

# Quantitative Methods in Political Science

## Recitation

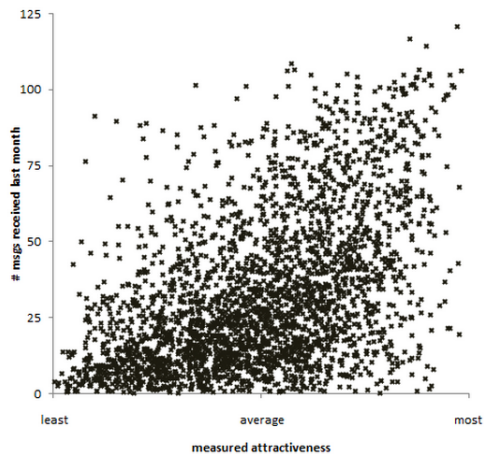
Mai Nguyen

New York University

February 12, 2014

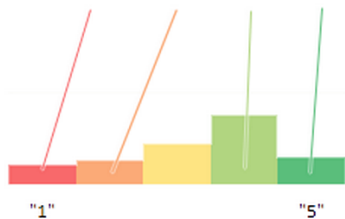
## Messages Received Last Month vs. Attractiveness

5,000 users: each x represents an actual person



# Introduction

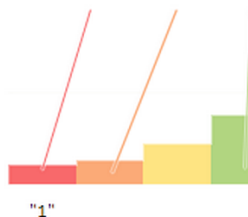
$$msgs = .4m_1 - .5m_2 - .1m_4 + .9m_5 + k$$



Source: OkTrends

# Introduction

$$msgs = .4m_1 - .5m_2 - .1m_4 + .9m_5 + k$$



## If You're Into Algebra

We did a regression on the votes for and messages to a sample of 43,000 women. To keep everything consistent, all the women were straight, between the ages of 20 and 27, and lived in the same city. The formula given in the body of the post was the best-fit we found on our second regression, after dropping the  $m_3$  term because its  $p$ -value was very near 1.

$msgs$  are the number of messages the woman received during the observation period. The constant  $k$  reflects her overall level of site activity. For this equation,  $R^2 = .28$ , which isn't great in a lab or on a problem set, but is actually very good in a real-world environment.

Source: OkTrends

## **TA: Mai Nguyen**

- Email: [mai.nguyen@nyu.edu](mailto:mai.nguyen@nyu.edu)
- Office: Room 318, 19 West 4th St.
- Office hours: Monday 2-4pm

## What is lab about?

- Reviewing what you learned in class.
- Reviewing your homework.
- Applying what you learned in class.
- Learning to use STATA.

## Requirements

- Weekly homework assignments
- Final Project: worth 20%
  - Due at the end of term
  - Exact format given mid-semester
  - Essentially a series of small assignments for the last month
- Attendance to recitations is mandatory.

# Why Quantitative Methods?

- Who is Billy Beane? Nate Silver?
- How do Google and Facebook know us so well?
- **Statistical Analysis** is a powerful tool
  - Make sense of the world in a rigorous fashion.
  - Allows for dispassionate analysis.



## Step one: Theory

- “A general statement about how the world works.”
- Attempts to (1) *explain* and (2) *predict* certain phenomena in the world.
- Implies a **causal relationship** between an **independent** and a **dependent** variable.
  - $X \rightarrow Y$
- To infer causality:
  - Correlation
  - Causal direction
  - Nonspuriousness

## Step two: Generate hypotheses

- Conditional statement; a claim or prediction that can be tested
- Hypothesis template:
  - In a comparison of [unit of analysis] those having/being [one value of the independent variable] will be more/less likely to have/be [one value of dependent variable] than those having/being [a different value of independent variable].

Examples:

- Democracy decreases conflict.
- Education affects voter turnout.
- Aid leads to corruption.

## Step three: Gathering data

- Choosing the appropriate data:
  - Qualitative
  - Quantitative
- Data sources:
  - Economic data: Penn World Tables, IMF, World Bank, WTO...
  - Political data: World Values Survey, Afrobarometer, Correlates of War, UCD-PRIO
  - Make sure your data is reliable.
  - Good website: [www.icpsr.umich.edu](http://www.icpsr.umich.edu)

Step four: Analyze the data!

- Stata
- R
- SPSS
- SAS
- Matlab
- Excel

- We will use Stata 12
- Available on NYU lab computers
- If you want to access Stata from home: Virtual Computer Lab
  - <http://www.nyu.edu/its/vcl/>
- Phillip Pollock. 2006. *A Stata Companion to Political Analysis*, CQ Press
- Great websites:
  - [www.ats.ucla.edu/stat/stata](http://www.ats.ucla.edu/stat/stata)
  - [www.stata.com/help.cgi?contents](http://www.stata.com/help.cgi?contents)

## Datasets:

- Przeworski's data on political institutions
- Persson and Tabellini on Constitutions
- NOTE: use the datasets we provide on NYU classes, not those that are found on the internet (we modified them).

Questions?