

# Key Problems for Data Science Projects

# Why Is This Problem Important?

## The Good Reasons:

Lack of Understanding Inhibits Decision Making

Prediction is Hard, Particularly About the Future

Where Is This Happening?

## The Bad Reasons:

Methodology (The cool tool problem)

Deliverable (Keeping up with the neighbors)

# Who Does This Problem Affect?

- Start with this question, not the data
- Good: Assistant Principal retention in our school district is much lower than we want
- Bad: Let's explore the factors associated with employee retention

Who will use the information gathered?

What will change about their work?



## Teacher Evaluation in Tennessee



*What We Have Learned from a Decade of Research*

*July 2021*

J. Edward Guthrie, Mónica Hernández, and Jason A. Grissom

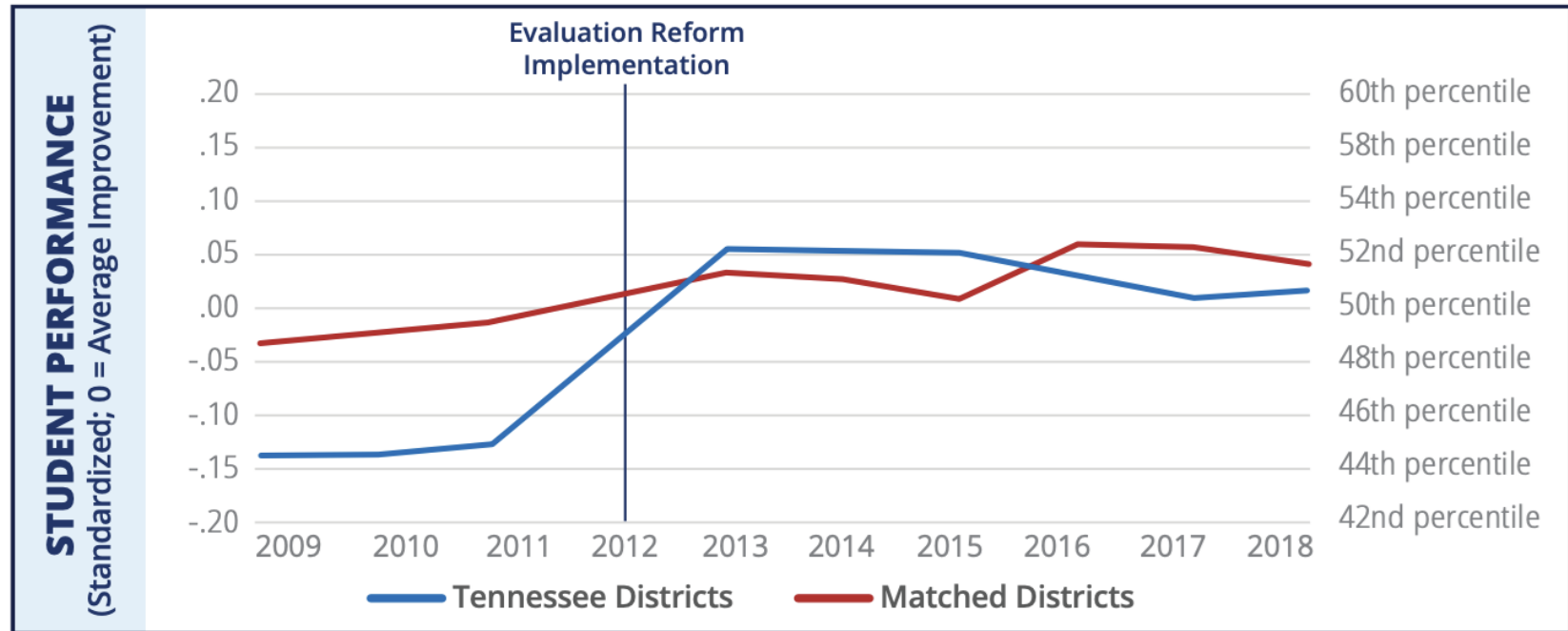
### Introduction

After piloting new teacher evaluation systems in the 2010-2011 academic year, the Tennessee Department of Education implemented a statewide comprehensive teacher evaluation system that took effect in 2011-2012.

**This brief unpacks four primary findings from the body of research on evaluation reform in Tennessee:**

 *Student achievement improved more after*

**FIGURE 1:** After evaluation reform, student performance in Tennessee districts increased at a much faster rate than similar districts in other states.



# What If We Don't Have the Right Data?

- VERY common issues
- Dependent variable: sampling on the outcome
  - Only successes
  - “Best Practices”
- Independent variable
  - No variation
  - Assuming Causality
  - Measurement, measurement, measurement

# Sampling on the Dependent Variable

OPINION

## Op-Ed: Why Finland has the best schools

BY WILLIAM DOYLE

MARCH 18, 2016 5 AM PT



The Harvard education professor Howard Gardner once advised Americans, “Learn from Finland, which has the most effective schools and which does just about the opposite of what we are doing in the United States.”

<https://www.latimes.com/opinion/op-ed/la-oe-0318-doyle-finnish-schools-20160318-story.html>

# When Is the Project Over?

The number one rule: don't define success as finding what you want to find.

One key way to get here is to define the number or set of numbers that you're looking for:

Good: To what extent did our placement of ads increase the number of applicants?

Bad: How do we get more applicants?

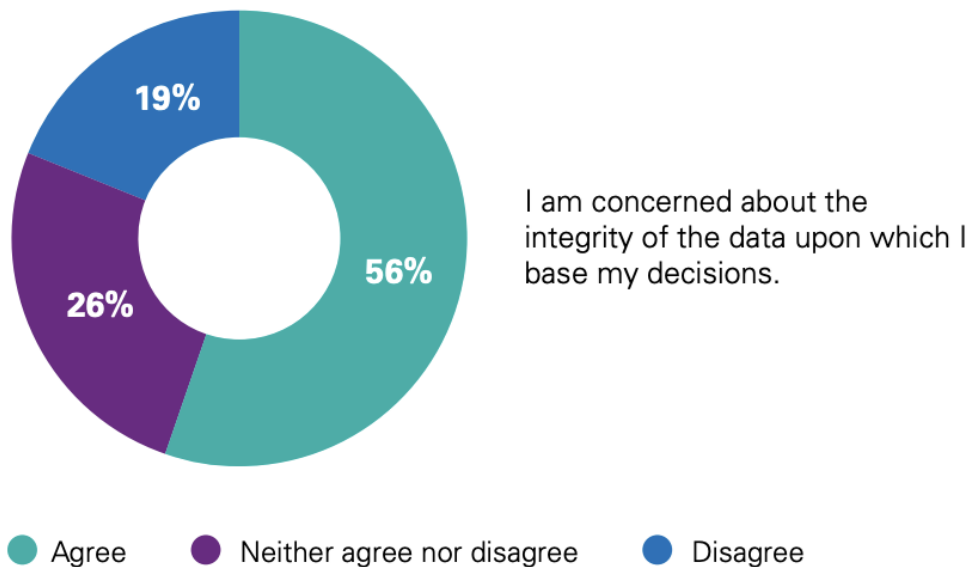




# Trusted analytics matter to CEOs

<https://assets.kpmg.com/content/dam/kpmg/xx/pdf/2017/08/da-trusted-analytics-ceo-outlook.pdf>

**Figure 3: CEOs doubt their data**



Source: 2017 Global CEO Outlook, KPMG International

# Why Do Data Science Projects Fail?

- Expectations of proving a “pet hypothesis”
- Links between results and decisions are not defined
- User stories are not utilized in the design phase: outputs are analyst focused