What You Can and Cannot Say With Data

Holly Stovall, Ph.D.
(Research Analyst)

Robert Lorick, M.S.M.R.
(Director of Academic Analysis)

Office of Institutional Intelligence & Research

Tarrant County College

Let's Make a Deal!

Tarrant
County
College

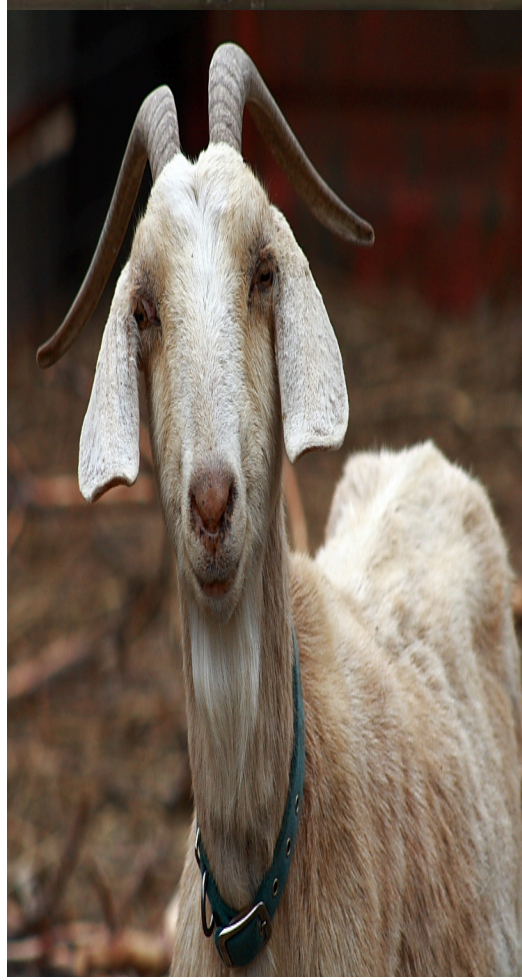# Let's Make a Deal!

