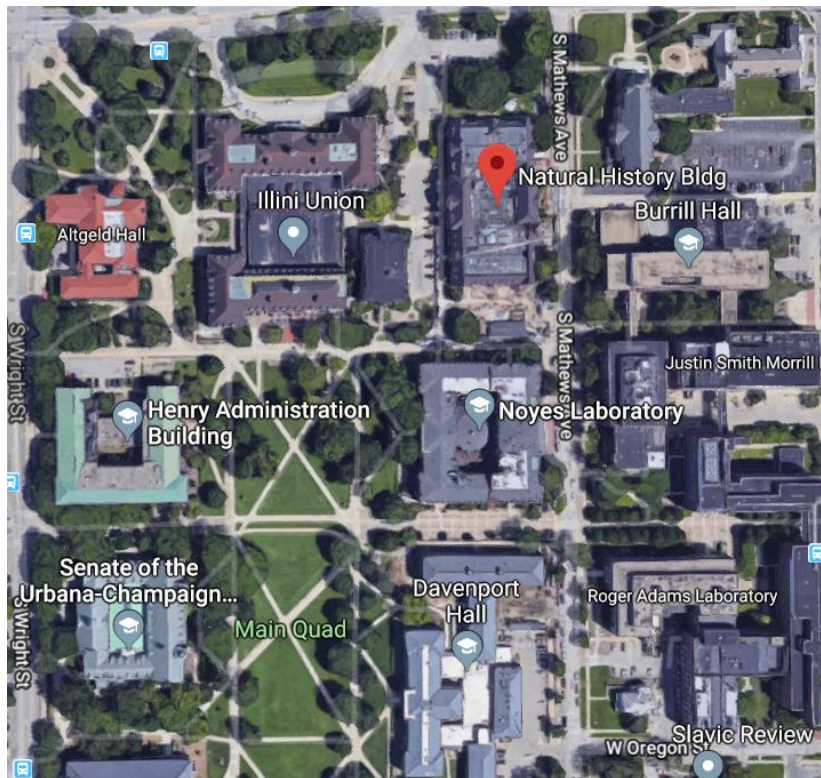


Programming for GIS



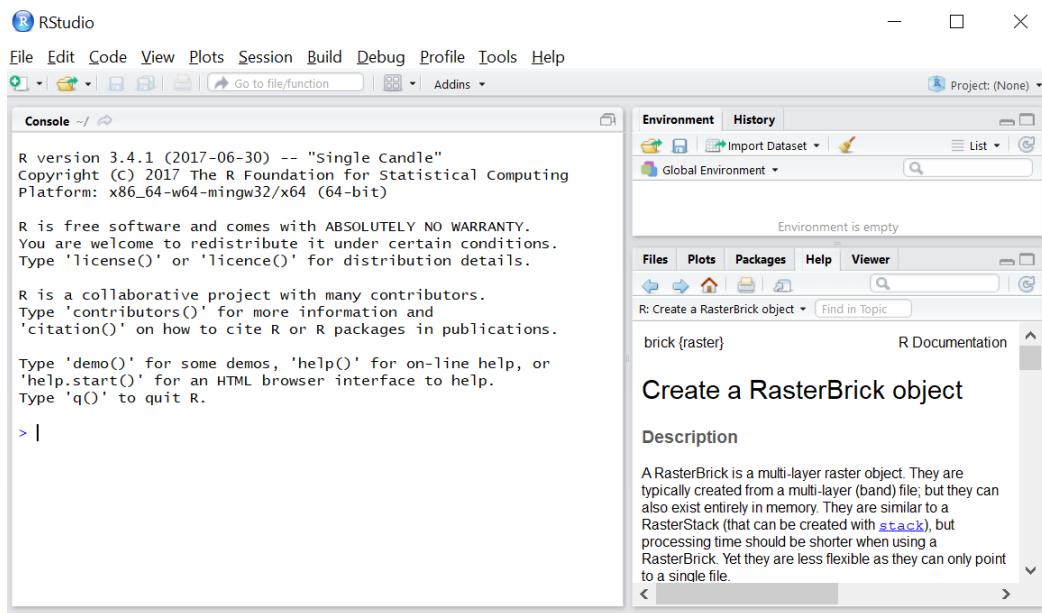
GEOG 489

Tuesday 2:00 – 3:20 PM

Thursday 2:00 – 3:20 PM

1020 Natural History Building

Programming for GIS



Instructor:

