

# Lecture 19: Descriptive Statistics

Wednesday, November 1, 2023

Your Teaching Fellows:

003/004:	Zahra Abolghasem	Bronwen Grocott
	Vasileia Karasavva	Ni An
010:	Thalia Lang	Malina Lemmons
	Ruoning Li	Irene Wen

Lectures: MWF 12:00 PM – 1:00 PM (003); 1:00 PM – 2:00 PM (004); 2:00 PM – 3:00 PM (010)

Office hours: Tuesdays 2:00 PM – 4:00 PM

## For the exam

- Bring pencil, pen, eraser, a simple or scientific calculator (no programmable calculator)

## Recall...

- Goals of science:
  - To describe behaviour
    - Descriptive statistics
    - Correlation
  - To predict behaviour
    - Regression
  - To determine the cause of behaviour
  - To understand or explain behaviour

## Cultural adaptation

- What do we need in order to predict someone's score on an outcome variable?
  - Relationship between two variables
  - Some “point of origin”

# Regression

- Extension of correlation:
  - Both measure relationships among variables
  - Neither implies causation
- Important characteristics:
  - Key terms, not just “variable 1” and “variable 2”
    - Use score on one variable (“Predictor”) to predict changes in another variable (“Criterion”)
  - Regression models = a set of theoretically relevant predictors predicting a criterion variable
    - Can look at how 1 or more predictors can uniquely predict variability in criterion, amongst a set of predictors

## Cultural adaptation

- Correlation question:
  - Is there a relationship between



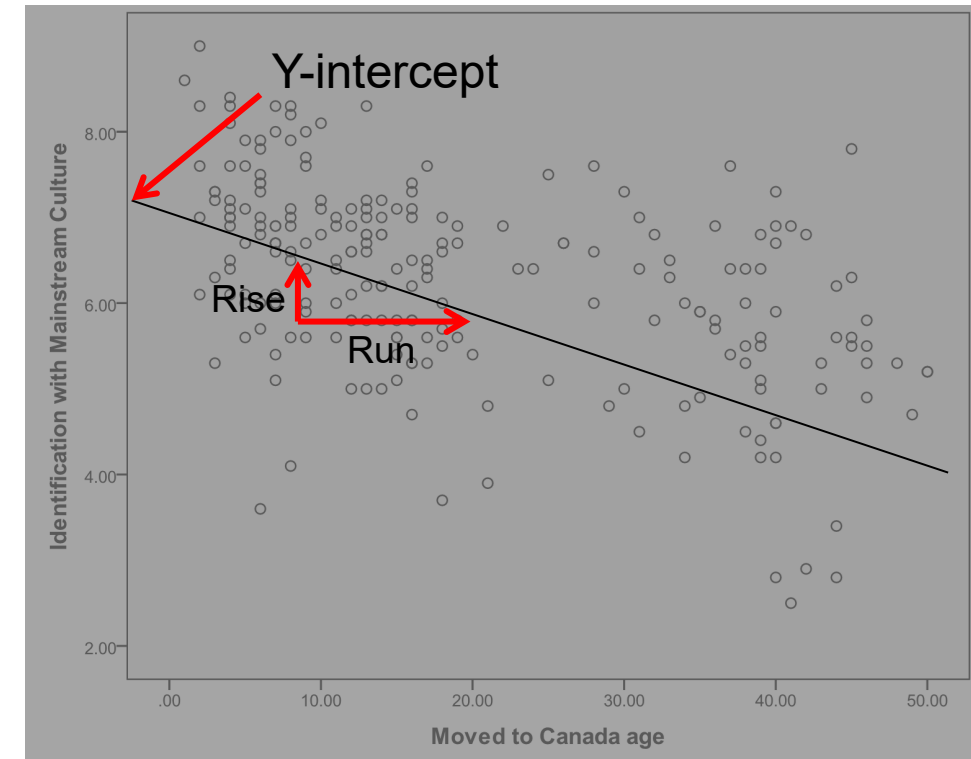
$$r = -.49$$

## Cultural adaptation

- What do we need in order to predict someone's score on an outcome variable?
  - Relationship between two variables
  - Some “point of origin”

