PROBLEM

Multiple regression is foundational to statistics, yet there are common misunderstandings.



Multicollinearity

Coefficient signs can flip
when correlated
predictors are added



p-Hacking

Enough tests guarantee a
'significant' result

THE PROBLEM

Multiple regression is foundational to statistics, yet there are common misunderstandings.



Multicollinearity

Coefficient signs can flip
when correlated
predictors are added



p-Hacking

Enough tests guarantee a
'significant' result



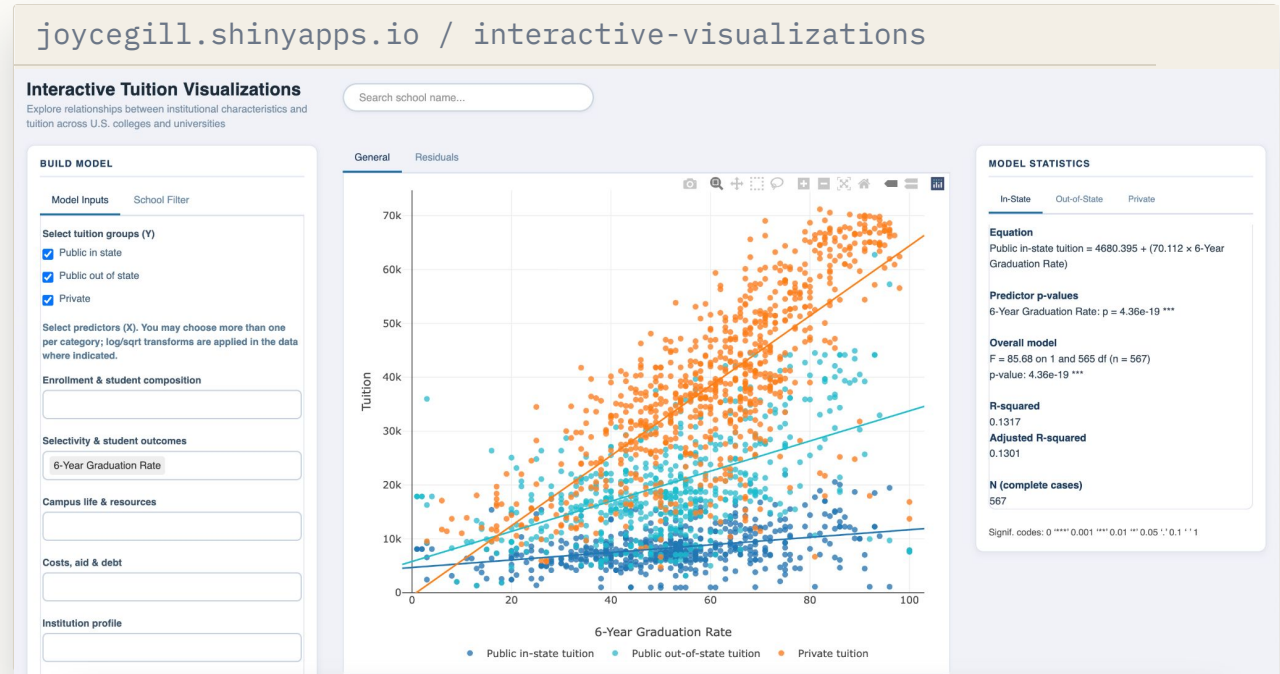
Stat vs Practical Significance

A reliable nonzero
coefficient can still mean
nothing real

OUR INTERVENTION

An interactive activity designed to visualize and conceptually explain common misconceptions

Built on **IPEDS**
(Integrated
Postsecondary
Education Data
System) data by the
NCES (National
Center for Education
Statistics)



Exploring the Cost of College Tuition: *a 16-question worksheet focused on misconceptions within multiple regression*

PART 1

Calculating a Simple Regression Model

- Fit a simple model
- Interpret the results

PART 2

Calculating a Multiple Regression Model

- Build a model one predictor at a time
- Watch coefficients shift

PART 3

Can You Trust p-values?

- Unreliable coefficients and p-values
- Solve questions to understand what happened and why

PART 4

Guided Discussion

- Multicollinearity
- Multiple comparisons
- Limits of variable selection

• LIVE

Live *demo.*

TRY IT NOW

`joycegill.shinyapps.io/interactive-visualizations`

Same data. Opposite conclusion.

MODEL A

Tuition ~ 6-year graduation rate

+ Positive coefficient

Higher graduation rates predict **higher** tuition.

p < 0.05

Question 3

vs

MODEL B

Tuition ~ 6-year grad rate + grad rate (men) + application fee + log(grants)

- Negative coefficient

Higher graduation rates predict **lower** tuition.

p < 0.05

Question 7

The survey shows that students didn't just enjoy the activity; they left questioning their results




■ Agree ■ Unsure ■ Disagree

Most students' misconceptions were clarified


MISCONCEPTION

% WHO **CORRECTLY** SELECTED "DISAGREE"


"A p-value < 0.05 means we can be confident the test was properly conducted" 88%



"If results match our expectations, we can assume the data was collected properly" 88%




"If the math is done correctly, we can trust the statistical results" 69%




GOOD PRACTICE

% WHO **CORRECTLY** SELECTED "AGREE"

"Before deciding with data, carefully evaluate how it was collected" 100%



"Statistics requires critical thinking, not just calculation" 100%



→ **31% are still unsure** that correct math can hide an untrustworthy result

What students valued and what they'd change

“It is a strong, visible example of how p-values can be misleading.”

“A helpful way to visually understand multicollinearity and p-hacking without just reading about them.”

“Seeing how easily a p-value can be tricked when you start piling variables into a model.”

WHAT THEY'D IMPROVE

Teach it earlier. Move the lab into the regression unit itself, rather than after as extra credit.

Ease into the variables. A short orientation to the predictor groups before diving in.

Thank you for listening!

Questions & feedbacks are welcomed.

APP

joycegill.shinyapps.io/interactive-visualizations

ACTIVITY

https://dataspace.sites.grinnell.edu/college_tuition.html

CODE

<https://github.com/joycegill/Advanced-Statistical-Modeling>

CONTACT

khanhphuongdo28@gmail.com · joyceygill@gmail.com