

# Title

Maria Tackett  
Duke University

JSM  
August 2020



# Outline

- The course
- Learning objectives + data analysis life cycle
- Undercounts in the US Census
  - Demographic Analysis
  - Dual-Systems Estimates
- Class activity
- Ideas for implementing in an online course

# The course: regression analysis

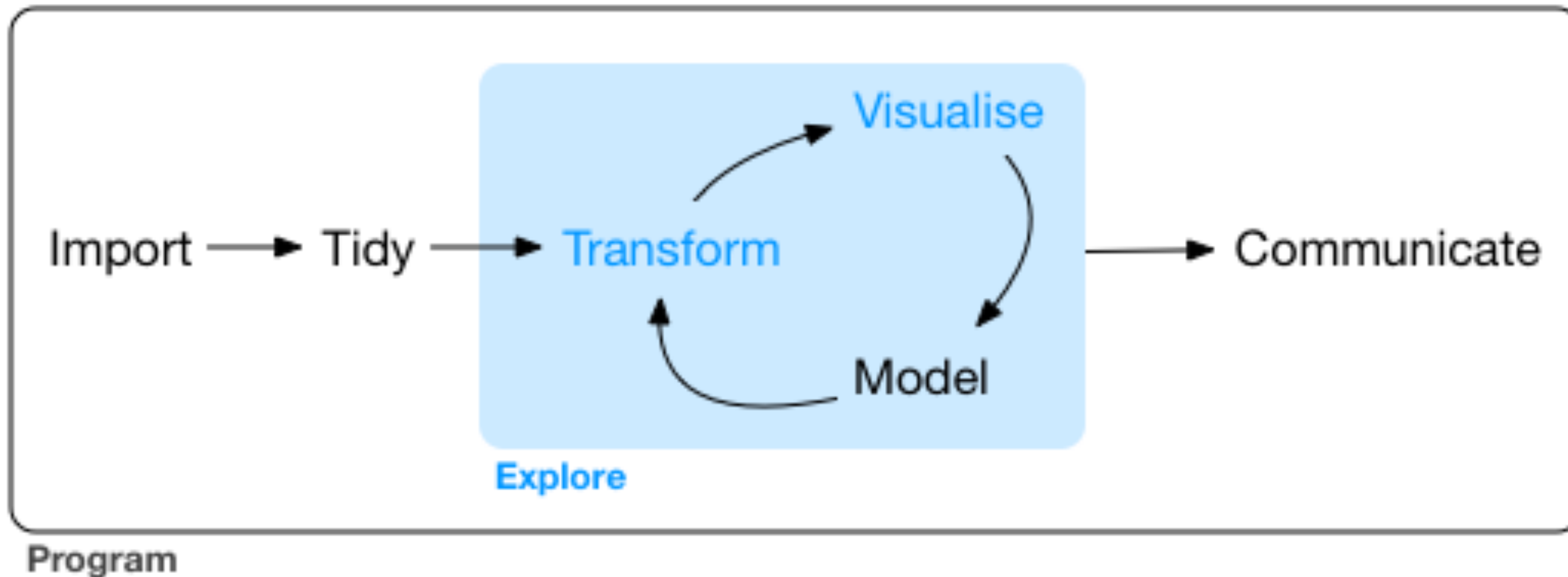
- Second semester undergraduate statistics course (~ 90 students)
  - Pre-req: Intro statistics, data science, or probability
- Topics: multiple linear regression, logistic regression, ANOVA
  - Focus on application rather than mathematical theory
- Computing using R and GitHub

# Missing data

- Part of a “modeling in practice” unit
- Learning objectives:
  - Identify different types of missingness
  - Understand why it’s important to understand missingness
  - Apply strategies for dealing with missingness + pros / cons for each

