

# Knowing when and where to get help

---

Jeff Leek

@jtleek

[www.jtleek.com](http://www.jtleek.com)

**We covered a tiny fraction of statistics**



## Statistics and R for the Life Sciences

An introduction to basic statistical concepts and R programming skills necessary for analyzing data in the life sciences.



Self-Paced

[Enroll Now](#)

☒ I would like to receive email from Harvard University and learn about its other programs.

### About this course

8 Reviews 4/5 ★★★★★

We will learn the basics of statistical inference in order to understand and compute p-values and confidence intervals. We will provide examples by programming in R in a way that will help make the connection between concepts and implementation. Python and R will be used to test

[PH525.1x: Statistics and R for the Life Sciences](#)

[PH525.2x: Introduction to Linear Models and Matrix Algebra](#)

[PH525.3x: Advanced Statistics for the Life Sciences](#)

[PH525.4x: Introduction to Bioconductor](#)

[PH525.5x: Case study: RNA-seq data analysis](#)

[PH525.6x: Case study: Variant Discovery and Genotyping](#)

[PH525.7x: Case study: ChIP-seq data analysis](#)

[PH525.8x: Case study: DNA methylation data analysis](#)

[About This Specialization](#)[Courses](#)[Pricing](#)[Creators](#)[FAQs](#)

## Data Science Specialization

From \$29

Enroll

Starts September 7, 2015

Share    

# Launch Your Career in Data Science

A nine-course introduction to data science, developed and taught by leading professors.

About This Specialization

Ask the right questions, manipulate data sets, and create

**Statistics requires lots of practice**

08  
JUN

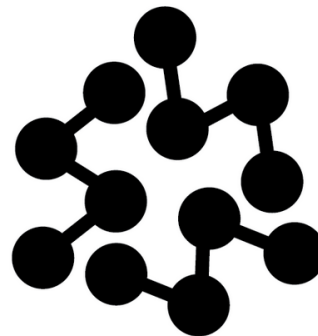
# I'm a data scientist - mind if I do surgery on your heart?

POSTED BY JEFF LEEK / UNCATEGORIZED



There has been a lot of recent interest from scientific journals and from other folks in creating checklists for data science and data analysis. The idea is that the checklist will help prevent results that won't reproduce or replicate from the literature. One analogy that I'm frequently hearing is the analogy with checklists for surgeons that **can help reduce patient mortality**.

The one major difference between checklists for surgeons and checklists I'm seeing for research purposes is the difference in credentialing between people allowed to perform surgery and people allowed to perform complex data analysis. You would never let me do surgery on you. I have no medical training at all. But I'm frequently asked to review papers that include complicated and technical data analyses, but have no trained data analysts or statisticians. The most common approach is that a postdoc or graduate student in the group is assigned to do the analysis, even if



Search..



## RECENT POSTS

[Interview with COPSS award Winner John Storey](#)

[The Next National Library of Medicine Director Can Help Define the Future of](#)

**It is ok to ask for help**  
**Locally**



JOHNS HOPKINS  
BLOOMBERG SCHOOL  
of PUBLIC HEALTH

# BIostatISTICS CENTER

Biostatistics supporting health research



[Home](#)

[News & Events](#)

[About Us](#)

[Meet Our Team](#)

[Services](#)

[Tools/Resources](#)

[Center Bibliography](#)

[Policies](#)

[Make a Gift](#)

[Contact Us](#)

## Welcome

The Johns Hopkins Biostatistics Center (JHBC) provides the latest in biostatistical and information science expertise to a wide range of clients in the Johns Hopkins Medical Institutions and beyond. We consult to health researchers at other academic health centers, pharmaceutical companies, medical research organizations, managed care organizations and government agencies. We provide a wide range of [services](#) and guidance; examples of our collaborative efforts can be found in our [Center Bibliography](#).

Our [team](#) is comprised of Masters and PhD level statisticians and computer scientists from the Johns Hopkins Department of Biostatistics and affiliated faculty from other quantitative departments. There are also opportunities for students to be involved in the Center, with training in consulting processes, biostatistical methods, and substantive areas. For statisticians who may be interested in working with us, see our [employment opportunities](#) listed under [News and Events](#).

We invite you to find out more [about us...](#)

***Questions about statistical consulting?***

**REQUEST  
A MEETING**



Our Biostatistics  
Center Team

## Upcoming Events

### **Student Walk-in Clinic**

Fridays 10:30am - 12  
noon

Room E3141

Clinic slots are limited.

*Sign-in begins at*



**It is ok to ask for help**



Questions

Tags

Users

Badges

Unanswered

Ask Question

Ten. Million. Questions. Let's celebrate [all we've done together](#).

Stack Overflow is a question and answer site for professional and enthusiast programmers. It's 100% free.

Here's how it works:



Anybody can ask a question



Anybody can answer



The best answers are voted up and rise to the top

[Take the 2-minute tour](#)

## Top Questions

interesting

405 featured

hot

week

month

0 votes  
0 answers  
1 view

### Specify the number of pattern repeats in JavaScript Regex

[javascript](#) [regex](#)

asked 1 min ago [user1613797](#) 340

0 votes  
0 answers  
7 views

### Enrolling failure

[wso2](#) [wso2-emm](#)

modified 1 min ago [user5302302](#) 1

0 votes  
0 answers  
1 view

### Typescript "Go to definition" jumps to .d.ts in Visual Studio 2015

[visual-studio](#) [typescript](#)

asked 1 min ago [Adrian Rosca](#) 532

Work on work you love.