# Cultural adaptation

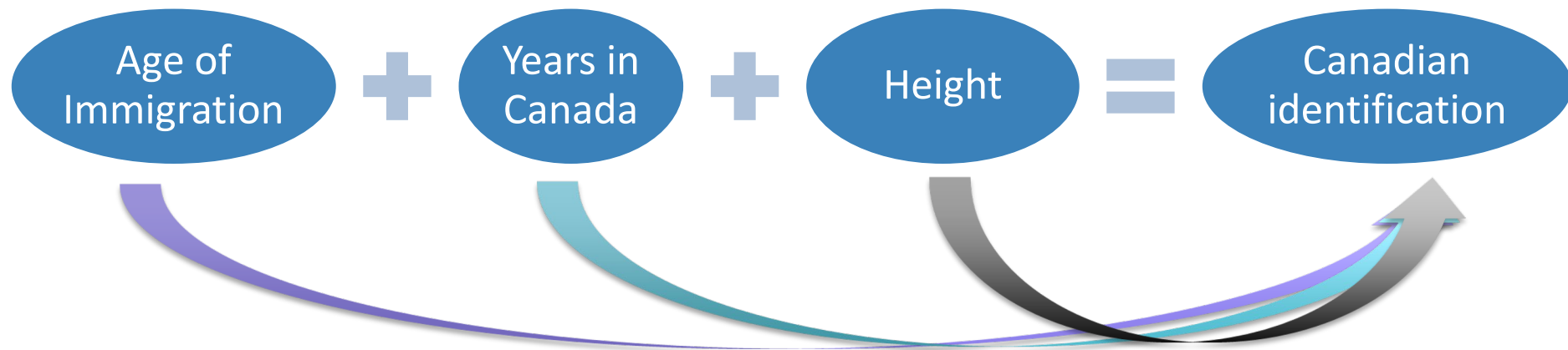
- Regression question:
  - Can we use one's age of immigration to predict current identification with Canadian culture?
  - Regression line:
    - $\hat{Y} = a + bX$ 
      - Relationship
    - $b = \text{slope}$ 
      - Rise over run
    - $a = \text{y-intercept}$ 
      - “Point of origin”
    - $\hat{Y} = 7.03 - .04X$





## Cultural adaptation

- Most important benefit of regression:
  - Can investigate role of *multiple predictors* in independently predicting the criterion



## Cultural adaptation

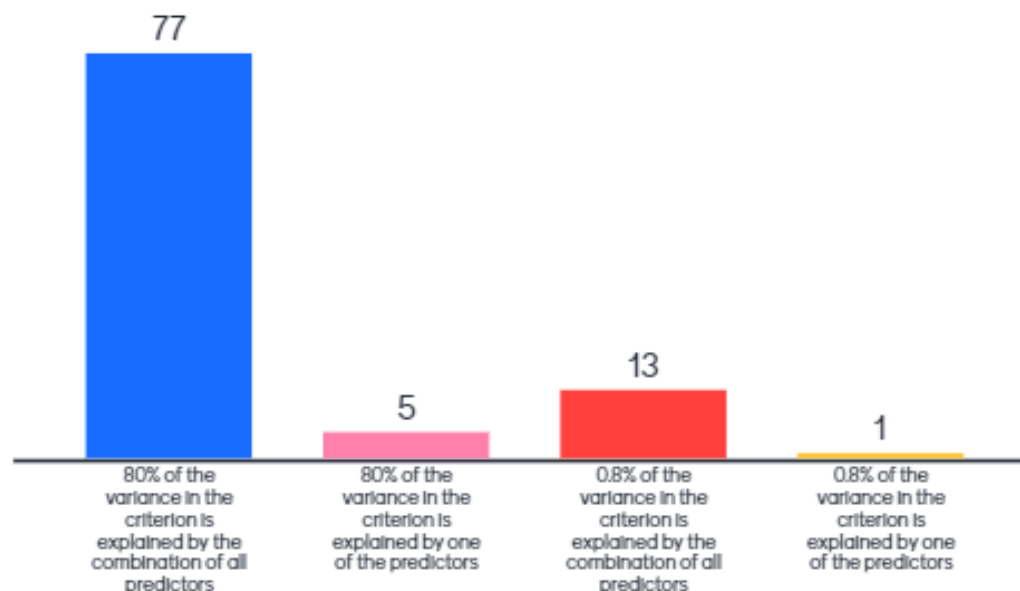
- To examine contribution of each predictor, we use the multiple regression equation

$$\hat{Y} = a + b_1X_1 + b_2X_2 + b_3X_3$$

$$\hat{Y} = 6.90 - .04(\text{age of immigration}) + .01(\text{years in Canada}) + .02(\text{height})$$

- Multiple correlation coefficient ( $R$ )       $R = .49$

The Multiple Correlation ( $R$ ) indicates the contribution of all predictors combined in predicting the criterion. What does an  $R^2$  of .80 mean?



