

**GEOG 400 Project 4**

Name:

**Dependent variable**

Which is your Y variable?

**Hypotheses**

X<sub>1</sub> (expstd) on Y:

X<sub>2</sub> (stdfac) on Y:

X<sub>3</sub> (teachpay) on Y:

X<sub>4</sub> (takesat) on Y:

X<sub>1</sub> (expstd) on Y<sub>2</sub> (takesat):

X<sub>2</sub> (stdfac) on Y<sub>2</sub> (takesat):

X<sub>3</sub> (teachpay) on Y<sub>2</sub> (takesat):

**Seven simple linear regressions**

*Four models using totalsat or mathsat or verbalsat as Y*

model

| X <sub>i</sub>            | R                    | R <sup>2</sup> <sub>adj</sub> | a                        | b <sub>i</sub>                        | t                    | sig t                | is t sig?            | direct or inverse?   | expected direction?  |
|---------------------------|----------------------|-------------------------------|--------------------------|---------------------------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| X <sub>1</sub> (expstd)   | <input type="text"/> | <input type="text"/>          | Y = <input type="text"/> | + <input type="text"/> X <sub>1</sub> | <input type="text"/> | <input type="text"/> | <input type="text"/> | <input type="text"/> | <input type="text"/> |
| X <sub>2</sub> (stdfac)   | <input type="text"/> | <input type="text"/>          | Y = <input type="text"/> | + <input type="text"/> X <sub>2</sub> | <input type="text"/> | <input type="text"/> | <input type="text"/> | <input type="text"/> | <input type="text"/> |
| X <sub>3</sub> (teachpay) | <input type="text"/> | <input type="text"/>          | Y = <input type="text"/> | + <input type="text"/> X <sub>3</sub> | <input type="text"/> | <input type="text"/> | <input type="text"/> | <input type="text"/> | <input type="text"/> |
| X <sub>4</sub> (takesat)  | <input type="text"/> | <input type="text"/>          | Y = <input type="text"/> | + <input type="text"/> X <sub>4</sub> | <input type="text"/> | <input type="text"/> | <input type="text"/> | <input type="text"/> | <input type="text"/> |

Three models using takesat as Y (instead of X<sub>4</sub>)

| X <sub>i</sub>            | R                    | R <sup>2</sup> <sub>adj</sub> | a                        | b <sub>i</sub>         | t              | sig t                | is t sig?            | direct or inverse?   | expected direction?  |
|---------------------------|----------------------|-------------------------------|--------------------------|------------------------|----------------|----------------------|----------------------|----------------------|----------------------|
| X <sub>1</sub> (expstd)   | <input type="text"/> | <input type="text"/>          | Y = <input type="text"/> | + <input type="text"/> | X <sub>1</sub> | <input type="text"/> | <input type="text"/> | <input type="text"/> | <input type="text"/> |
| X <sub>2</sub> (stdfac)   | <input type="text"/> | <input type="text"/>          | Y = <input type="text"/> | + <input type="text"/> | X <sub>2</sub> | <input type="text"/> | <input type="text"/> | <input type="text"/> | <input type="text"/> |
| X <sub>3</sub> (teachpay) | <input type="text"/> | <input type="text"/>          | Y = <input type="text"/> | + <input type="text"/> | X <sub>3</sub> | <input type="text"/> | <input type="text"/> | <input type="text"/> | <input type="text"/> |

Kitchen sink multiple regression using totalsat or mathsat or verbalsat as Y

| R                    | R <sup>2</sup> <sub>adj</sub> | model  | F                    | sig F                |
|----------------------|-------------------------------|--|----------------------|----------------------|
| <input type="text"/> | <input type="text"/>          | Y = <input type="text"/> + <input type="text"/> X <sub>1</sub> + <input type="text"/> X <sub>2</sub> + <input type="text"/> X <sub>3</sub> + <input type="text"/> X <sub>4</sub> | <input type="text"/> | <input type="text"/> |

Is this model significant?  yes  no

Individual X<sub>i</sub> performances

|                           | t                    | sig t                | Is this X <sub>i</sub> contributing significantly? | Do you expect SPSS to retain this X <sub>i</sub> during backwards elimination? |
|---------------------------|----------------------|----------------------|--|--|
| X <sub>1</sub> (expstd)   | <input type="text"/> | <input type="text"/> | <input type="text"/>                               | <input type="text"/>   |
| X <sub>2</sub> (stdfac)   | <input type="text"/> | <input type="text"/> | <input type="text"/>                               | <input type="text"/>   |
| X <sub>3</sub> (teachpay) | <input type="text"/> | <input type="text"/> | <input type="text"/>                               | <input type="text"/>   |
| X <sub>4</sub> (takesat)  | <input type="text"/> | <input type="text"/> | <input type="text"/>                               | <input type="text"/>   |

Moving toward a more elegant model through backwards elimination

How many models did SPSS build using the Backwards method?

Fill out this model for that model with the fewest variables (I'm leaving space for all four variables so not to bias you).