Chunyuan Diao

**2034 Natural History Building
chunyuan@Illinois.edu**

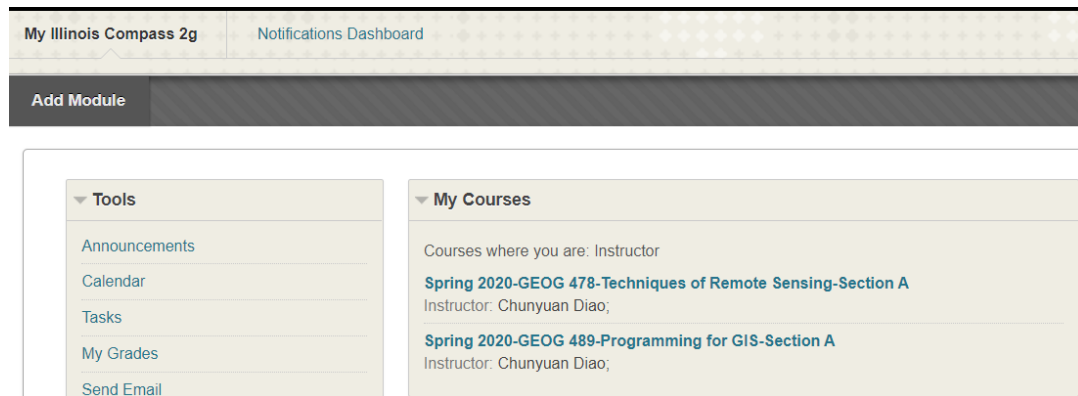
Office hours:

**Thursday 3:30 – 4:30 PM
or by appts**

Course Website

<https://compass2g.illinois.edu/>

- Login using your UIUC username/password
- Click on GEOG 489 in the “Course List” box
- I will be archiving all the course material on this site.



Course Description

- To learn the basics of programming using the R statistical computing language.
- To learn to expand GIS capabilities programmatically, including advanced techniques to import, manipulate, analyze, and export geospatial data.

Course Description

- To learn the basics of programming using the R statistical computing language.
 - Introduction, Vectors and Matrices, Debugging
 - Lists, Data Frames, Factors and Tables
 - R Programming Structures
 - Math, Simulations, and Object Oriented Programming
 - Basic I/O, String Manipulation, and Basic Graphics
 - Code Performance, Parallel Code, Introduction to IDEs

Course Description

- To learn to expand GIS capabilities programmatically, including advanced techniques to import, manipulate, analyze, and export geospatial data.
 - Introduction to GIS in R, Spatial Classes, Visualizing Spatial Data
 - Spatial Data I/O, Basic Vector Geoprocessing, Customizing Spatial Classes and Methods
 - Vector analysis
 - Raster Analyses 1: Basic Raster Math and Parallel Code
 - Raster Analyses 2: Remote Sensing
 - Raster Analyses 3: GLM and GAMs
 - Raster Analyses 4: Local Window Analysis

Prerequisite

- The course is open to graduate students or senior undergraduate students.
- Familiarity with basic GIS concepts (raster, vector, geographic projection, etc.) will be assumed, but no prior coding experience is required.

Course Structure

- The course will focus on programming in the R language.
- Typical class sessions will consist of a short (<30 minute) lecture followed by interactive exercises and activities.
- All teaching and exercises are done from within RStudio (<https://www.rstudio.com/>).

Course Grading

- Class attendance/class quiz: 10%
- Package introduction: 10%
- Assignments: 35%
- Final project: 45%

Course Grading

- Class attendance/class quiz: 10%

Class participation/attendance is **mandatory**. There will be 10 class quizzes to take attendance, each worth 1% of your grade (1 x 10% = 10% of your grade)

Missing a class quiz without an Absence Letter will result in a loss of 1% of your grade.

You can get an Absence Letter from the Emergency Dean:
<https://odos.illinois.edu/community-of-care/resources/students/absence-letters/>

Course Grading

- Package introduction: 10%

Each student is expected to introduce a package (or two) that is relevant to your research interests during a class session

The objectives are:

- Learn how to find/download/install a new package and learn how to use it
- Teach your peers about existing R packages that may be useful in their research

Course Grading

- Assignments: 35%
- We will have 5 graded assignments worth 7% each during the semester ($5 \times 7\% = 35\%$ of your grade).
- Typically you will have an entire week to complete the assignment. Late assignments will lose 1% per day they are late (each day ends at 5pm), so after 7 days late you can't get credit for them.

Course Grading

- Final project: 45%

Each student will plan and execute a start-to-finish research project using the R programming language

- Begin by becoming familiarized with the R programming language and its advanced GIS capabilities.
- Formulate a scientific question that can be answered using the R programming language.
- Design the R programming workflow to answer your question, and preprocess the data properly.