## Cultural adaptation

- To examine contribution of each predictor, we use the multiple regression equation

$$\hat{Y} = a + b_1X_1 + b_2X_2 + b_3X_3$$

$$\hat{Y} = 6.90 - .04(\text{age of immigration}) + .01(\text{years in Canada}) + .02(\text{height})$$

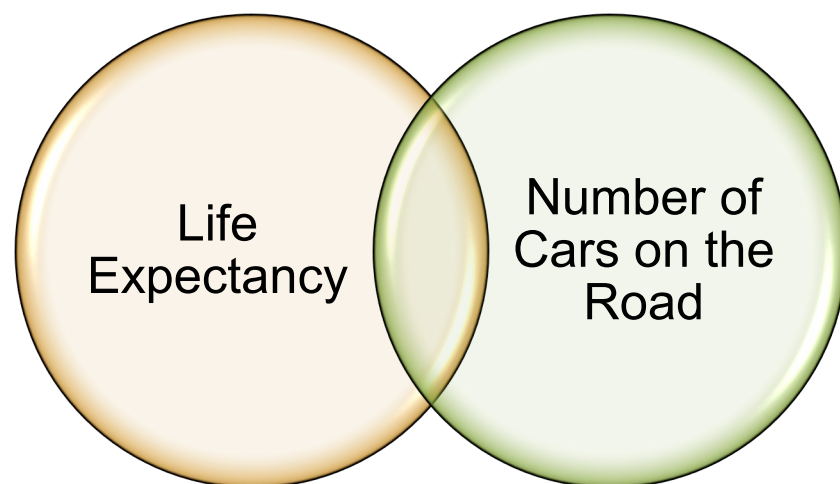
- Multiple correlation coefficient ( $R$ )  
 $R = .49$   
 $R^2 = .24$

## Partial correlation

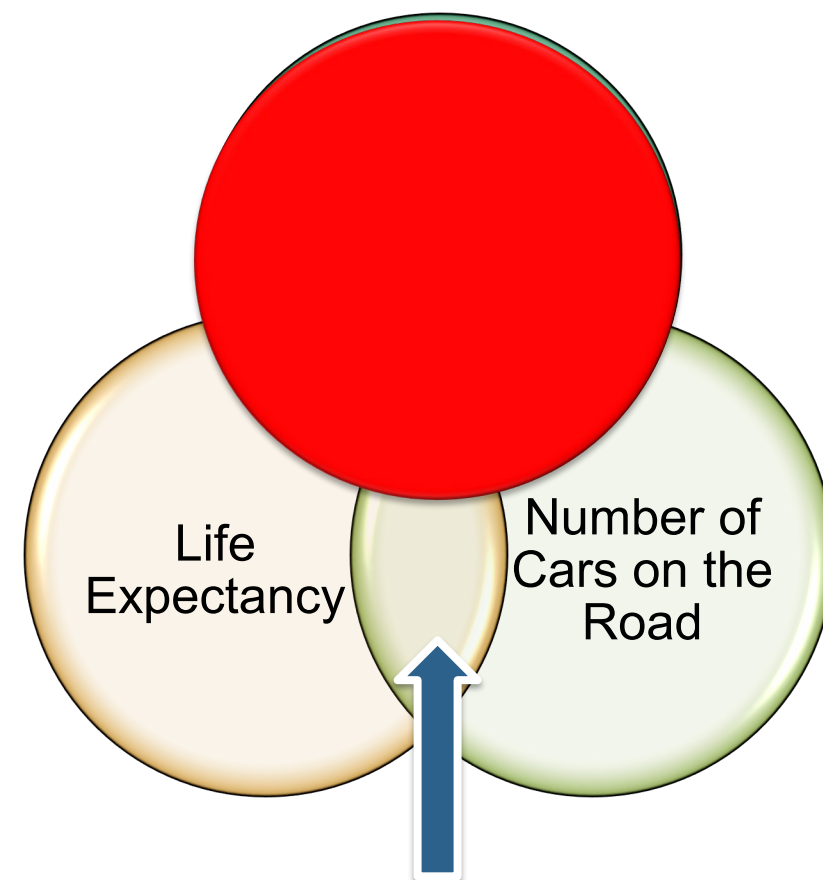
- 3<sup>rd</sup> variable problem
  - Some other variable is responsible for the relationship between X & Y
- Identify and statistically control for 3<sup>rd</sup> variable!

## Example of partial correlations

“SOCIETIES IN WHICH PEOPLE LIVE THE LONGEST HAVE THE MOST CARS! CARS IMPROVE LIFE EXPECTANCY!”



