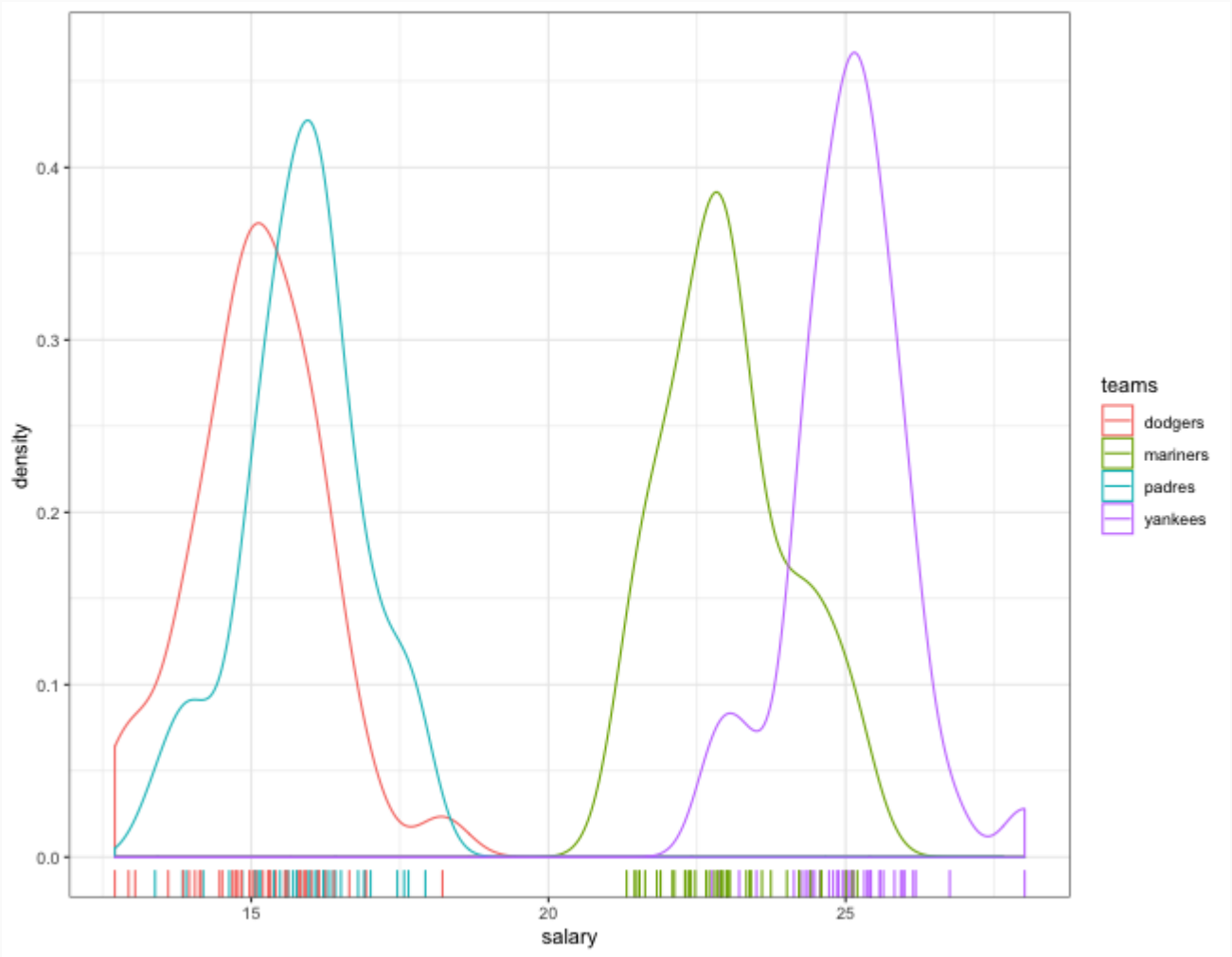
# Experimental Design

# Prelude: Sampling

```
teams <- c(rep("yankees", 40),
           rep("padres", 40),
           rep("mariners", 40),
           rep("dodgers", 40))
salary <- c(rnorm(40, mean = 25),
            rnorm(40, 16),
            rnorm(40, 23),
            rnorm(40, 15))
df <- data.frame(teams, salary)
head(df)
```

```
##     teams   salary
## 1 yankees 23.49881
## 2 yankees 24.99394
## 3 yankees 24.88391
## 4 yankees 26.17822
## 5 yankees 25.42410
## 6 yankees 24.85884
```