Course Grading

- Final project: 45%
 - Final project proposal: 10%
 - Final project presentation: 10%
 - Final project paper: 25%

Course Grading

- Final project: 45%
 1. Title (<25 words)
 2. Introduction [~ 200 words]
 3. Materials and methods [~ 200 words]
 - 1). Narrative
 - 2). Code
 - 3). Data
 4. Results [~200 words]
 5. Conclusions [~200 words]
 6. References

Course Grading

A	93.33-100.0%
A-	90.00-93.32%
B+	86.67-89.99%
B	83.33-86.66%
B-	80.00-83.32%
C+	76.67-79.99%
C	73.33-76.66%
C-	70.00-73.32%
D+	66.67-69.99%
D	60.00-66.66%
F	<60%

If the mean final percentage for the class is under 75%, I will adjust the final grade such that the mean score is 75%. The curve will be calculated separately for undergraduates and graduates.

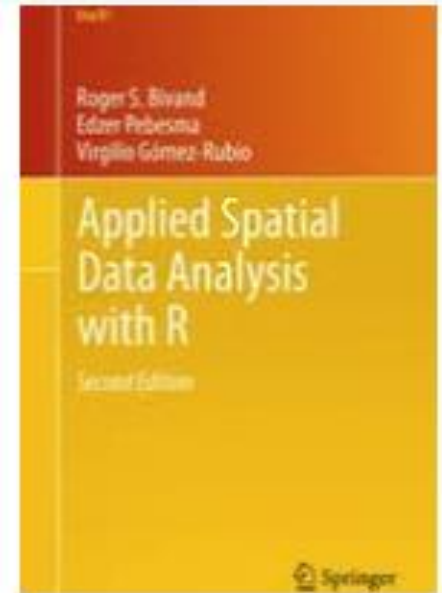
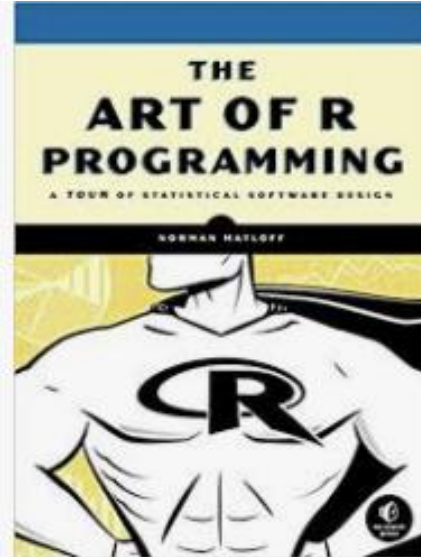
Email Policy

- During the week, I will attempt to respond to emails within 48 hours of receiving them (not including weekends). Do not expect an immediate response (please plan accordingly).
- For example, do not send an email with a question about a homework the same day that the homework is due.
- If you send an email over the weekend, do not expect any response until Monday or Tuesday.

Textbook

Textbooks (optional, but highly recommended)

- Matloff, N. (2011). *The Art of R Programming* . No Starch Press.
- Bivand, R. S., Pebesma, E. J., & GómezRubio, V. (2013). *Applied spatial data analysis with R, Second Edition* . Springer.



Course Policies

- Class participation is **mandatory**. Cell phones must be silenced and may not be used for ANY reason during the lecture.
- Errors in grading: if you believe an error was made, you must submit via email to the instructor a clear description of what grade was given, and why it was in error. The instructor will NOT discuss grades without this written explanation.

Course Policies

- This course, as with all courses at UIUC, expects the students to abide by the “Academic Integrity and Procedure” guidelines:

http://admin.illinois.edu/policy/code/article1_part4_1-401.html

- Ignorance of these rules is not an excuse to cheat!
- **Important**: Labs may be discussed with your classmates, but the writing of the assignment should be yours, and yours alone. We are checking for identical (or nearly identical) answers, so be aware.

Tentative Schedule

I. Basics of the R programming languages

1/20 Introduction, Vectors and Matrices, Debugging

1/27 Lists, Data Frames, Factors and Tables

2/3 R Programming Structures

2/10 Math, Simulations, and Object Oriented Programming

2/17 Basic I/O, String Manipulation, and Basic Graphics

2/24 Code Performance, Parallel Code, Introduction to IDEs

Tentative Schedule

II. Advance GIS capabilities using the R programming languages

3/2 Introduction to GIS in R, Spatial Classes, Visualizing Spatial Data

3/9 Spatial Data I/O, Basic Vector Geoprocessing, Customizing Spatial Classes and Methods