1

2

3

Let's Make a Deal!
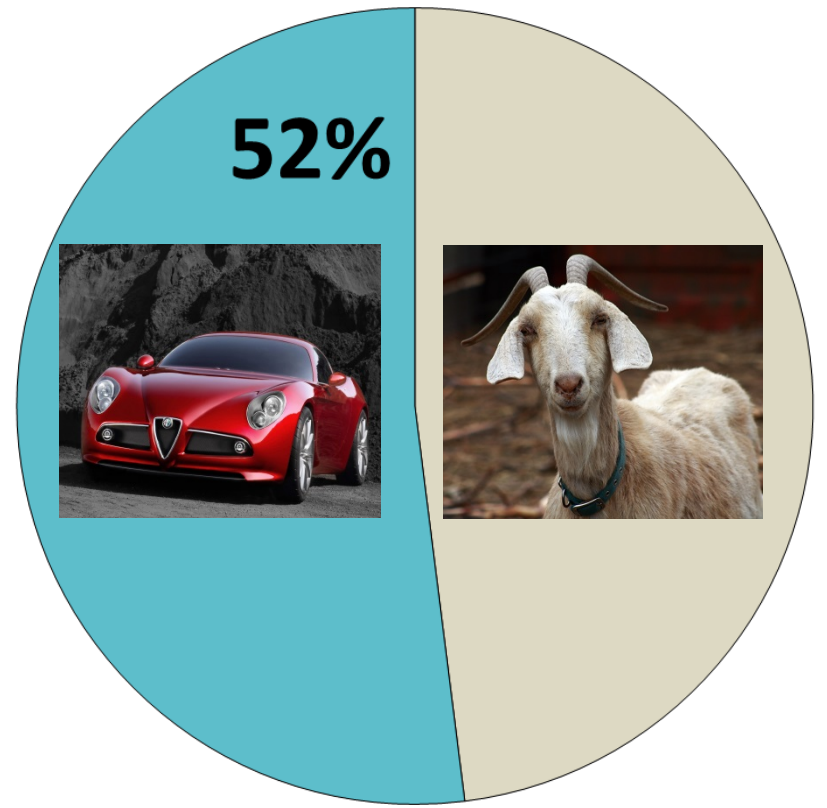
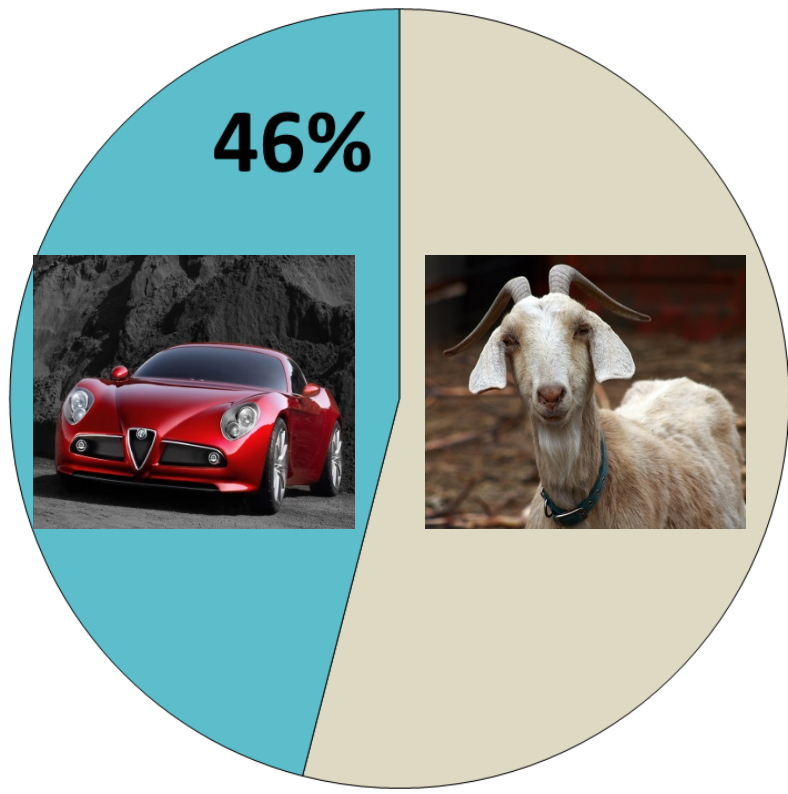# Let's Make a Deal!