# Population view



# Simple Random Sample (SRS)

```
# population mean
mean(df$salary)
```

```
## [1] 19.75558
```

```
# SRS
df %>% sample_n(40) %>% summarize(mean(salary))
```

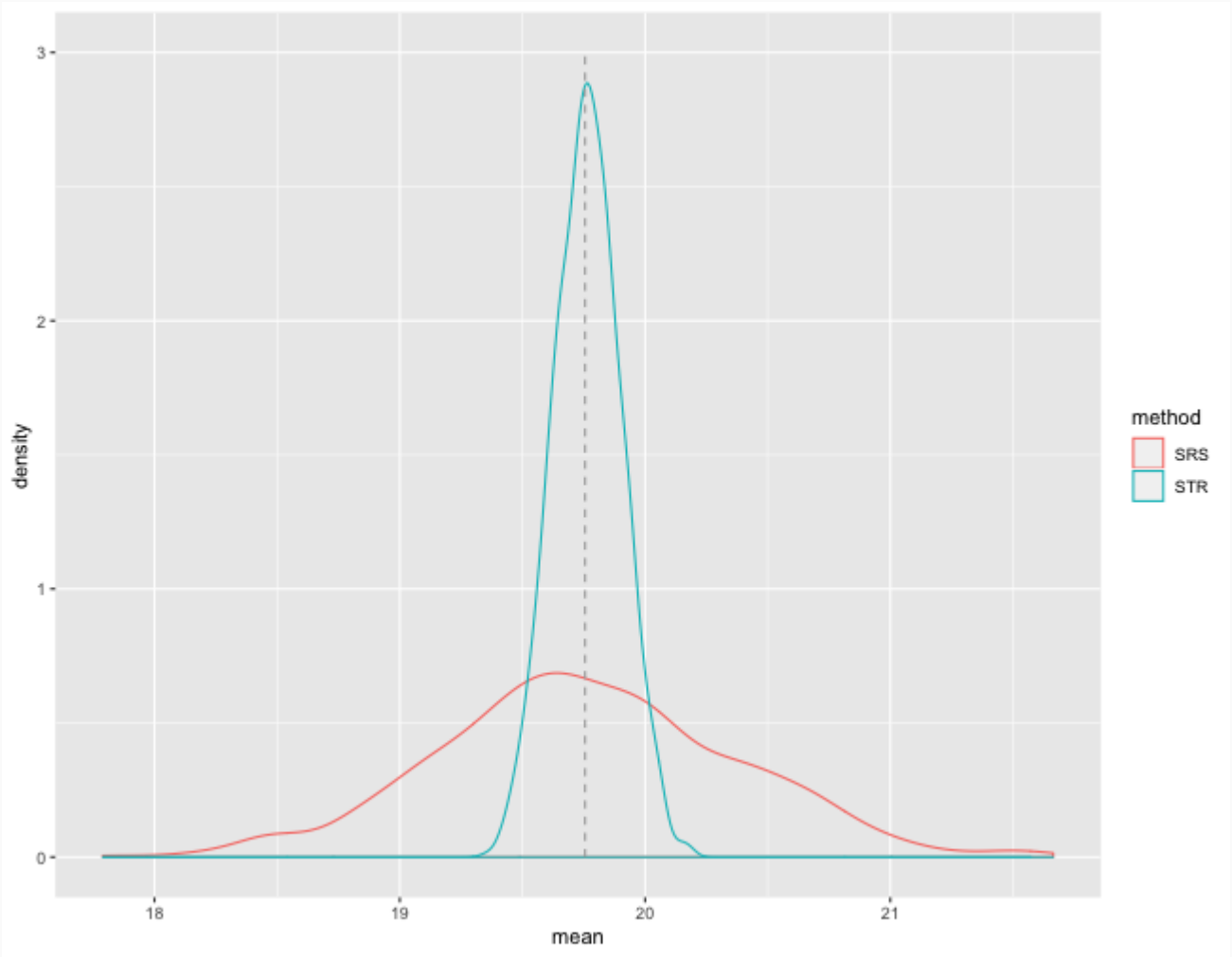
```
##   mean(salary)
## 1     17.96994
```