To keep from going insane later, try to use the same X variable number for each surviving X variable as you used above (SPSS will renumber them, so it can get confusing)

(X<sub>1</sub> = expstd, X<sub>2</sub> = stdfac, X<sub>3</sub> = teachpay, X<sub>4</sub> = takesat)

| R                    | R <sup>2</sup> <sub>adj</sub> | model  | F                    | sig F                |
|----------------------|-------------------------------|--|----------------------|----------------------|
| <input type="text"/> | <input type="text"/>          | Y = <input type="text"/> + <input type="text"/> X <sub>1</sub> + <input type="text"/> X <sub>2</sub> + <input type="text"/> X <sub>3</sub> + <input type="text"/> X <sub>4</sub> | <input type="text"/> | <input type="text"/> |

Name of X<sub>i</sub> IF you use it:



**Kitchen sink using *takesat for the Y variable***

<sup>R</sup>  <sup>R<sup>2</sup><sub>adj</sub></sup> Y =  +  <sup>model</sup> X<sub>1</sub> +  X<sub>2</sub> +  X<sub>3</sub>
 <sup>F</sup>  <sup>sig F</sup>

Name of X<sub>i</sub> IF you use it:

Is this model significant?  <sup>yes</sup>  <sup>no</sup>

Individual X<sub>i</sub> performances

|                           | t                    | sig t                | Is this X <sub>i</sub> contributing significantly? | Do you expect SPSS to retain this X <sub>i</sub> during backwards elimination? |
|---------------------------|----------------------|----------------------|--|--|
| X <sub>1</sub> (expstd)   | <input type="text"/> | <input type="text"/> | <input type="text"/>                               | <input type="text"/>   |
| X <sub>2</sub> (stdfac)   | <input type="text"/> | <input type="text"/> | <input type="text"/>                               | <input type="text"/>   |
| X <sub>3</sub> (teachpay) | <input type="text"/> | <input type="text"/> | <input type="text"/>                               | <input type="text"/>   |

**Backwards elimination and kids taking the SATs**

How many models did SPSS build using the Backwards method?

How many X variables did SPSS leave in the final model?

Fill out the model with the fewest variables:

<sup>R</sup>  <sup>R<sup>2</sup><sub>adj</sub></sup> Y =  +  <sup>model</sup> X<sub>1</sub> +  X<sub>2</sub> +  X<sub>3</sub> +  X<sub>4</sub>
 <sup>F</sup>  <sup>sig F</sup>

Name of X<sub>i</sub> IF you use it:

Fate of the variables in your final Backwards model

|                           | t                    | sig t                | excluded?            | included?            |
|---------------------------|----------------------|----------------------|----------------------|----------------------|
| X <sub>1</sub> (expstd)   | <input type="text"/> | <input type="text"/> | <input type="text"/> | <input type="text"/> |
| X <sub>2</sub> (stdfac)   | <input type="text"/> | <input type="text"/> | <input type="text"/> | <input type="text"/> |
| X <sub>3</sub> (teachpay) | <input type="text"/> | <input type="text"/> | <input type="text"/> | <input type="text"/> |

### Forward modelling and kids taking SATs

How many models did SPSS build using the Forward method?

How many X variables did SPSS put into the final model?

Fill out the model with the fewest variables (entering only those X variables it kept):

|   |                               |     |       |   |   |                  |  |                  |  |                  |  |                |  |  |  |
|---|-------------------------------|-----|-------|---|---|------------------|--|------------------|--|------------------|--|----------------|--|--|--|
| R | R <sup>2</sup> <sub>adj</sub> |     | model |   | F | sig F            |  |                  |  |                  |  |                |  |  |  |
|   |                               | Y = |       | + |   | X <sub>1</sub> + |  | X <sub>2</sub> + |  | X <sub>3</sub> + |  | X <sub>4</sub> |  |  |  |

Name of X<sub>i</sub> IF you use it:

|  |  |  |  |
|--|--|--|--|
|  |  |  |  |
|--|--|--|--|

Fate of the variables in your final Forward model

|                           | t | sig t | excluded? | included? |
|---------------------------|---|-------|-----------|-----------|
| X <sub>1</sub> (expstd)   |   |       |           |           |
| X <sub>2</sub> (stdfac)   |   |       |           |           |
| X <sub>3</sub> (teachpay) |   |       |           |           |

### Interpretation of student performance on the SATs

Why is there a significant negative relationship between the percentage of students taking the SAT and the average performance on the SAT?

In light of this, why is there a significant negative association between expenditures per student and SAT performance?

Again in light of this, why is there a significant negative association between teacher pay and student performance on the SATs?

Why is there no significant association between student:faculty ratio and anything else?

Of the three multiple regression models you built to explain student performance on the SATs, which one makes the most sense to you theoretically? Be sure to consider everything in the list of hints provided in the lab.

### Interpretation of students taking the SATs

What is the association between:

|   | R                    | sig t                |
|---|----------------------|----------------------|
| expenditures per student and percentage of students taking the SAT? | <input type="text"/> | <input type="text"/> |
| student to faculty ratio and percentage of students taking the SAT? | <input type="text"/> | <input type="text"/> |
| teacher pay and percentage of students taking the SAT?              | <input type="text"/> | <input type="text"/> |

Why does expenditure per student drop out in the multiple regression models?

Which of the three multiple regression models you built to explain the percentage of students taking the SATs makes the most sense to you theoretically? Again, be sure to consider everything on the list of hints (e.g., diagnostics, simplicity)

### Science meets policy

If you were a conservative Republican wanting to reduce government spending, what would you pick out of this analysis and turn into a helpful slogan?

If you were a teacher's union representative, what would you pick out of this study that might make some great placards for the picket line?

If you were the parent of a kid in school, what would you want public policy to consider? What's in the best interest of the kids?