Let's Make a Deal: Data

100 people

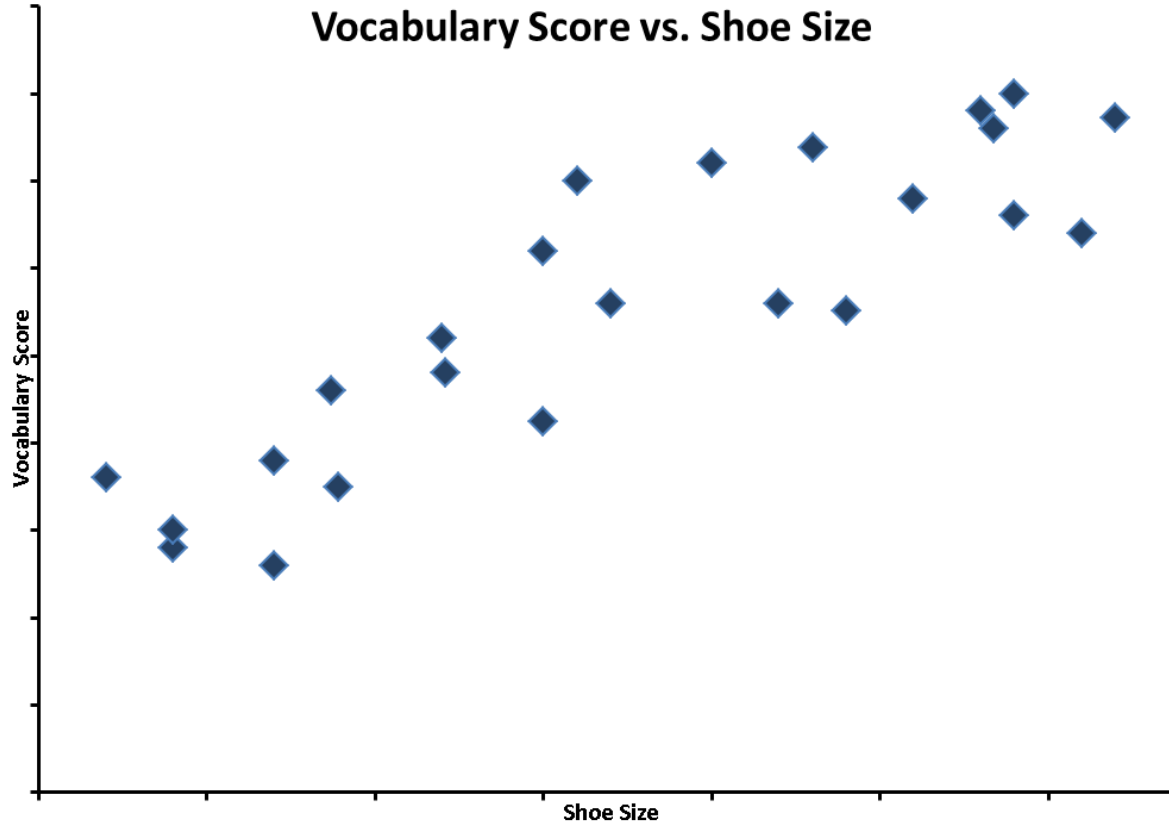- 50 players stay

- 50 players switch

Outline

- Points of Confusion
  - Correlation
  - Modeling
  - Sampling
  - Hypothesis Testing

- Group Assignment
  - Bias
  - Establishing Cause-and-Effect

- Group Selection
  - Bias
  - Generalizability

Correlation

Tarrant County College

Modeling: Linear Regression

X: Entering Exam Score
Y: Words per Minute

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	88.063	17.757		4.959	.000
	EnteringExam	-.456	.177	-.251	-2.571	.012