# Stratified Sample

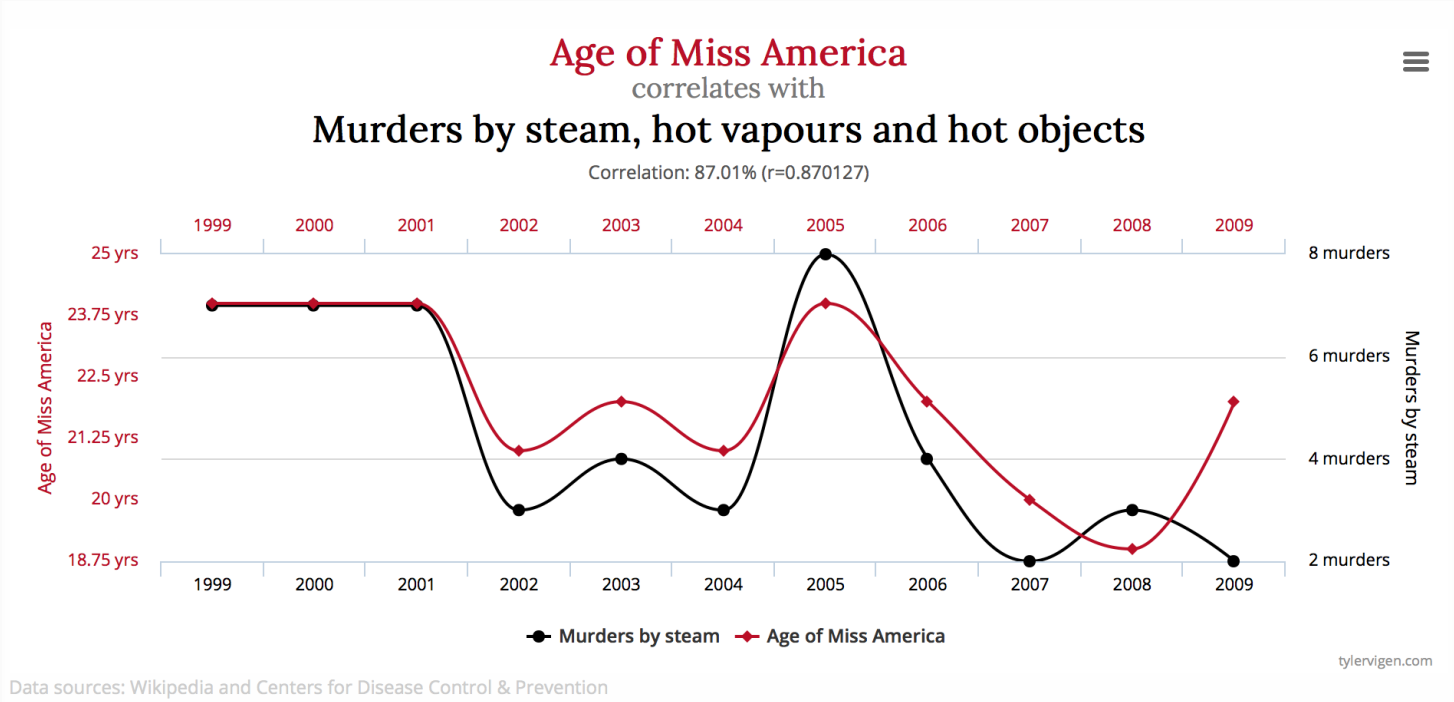
```
# Stratified sample
df %>%
  group_by(teams) %>%
  sample_n(10) %>%
  ungroup() %>%
  summarize(mean(salary))
```