stackoverflowcareers

Limit ▾

Sort ▾

Search

1 vote 3 answers 25 views **How to get duplicate IDs numbered?**  
plyr written 12 hours ago by b.nota • 70

0 votes 1 answer 23 views **Permutations of GWAS for GWASTools**  
gwasools written 15 hours ago by chintanm • 0 • updated 11 hours ago by Stephanie M. Gogarten • 430

0 votes 0 answers 11 views **Change column width in ReportingTools**  
reportingtools written 16 hours ago by benedikt.athmer • 0

0 votes 1 answer 13 views **RnBeads: Error in as.list.default(data.source) : no method for coercing this S4 class to a vector**  
s4 s4vectors error as.vector(data) rnbeads written 18 hours ago by schasse • 0

3 votes 2 answers 52 views **limma p-value and adjusted p-value meaning**  
limma written 19 hours ago by arfranco • 60

3 votes 2 answers 68 views **SomaticSignatures mutationContext (char ',') not in lookup table**  
vranges somaticsignatures mutationcontext written 3 days ago by tesa.severson • 10 • updated 1 day ago by Julian Gehring • 950

3 votes 1 answer 40 views **edgeR Differentially Expressed Genes**  
edgeR written 1 day ago by p\_das • 0 • updated 1 day ago by Aaron Lun • 3.6k

1 vote 3 answers 40 views **Fixing download.kegg, from pathview(), which returns empty xml**  
pathview download.kegg() xml written 4 days ago by lucmoitinho • 0

0 votes 0 answers 19 views **sva and svaseq for unsupervised analysis of expression data**  
sva combat batch effect gene expression rnaseq written 1 day ago by Farshad Farshidfar • 10

## Recent...

### Replies

- [A: How to get duplicate IDs...](#) by b.nota • 70
- [C: Problem with summarizeOv...](#) by Hervé Pagès ♦♦ 9.7k
- [A: How to get duplicate IDs...](#) by Charles Berry • 280
- [C: Fixing download.kegg, fr...](#) by Luo Weijun • 1.2k
- [A: Permutations of GWAS for...](#) by Stephanie M. Gogarten • 430

### Votes

- [A: How to get duplicate IDs...](#)
- [Analyzing Open source PPI d...](#)
- [A: edgeR Differentially Exp...](#)
- [A: RDAVIDWebService Error 3...](#)
- [A: Fail to Install package ...](#)

### Awards • All »

- [Teacher ☺](#) to Aaron Lun • 3.6k
- [Teacher ☺](#) to Aaron Lun • 3.6k
- [Scholar ☺](#) to Aaron Lun • 3.6k
- [Supporter 🍀](#) to jessica.hekman • 40
- [Teacher ☺](#) to Gordon Smyth ♦ 24k
- [Scholar ☺](#) to Gordon Smyth ♦ 24k

### Locations • All »

- United Kingdom, just now
- Sydney, Australia, 30 minutes ago
- United States, 31 minutes ago
- Netherlands, 1 hour ago
- United States, 2 hours ago

 **SEQanswers Home**

User Name  ☐ Remember Me?

Password

[Register](#)

[FAQ](#)

[Community](#) ▼

[Calendar](#)

[Today's Posts](#)

[Search](#)

You are currently viewing the SEQanswers forums as a guest, which limits your access. [Click here to register now](#), and join the discussion

## » Site Navigation

- » [About SEQanswers](#)
- » [Next Gen Summaries](#)
- » **[Forums](#)**
- » [Wiki](#)
- » [Instrument Map](#)

## » Log in

User Name:

Password:

☒ Remember Me?

Not a member yet?  
[Register Now!](#)

## » Online Users: 94

1 members and 93 guests

[slefevre](#)

## » New Posts

Title, Username, & Date	Last Post	Replies	Views	Forum
<a href="#">Introducing BBNorm, a read normalization and error-correction tool</a> Brian Bushnell	Yesterday 11:48 PM by <a href="#">suzumar</a>	18	2,929	<a href="#">Bioinformatics</a>
<a href="#">RNA-seq from single cells?</a> irit	Yesterday 11:27 PM by <a href="#">Simone78</a>	147	25,489	<a href="#">Sample Prep / Library Generation</a>
<a href="#">How to use GNU parallel command for running jobs in parallel</a> Gyan_prakash	Yesterday 11:19 PM by <a href="#">Gyan_prakash</a>	0	29	<a href="#">Bioinformatics</a>
<a href="#">vcf convert to fasta, the fasta file is not an alignment file</a> wangxueqin	Yesterday 10:36 PM by <a href="#">wangxueqin</a>	1	47	<a href="#">Bioinformatics</a>
<a href="#">Yes today is holiday</a> kahorohai	Yesterday 10:35 PM by <a href="#">kahorohai</a>	0	14	<a href="#">Introductions</a>
<a href="#">Normalizing mutation counts by Gene size</a> Kotoro	Yesterday 09:11 PM by <a href="#">Kotoro</a>	1	54	<a href="#">Bioinformatics</a>

## » New Ion Torrent Machines....media roundup

Sep 01, 2015 - 10:15 AM - by [ECO](#)

Ion just lifted the embargo so all the bloggers are furiously tapping out their preview details and opinions of Thermo's new sequencers...

## » Our Sponsors

**NEXTflex™ Small RNA-Seq**  
**DOWNLOAD PROTOCOL**  
BIOO SCIENTIFIC®

**BioLabs**  
• Superior library quantitation accuracy and consistency

**ThruPLEX® Plasma-seq Kit**  
• Liquid Biopsy  
• ctDNA Analysis  
• Targeted Sequencing  
• NIPT

**Birmingham UK**  
10-11 Sept 2015