a. Dependent Variable: WordPerMinute

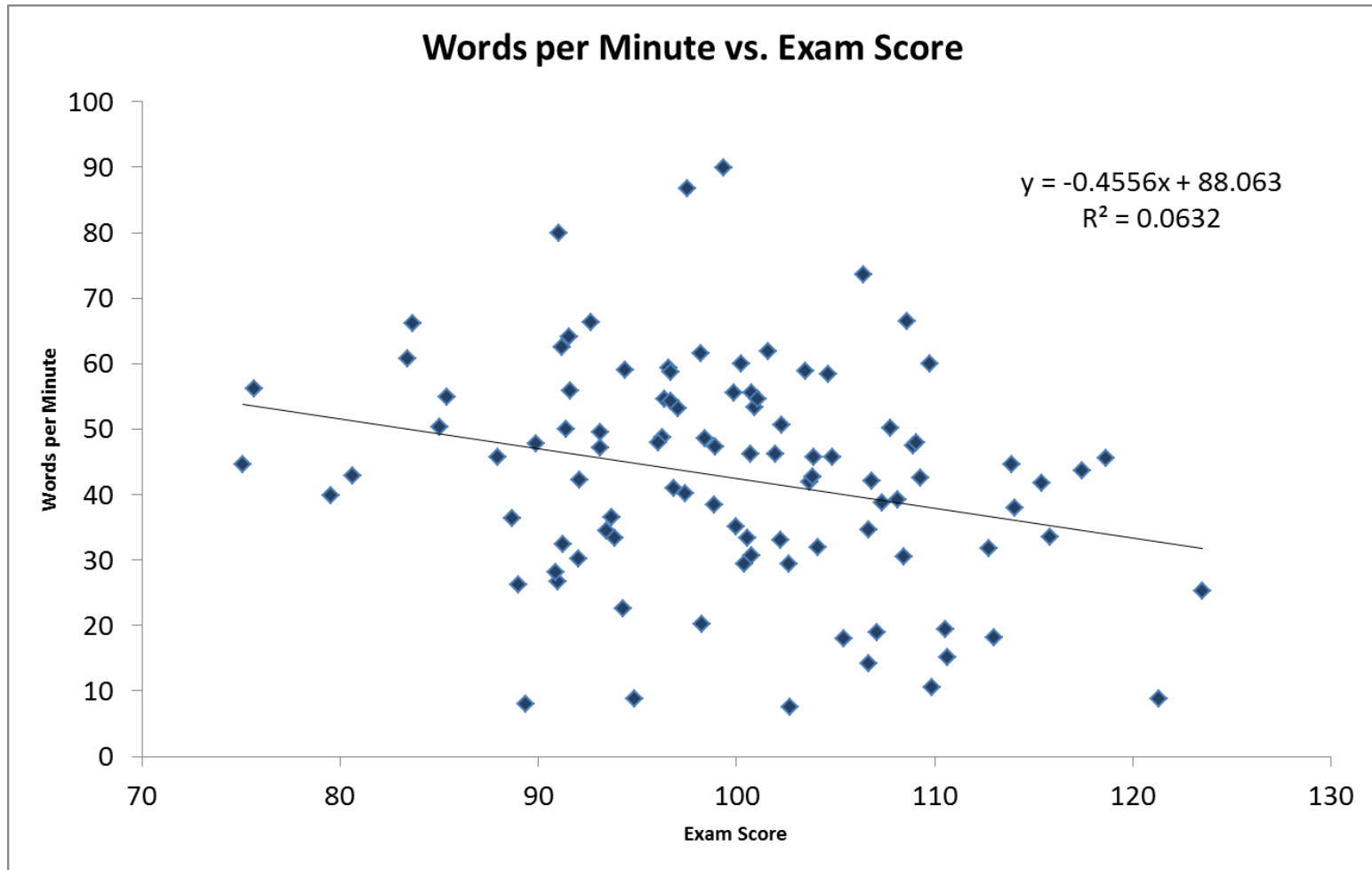# Modeling: Linear Regression

X: Entering Exam Score
Y: Words per Minute

Statistically Significant! 😊

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	88.063	17.757		4.959	.000
	EnteringExam	-.456	.177	-.251	-2.571	.012