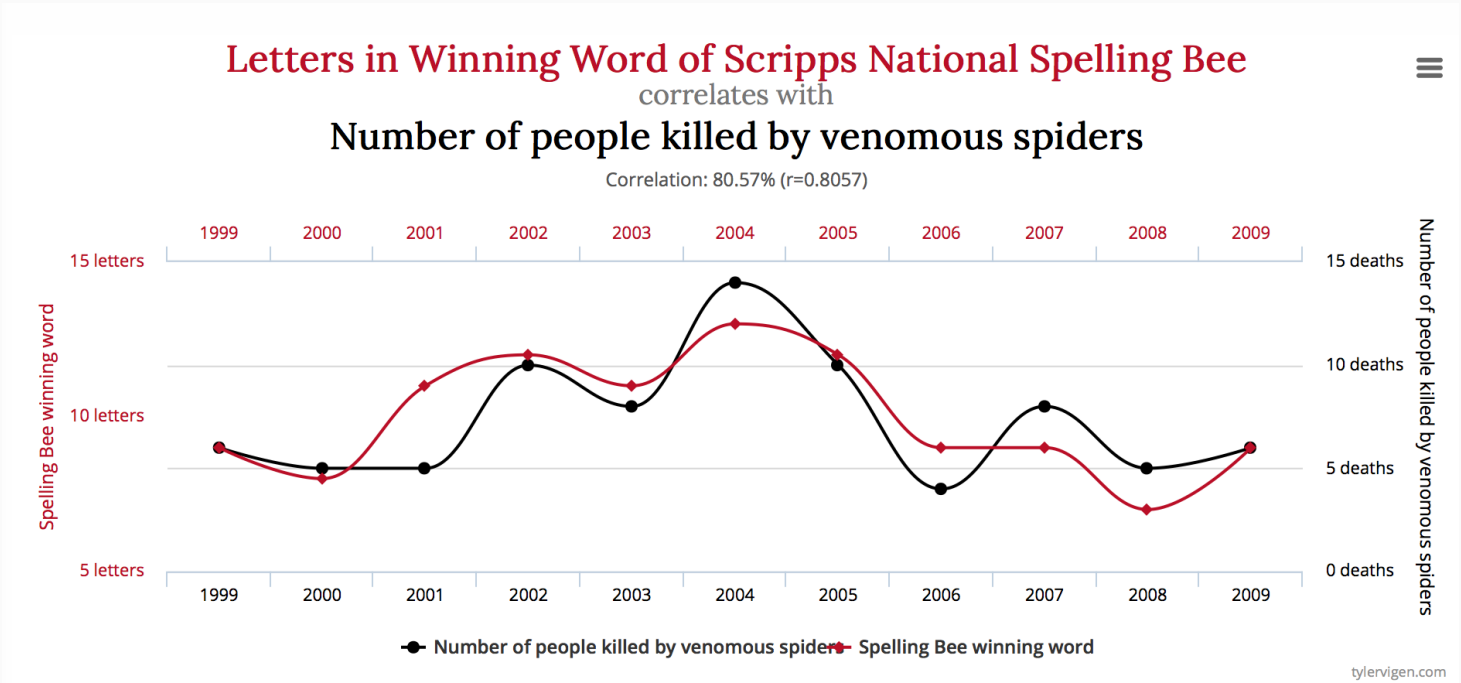
```
## # A tibble: 1 x 1
##   mean(salary)
##          <dbl>
## 1         19.4
```

