

# permute: A Python Package for Randomization Inference

Kellie Ottoboni

Department of Statistics, UC Berkeley  
Berkeley Institute for Data Science

June 14, 2016



University of California, Berkeley  
**DEPARTMENT OF STATISTICS**



# Outline

## 1 Introduction

## 2 Examples

- Gender bias in teaching evaluations
- Salt and mortality at the level of nations
- Inter-rater reliability

## 3 The role of software development in Statistics

In R:

- **ri** by Peter Aronow and Cyrus Samii

*"This package provides a set of tools for conducting exact or approximate inference for randomized experiments of arbitrary design. The primary functionality of the package is in the generation, manipulation and use of permutation matrices implied by given experimental designs..."*

- **RIttools** by Mark Fredrickson

*"The RIttools package implements useful functions for implementing randomization inference based statistical tests. The package provides tools for testing balance of observed covariates in observational studies using the methodology of:...The package also provides outcome analysis of simple or block randomized trials (or matched observational studies) based on user defined models and test statistics."*

- **coin** by Torsten Hothorn, Kurt Hornik, Mark A. van de Wiel, and Achim Zeileis

*The R package coin implements a unified approach to permutation tests providing a huge class of independence tests for nominal, ordered, numeric, and censored data as well as multivariate data at mixed scales. Based on a rich and flexible conceptual framework that embeds different permutation test procedures into a common theory, a computational framework is established in coin that likewise embeds the corresponding R functionality in a common S4 class structure with associated generic functions.*

- **perm** by Michael Fay

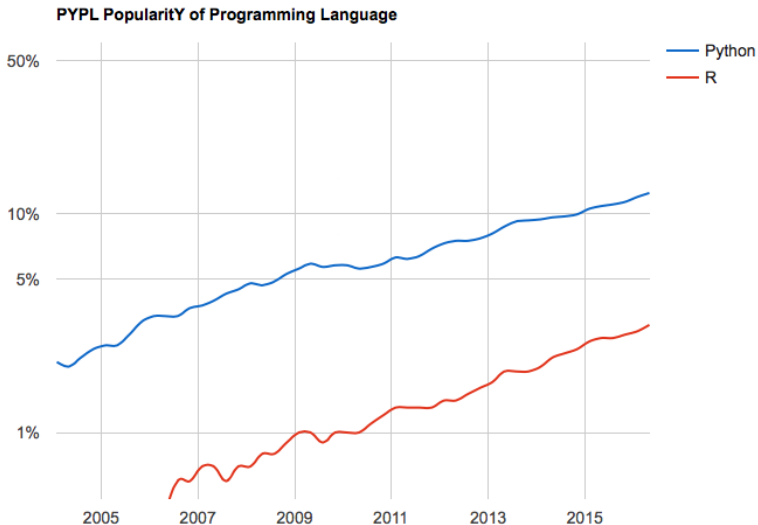
*The package has three main functions, to perform linear permutation tests. These tests are tests where the test statistic is the sum of the product of a covariate (usually group indicator) and the scores.*

In Python, statistics packages are limited.

- `numpy.random` generates random variables from common distributions
- `scipy.stats` computes moments, evaluates distribution functions, and generates random variables from common distributions and does common tests
- `StatsModels` is a Python module that provides classes and functions for the estimation of many different statistical models, as well as for conducting statistical tests, and statistical data exploration.
- `scikit-learn` is a module for machine learning

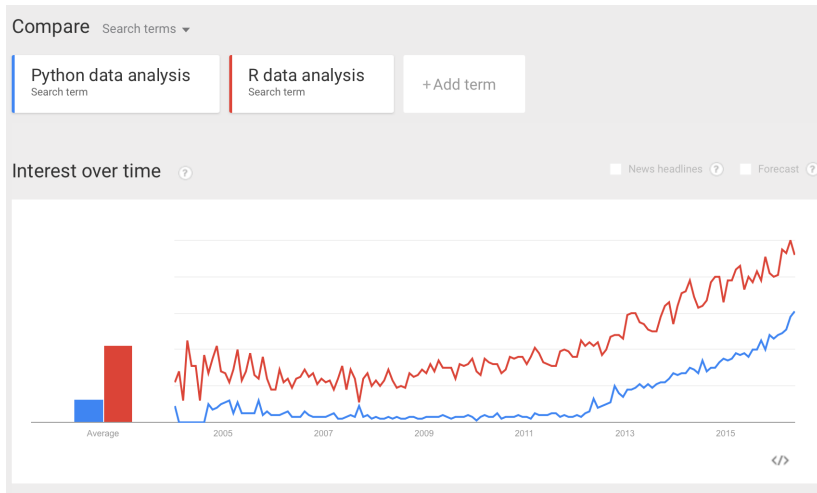
# Python is gaining popularity for doing data analysis