a. Dependent Variable: WordPerMinute

Modeling: Linear Regression

Words per Minute vs. Exam Score

$y = -0.4556x + 88.063$
$R^2 = 0.0632$

Modeling: Linear Regression

Modeling: Logistic Regression
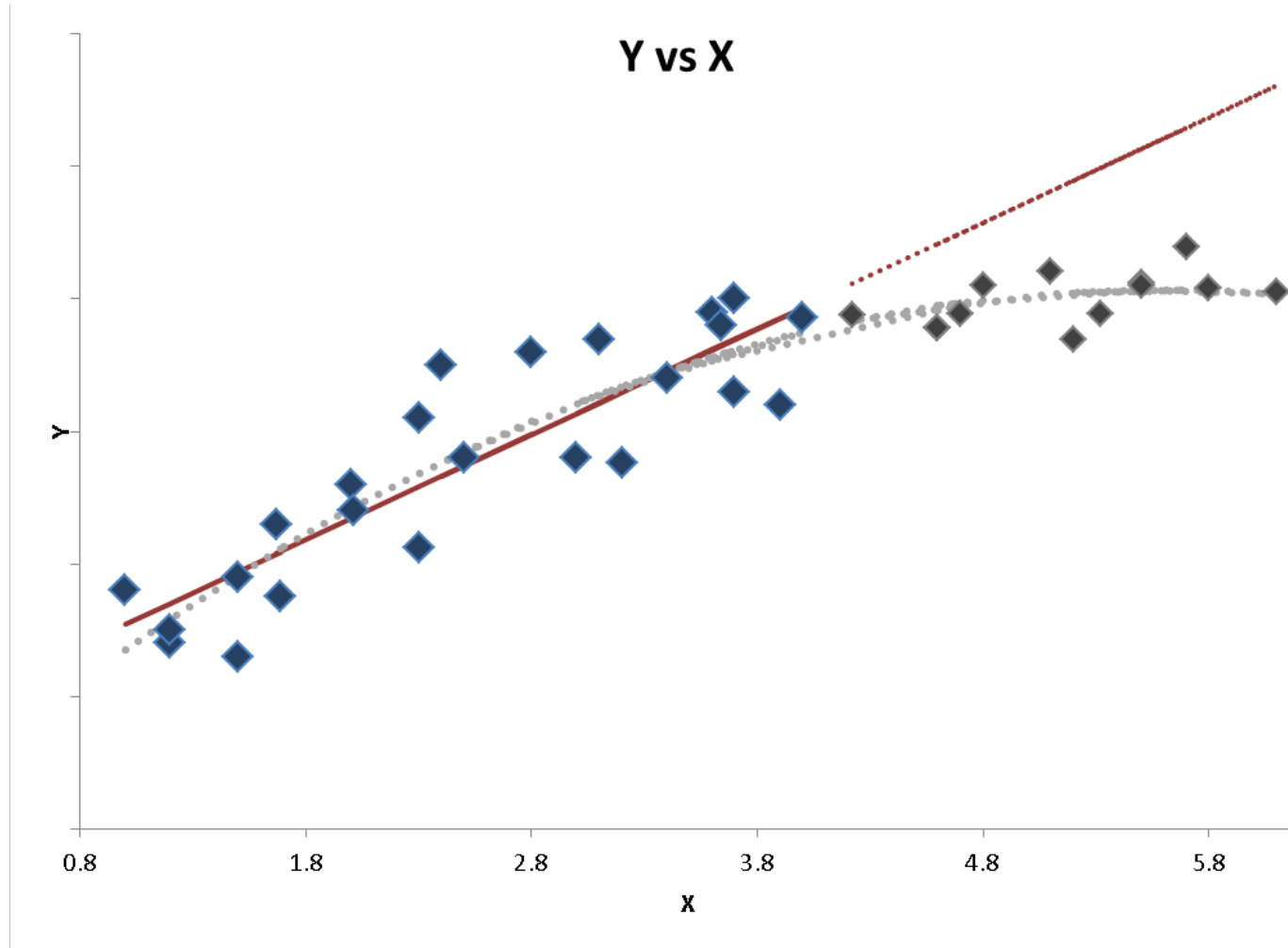
X: Gender
Y: Retained

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	Gender(1)	.046	.020	5.198	1	.023	1.047
	Constant	-.034	.014	5.644	1	.018	.967

a. Variable(s) entered on step 1: Gender.

Tarrant County College

Modeling: Logistic Regression

X: Gender
Y: Retained

Statistically Significant! 🙂

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	Gender(1)	.046	.020	5.198	1	.023	1.047
	Constant	-.034	.014	5.644	1	.018	.967

