

# Intro to Social Science Data Analysis

## Lecture 5: Descriptive Statistics

**Christopher Gandrud**

September 24, 2012

1 Recap

2 Describing Data Overall

# Review Questions 1

- ▶ What is reproducible research?
- ▶ Why is it important?

# Review Questions 1

- ▶ What is reproducible research?
- ▶ Why is it important?

## Review Questions 2

- ▶ What is the Markdown markup language?
- ▶ What does the *knitr* package do?

## Review Questions 2

- ▶ What is the Markdown markup language?
- ▶ What does the *knitr* package do?

## Review Questions 3

- ▶ What is a code chunk?
- ▶ How do you make a code chunk in a Markdown document?

## Review Questions 3

- ▶ What is a code chunk?
- ▶ How do you make a code chunk in a Markdown document?



So far we have learned how to gather data  
and get it into R.

Today we will start to learn tools for  
**describing** our data.

Why do we need tools for describing our data?

# Why do we need tools for describing our data?

- ▶ To find the **patterns** we are interested but too difficult to find by just looking at the raw data.
- ▶ Find potential **data biases**.

# Why do we need tools for describing our data?

- ▶ To find the **patterns** we are interested but too difficult to find by just looking at the raw data.
- ▶ Find potential **data biases**.

When describing data, **ALWAYS** look at **BOTH**

- ▶ The Central Tendency,
- ▶ The Variance (dispersion).

When describing data, **ALWAYS** look at **BOTH**

- ▶ The Central Tendency,
- ▶ The Variance (dispersion).

## References I

Diaz, David M., Christopher D. Barr, and Mine Çetinkaya-Rundel.  
OpenIntro Statistics. 1st ed.

<http://www.openintro.org/stat/downloads.php>.