# Data science life cycle

- Focus on explore, specifically on modeling

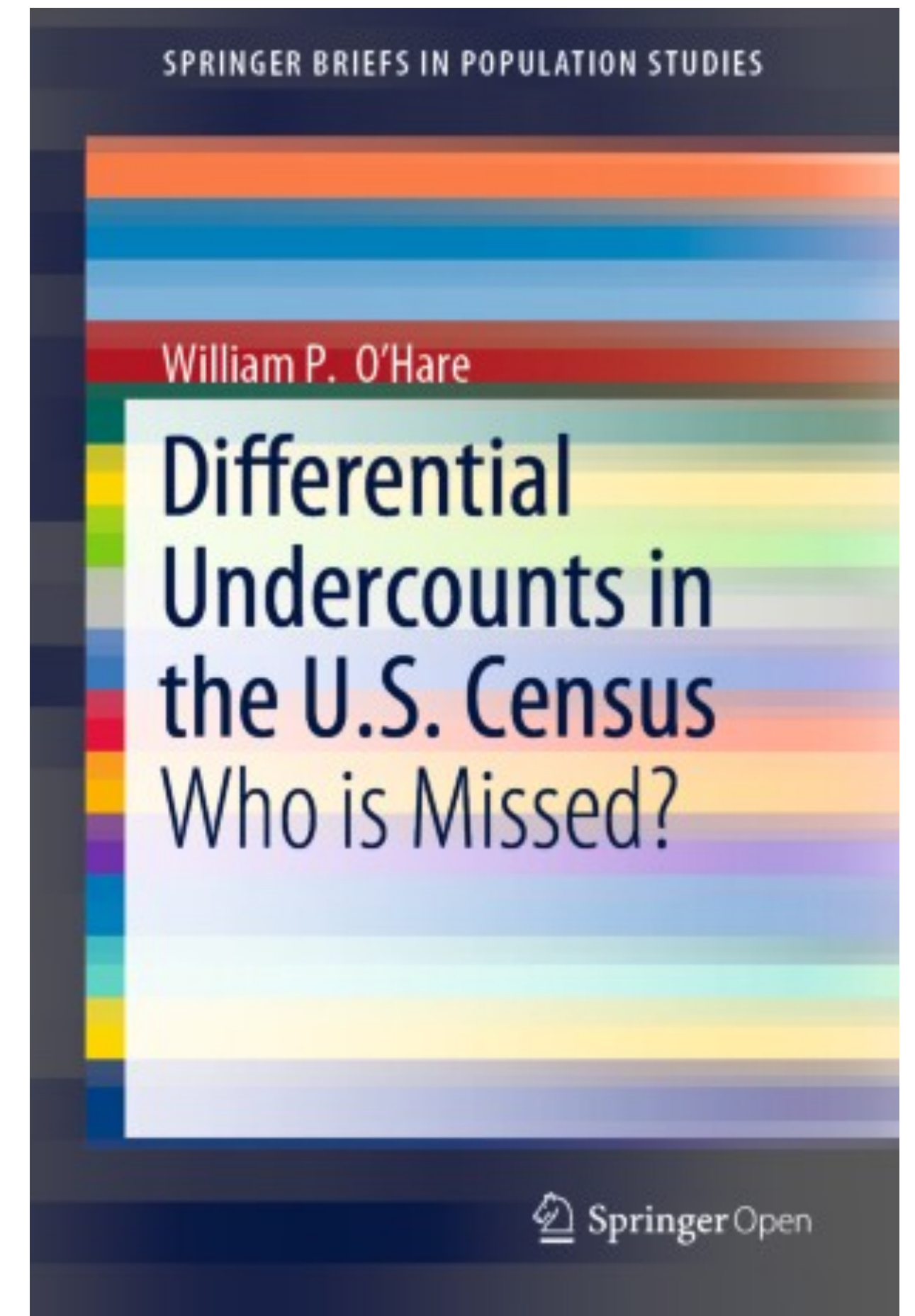


# US Census

- Objective: Get a count of everyone living in the United States
- How census is used
- Issue: Some populations are undercounted and thus underrepresented in the census
- Models to measure undercount:
  - Demographic Analysis
  - Dual-Systems Estimates

# Activity outline

- Learn the different types of missingness and strategies to handle missing data
- Discuss the US Census + implications of missing
- Critique DA and DSE models
- Use ACS data to fit a model that can be used to identify undercounts
- Compare model results to those of DA and DSE





# Activity details

- Fill in details for each step of the activity

# Implementing activity

- Ideas to implement this activity in an online course

# Recap of learning objectives

- Identify different types of missingness
- Understand why it's important to understand missingness
- Apply strategies for dealing with missingness + pros / cons for each

# Find lesson materials

- Add link

# Thank You!

 maria.tackett@duke.edu

 @MT\_statistics