a. Variable(s) entered on step 1: Gender.
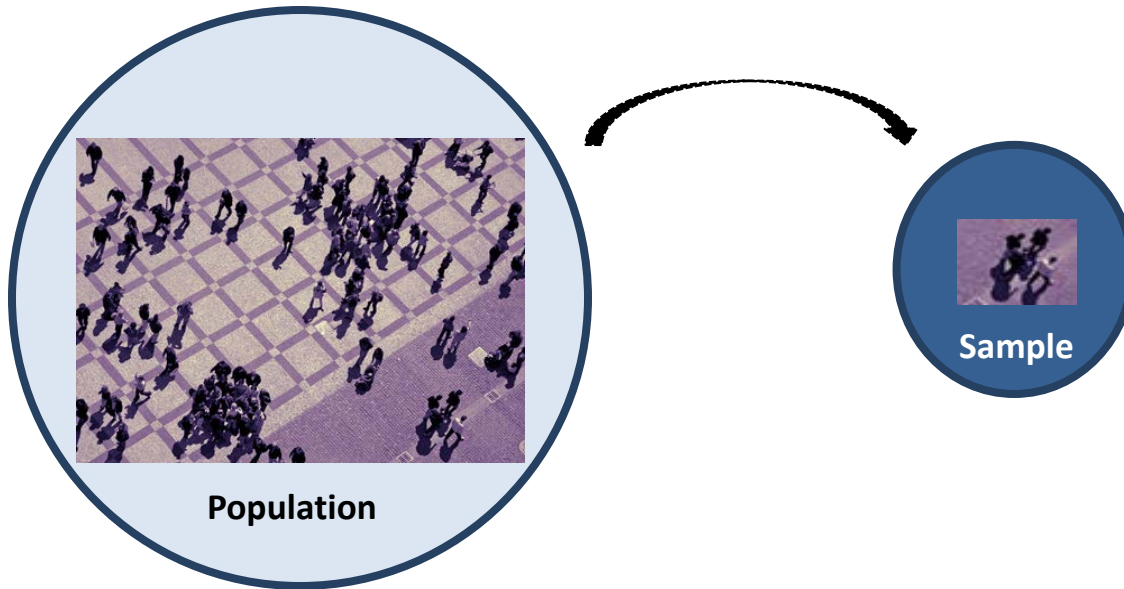