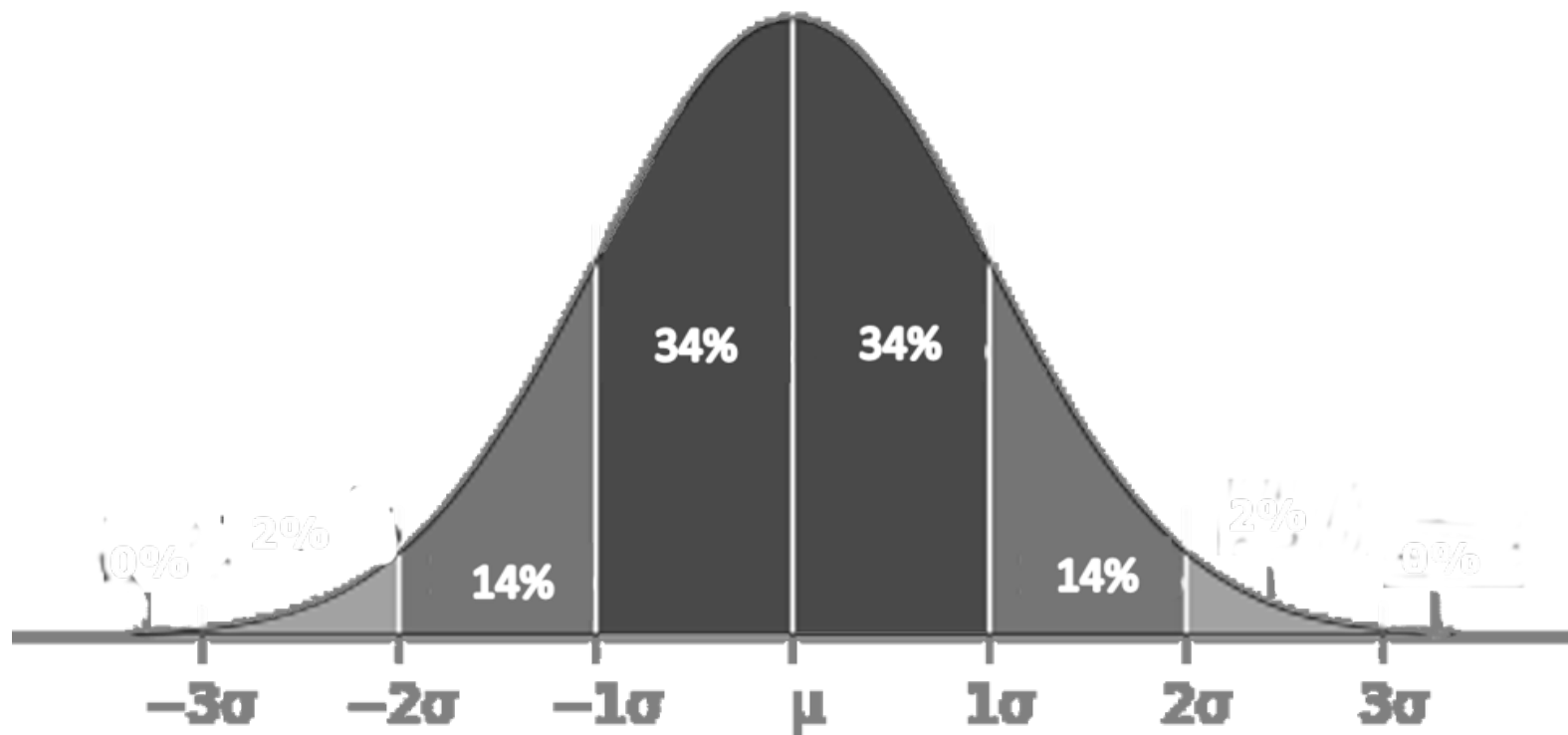
REMOVE THE EFFECTS OF...



# Measures of relationship

- Bivariate correlation ( $r$ )
  - Standardized index of how much two variables change with each other
- Squared Correlation ( $r^2$ )
  - Proportion of variability shared by 2 variables
- Multiple Regression
  - A technique used when we want to test how well one or more predictors individually predict the criterion
- Multiple Correlation ( $R$ )
  - A type of correlation coefficient that indexes how much a set of predictors, when combined, is related to the criterion
- Squared Multiple Correlation ( $R^2$ )
  - Proportion of variability in criterion accounted for by set of predictors
- Partial Correlation
  - A correlation between X and Y that statistically removes the effect of 3<sup>rd</sup> variable

# Any questions?





# Study, ask questions