PYPL Popularity of Programming Language Index, Worldwide



# Python is gaining popularity for doing data analysis

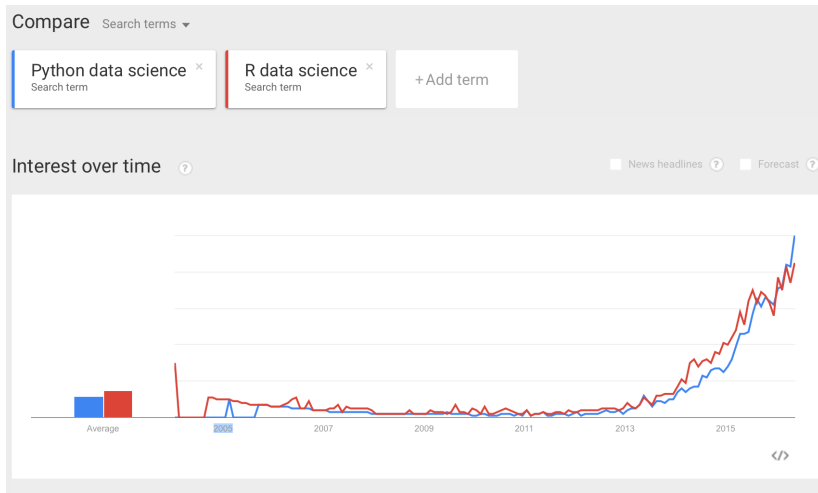
Google trends on May 22, 2016



Keyword: data analysis

# Python is gaining popularity for doing data analysis

Google trends on May 22, 2016



Keyword: data science

# Outline

## 1 Introduction

## 2 Examples

- Gender bias in teaching evaluations
- Salt and mortality at the level of nations
- Inter-rater reliability

## 3 The role of software development in Statistics



Data comes from MacNell et al. [2014]

salt and mortality

NSGK IRR stuff Millman et al. [2016]

# Outline

## 1 Introduction

## 2 Examples

- Gender bias in teaching evaluations
- Salt and mortality at the level of nations
- Inter-rater reliability

## 3 The role of software development in Statistics

## Reproducibility crisis:

- Why Most Published Research Findings Are False (Ioannidis, 2005)
- 30–50% **TO DO:** of studies fail to replicate (**TO DO: CITE**)

## Why?

- File drawer problem
- Publication bias: positive findings are more likely to get published
- P-hacking and trying many models before reporting one
- Inappropriate statistical tests

Randomization inference may ameliorate the last problem

# Download permute!

## ↳ Permutation tests and confidence sets

build passing coverage 99%

Permutation tests and confidence sets for a variety of nonparametric testing and estimation problems, for a variety of randomization designs.

- **Website (including documentation):** <http://statlab.github.io/permute>
- **Mailing list:** <http://groups.google.com/group/permute>
- **Source:** <https://github.com/statlab/permute>
- **Bug reports:** <https://github.com/statlab/permute/issues>

## Installation from binaries

```
$ pip install permute
```

<https://github.com/statlab/permute>

# Collaborators



**Jarrod Millman**  
jarrodmillman



**Philip B. Stark**  
pbstark



**Stefan van der  
Walt**  
stefanv



# References

- L. MacNell, A. Driscoll, and A. N. Hunt. What's in a name: Exposing gender bias in student ratings of teaching. *Innovative Higher Education*, pages 1–13, 2014.
- K. J. Millman, P. B. Stark, K. Ottoboni, and Naomi A. P. Stark. A case study in reproducible applied statistics: Is tagging of therapist-patient interactions reliable? Technical report, University of California, Berkeley, 2016.  
URL <https://github.com/statlab/nsgk>.