Modeling: Logistic Regression
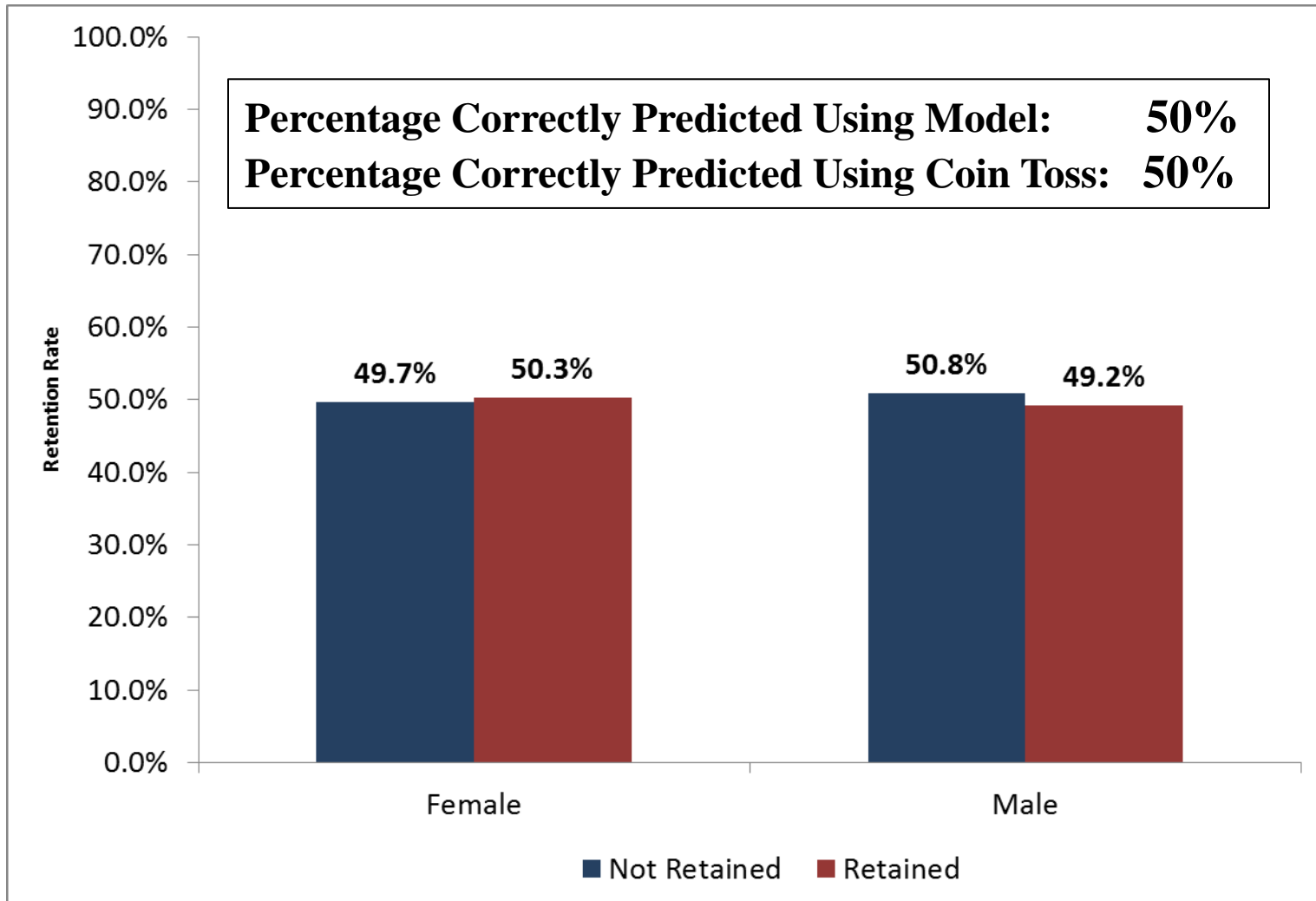
Sampling

