

# STAT 113

## Data, Variables and Sampling

Colin Reimer Dawson

Oberlin College

29 August 2017

# Race and the Death Penalty

Data from 1981 Florida Homicide Convictions

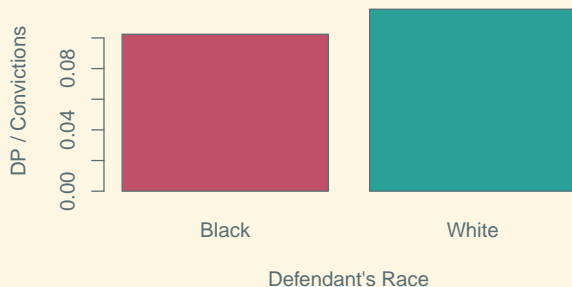


Figure: Proportions of Death Sentences out of Total Convictions, by Defendant's Race. Source: Agresti (2002)

## Race and the Death Penalty

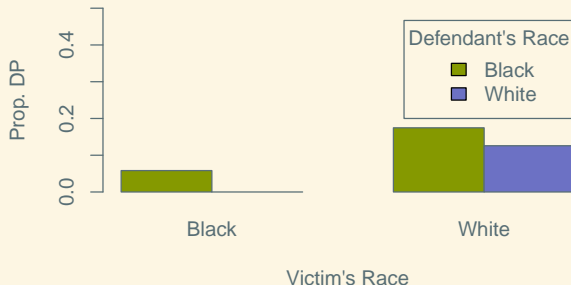


Figure: Proportions of Death Sentences out of Total Convictions, by both Victim's and Defendant's Race. Source: Agresti (2002)

Black defendants are sentenced to death at higher rates with both white and black victims. But combined, white defendants have a (slightly) higher rate. Why the seeming contradiction? 4 / 9

## Cases, AKA “Observational Units”

**Cases** When we collect data, we write down some measurements or characteristics of our **cases** — the individual “entities” or “observational units” that make up our dataset.

- The people in a survey or research study
- Plots of land in an agricultural experiment
- Days, in a weather dataset

## Categorical vs. Quantitative Variables

For each case we record one or more **variables**. One of the most basic distinctions is between **categorical** (or “qualitative”) and **quantitative** data.

**Categorical:** “Qualitative” variable that divides cases into groups

**Quantitative:** Measures something on a scale; arithmetic makes sense

# Categorical vs. Quantitative Variables

What are each of these?

- Favorite color **Categorical**
- Height in centimeters **Quantitative**
- Age in years **Quantitative**
- T-number **Categorical**

Notice that just using numbers isn't enough to make a variable numeric. Need to ask: "do the numbers represent an underlying *quantity* or *scale*?"

## Pair Quiz! (5 min.)

1. Which of the following are the cases that make up the Florida death penalty data?
  - (A) Victim's Race
  - (B) Defendant's Race
  - (C) Defendants
  - (D) Convicted Defendants
2. Sketch an appropriate table layout (like a "data frame") for this data. Rows are cases, columns are variables.
3. What are the variables we have looked at?
4. What type is each one? (Categorical/Quantitative)