# Long-run performance



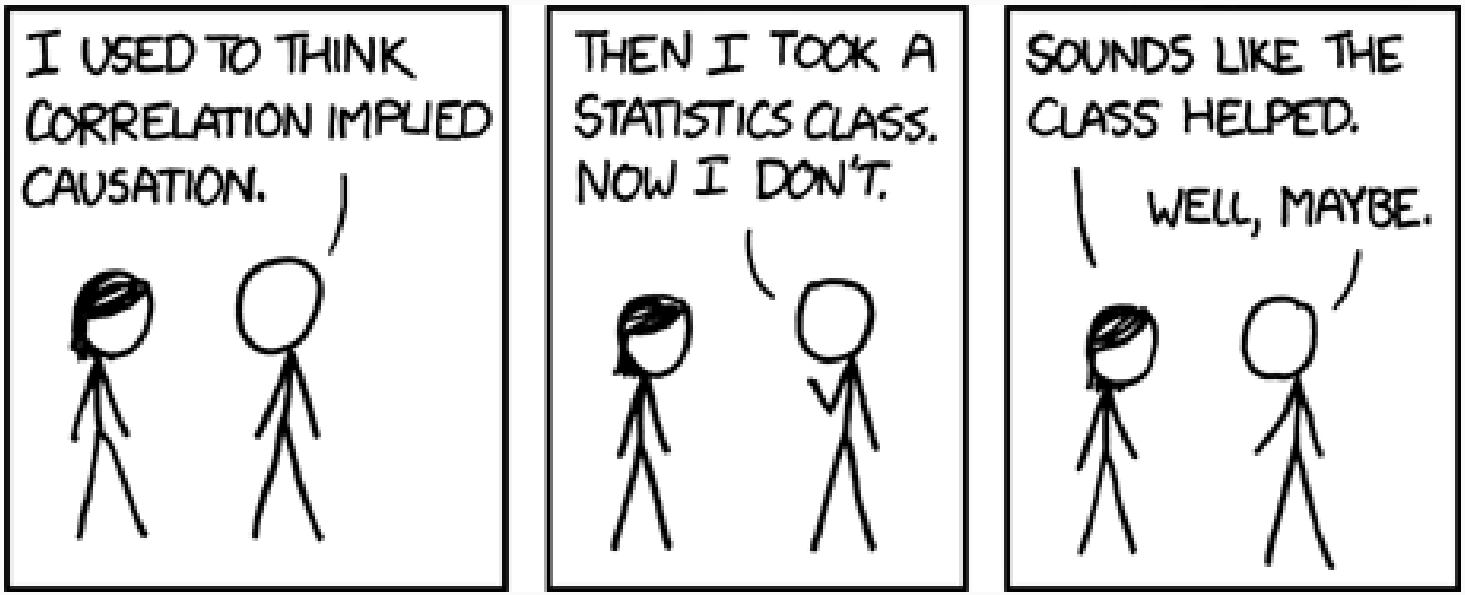
# Back to Experimental Design


Age of Miss America
correlates with
Murders by steam, hot vapours and hot objects
Correlation: 87.01% (r=0.870127)
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
25 yrs
23.75 yrs
22.5 yrs
21.25 yrs
20 yrs
18.75 yrs
Age of Miss America
8 murders
6 murders
4 murders
2 murders
Murders by steam
Murders by steam
Age of Miss America
tylervigen.com
Data sources: Wikipedia and Centers for Disease Control & Prevention




Data sources: National Spelling Bee and Centers for Disease Control & Prevention

# If you learn one thing in this class...



# Principles of Experimental Design

**Control**: Compare treatment of interest to a control group.

**Randomization**: Randomly assign subjects to treatments.

**Replication**: Within a study, replicate by collecting a sufficiently large sample. Or replicate the entire study.

**Blocking**: If there are variables that are known or suspected to affect the response variable, first group subjects into blocks based on these variables, and then randomize cases within each block to treatment groups.

# Replication

SCIENCE

## *Many Psychology Findings Not as Strong as Claimed, Study Says*

By BENEDICT CAREY AUG. 27, 2015

Email

Share

Tweet

Save

More


The past several years have been bruising ones for the credibility of the social sciences. A star social psychologist was caught fabricating data, leading to more than 50 retracted papers. A top journal published a study supporting the existence of ESP that was widely criticized. The journal Science pulled a political science paper on the effect of gay canvassers on voters' behavior because of concerns about faked data.

Now, a painstaking yearslong effort to reproduce 100 studies published in three leading psychology journals has found that more than half of the findings did not hold up when retested. The analysis was done by research psychologists, many of whom volunteered their time to double-check what they considered important work. Their conclusions, reported Thursday in the journal Science, have confirmed the worst fears of scientists who have long worried that the field needed a strong correction.

# Blocking

A study is designed to test the effect of light level and noise level on exam performance of students. The researcher also believes that light and noise levels might have different effects on males and females, so wants to make sure both genders are represented equally under different conditions. Which of the below is correct?

1. There are 3 explanatory variables (light, noise, gender) and 1 response variable (exam performance)
2. There are 2 explanatory vars (light and noise), 1 blocking var (gender), and 1 response var (exam performance)
3. There is 1 explanatory var (gender) and 3 response vars (light, noise, exam performance)
4. There are 2 blocking vars (light and noise), 1 explanatory var (gender), and 1 response var (exam performance)

# Other key ideas

**Placebo**: fake treatment, often used as the control group for medical studies

**Placebo effect**: experimental units showing improvement simply because they believe they are receiving a special treatment

**Blinding**: when experimental units do not know whether they are in the control or treatment group

**Double-blind**: when both the experimental units and the researchers do not know who is in the control and who is in the treatment group

# Consider acupuncture


How do you test if acupuncture reduces pain?

"Sham acupuncture" is a good control.

# Practice

1. Find a partner nearby and introduce yourself.
2. Discuss the problems on the handout and record your thoughts.