- Are you analyzing a population or a sample?

Hypothesis Testing: Trial by Jury

H_0 = Not guilty

H_A = Guilty

Jury's Decision

	Not Guilty	Guilty
Innocent		
Guilty		

Truth

- Jury's decision does not prove the truth.
- Jury can get the decision wrong.

Hypothesis Testing: Trial by Jury

H_0 = Not guilty

H_A = Guilty

Jury's Decision

- Jury's decision does not prove the truth.
- Jury can get the decision wrong.

Hypothesis Testing: Trial by Jury

H_0= Not guilty

H_A= Guilty

Jury's Decision

Not Guilty Guilty

Truth

Innocent

Guilty

- Jury's decision does not prove the truth.
- Jury can get the decision wrong.

Statistical Significance

- Does not mean results are correct

- Does not mean results are more powerful or more meaningful

- Does not mean results have practical implications

Group Assignment

Tarrant County College
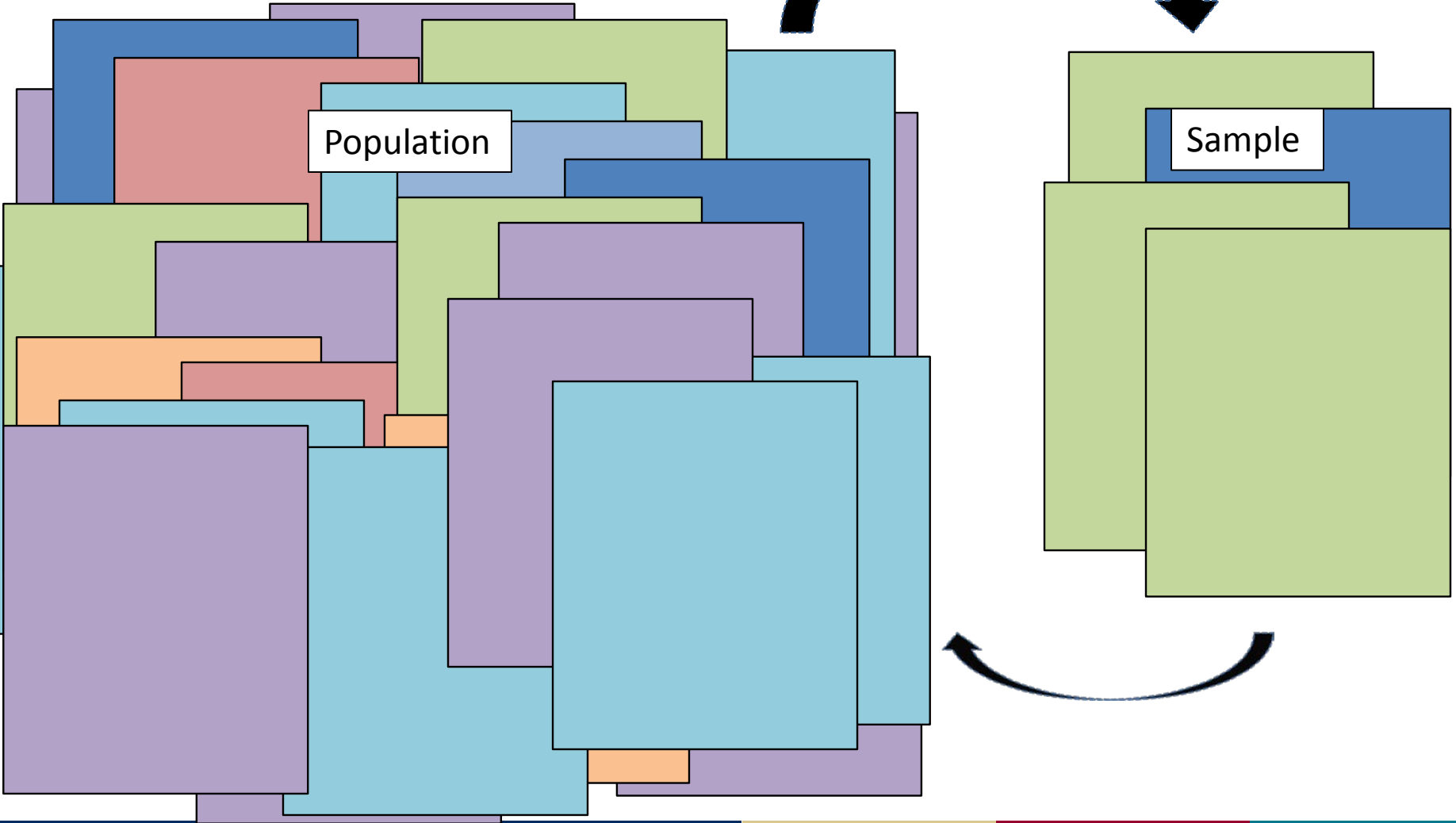
Implying Causality

Non-Review Group (Control)	Review Group (Treatment)

Tarrant County College

Group Selection

Tarrant County College

Generalizability

Population

Sample

Summary

Subjects are randomly assigned to groups.

	Yes	No
Yes Subjects are randomly chosen.	• **Generalizable** • **Causal implications**	• **Generalizable** • **No Causal implications**
No	• **Not Generalizable** • **Causal implications**	• **Not Generalizable** • **No Causal implications**

Final Thoughts

$$\text{Data} \overset{?}{=} \text{Proof}$$

$$\text{Data} \overset{?}{=} \text{Evidence}$$

Not Suggestive   Somewhat Suggestive   Highly Suggestive

Questions?

Tarrant
County
College

Contact Information

Holly Stovall, Ph.D.
Holly.Stovall@tccd.edu

Robert Lorick, M.S.M.R.
Robert.Lorick@tccd.edu

Office of Institutional Intelligence & Research

Tarrant County College