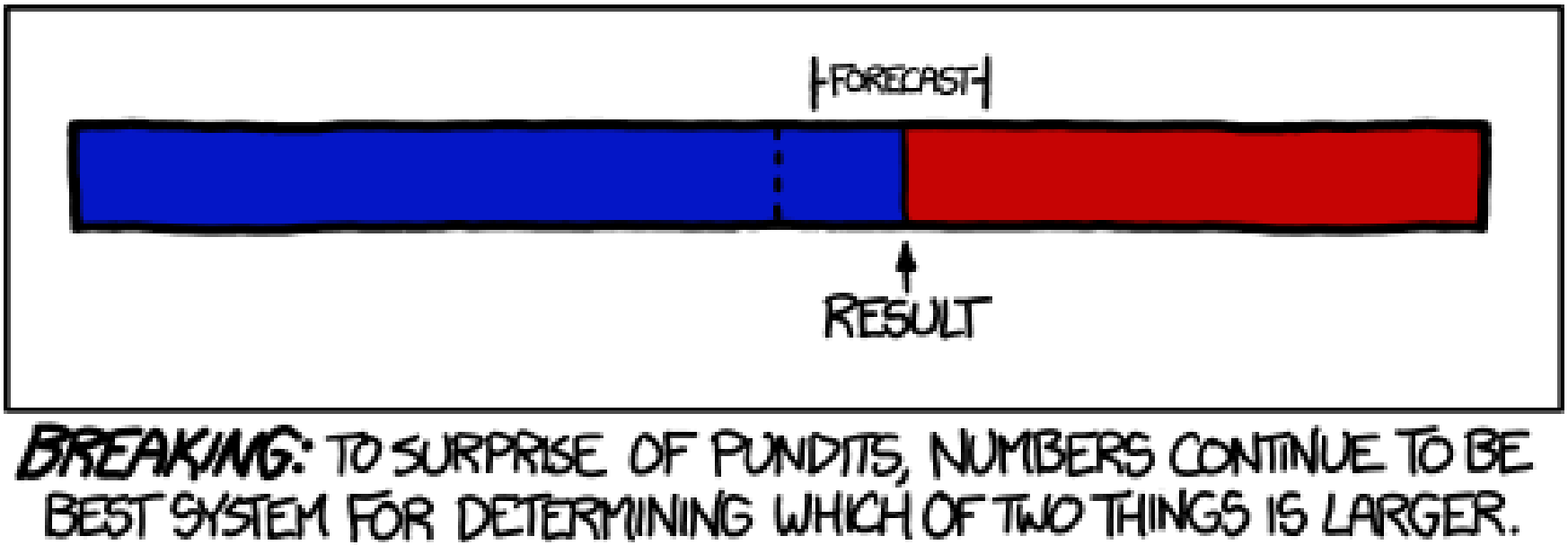
Intro to Social Science Data Analysis

Seminar 10: Comparing Proportions & Simple Linear Regression

Christopher Gandrud

November 8, 2012

A Win for R



Source: `http://xkcd.com/1131/`

Data

Load UK Smoking Data from the *openintro* package:

```
# Load package
library(openintro)

# Load Data
data(smoking)
```

Frequency Table

Create a One-way Contingency Table for Smoking Status

```
# Load package
library(MASS)

# Create Table
SmokingTable <- table(smoking$smoke)
```

Question

If 27% of South Korean's smoke, are the proportions of smokers and non-smokers different in the UK than in South Korea?

Independence

Determine if UK smokers are different than non-smokers in terms of their:

- highest educational qualification,
- nationality,
- gross income,
- ethnicity.

Independence

Determine if UK smokers are different than non-smokers in terms of their:

- highest educational qualification,
- nationality,
- gross income,
- ethnicity.

Independence

Determine if UK smokers are different than non-smokers in terms of their:

- highest educational qualification,
- nationality,
- gross income,
- ethnicity.

Independence

Determine if UK smokers are different than non-smokers in terms of their:

- highest educational qualification,
- nationality,
- gross income,
- ethnicity.

Two-way Relationships

Download data on 4 variables from the World Bank.

Graphically describe the relationships between these variables, including using linear regression lines.