3/23 Vector analysis

3/30 Raster Analyses 1: Basic Raster Math and Parallel Code

4/13 Raster Analyses 2: Remote Sensing

4/20 Raster Analyses 3: GLM and GAMs

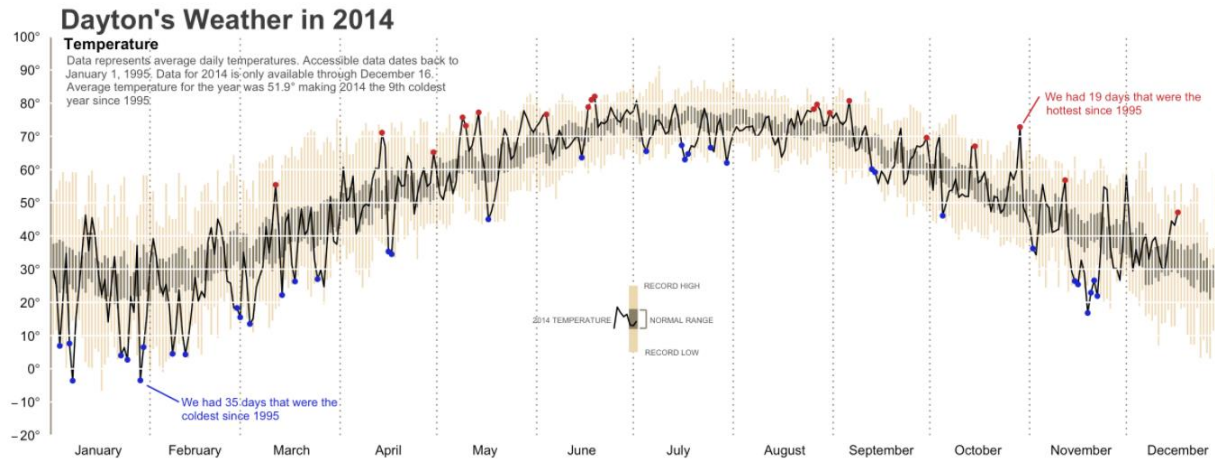
4/27 Raster Analyses 4: Local Window Analysis

Week beginning with 4/27: Final project presentations

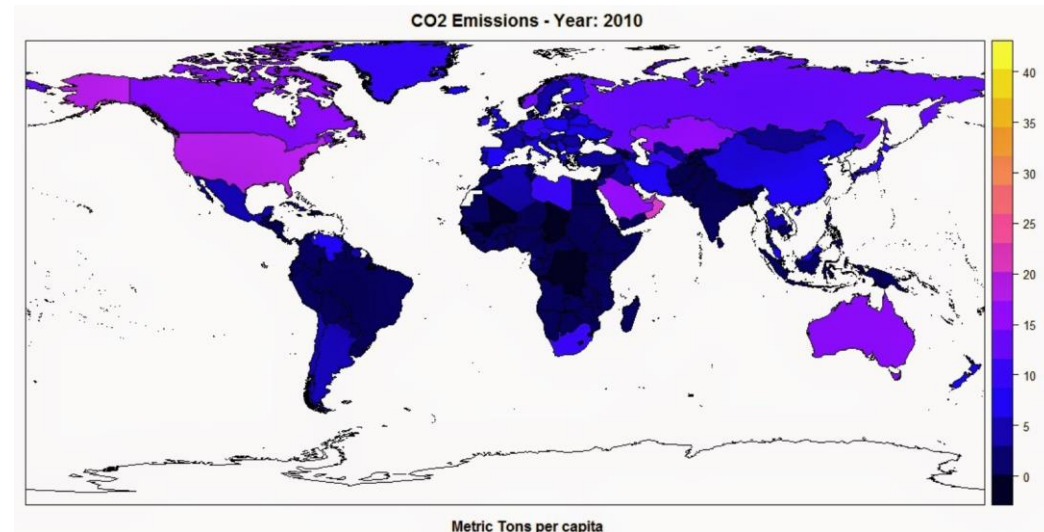
R Programming Language

- Free and Open source
- Data manipulation
- Data analysis tools
- Great graphics
- Programming language
- 10,000+ free, community-contributed packages
- A supportive and increasing user community

R Graphics



[http://rpubs.com/bradleyboehmke/
weather_graphic](http://rpubs.com/bradleyboehmke/weather_graphic)



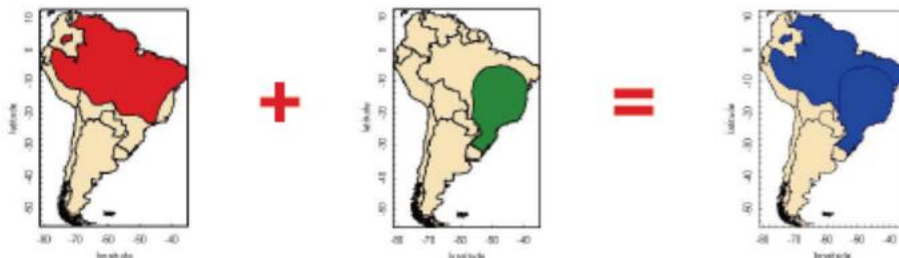
[http://r-video-tutorial.blogspot.com/2015/05/interactive-maps-for-
web-in-r.html](http://r-video-tutorial.blogspot.com/2015/05/interactive-maps-for-web-in-r.html)

Spatial Data in R

Packages: sp, maptools, rgeos, raster, ggmap

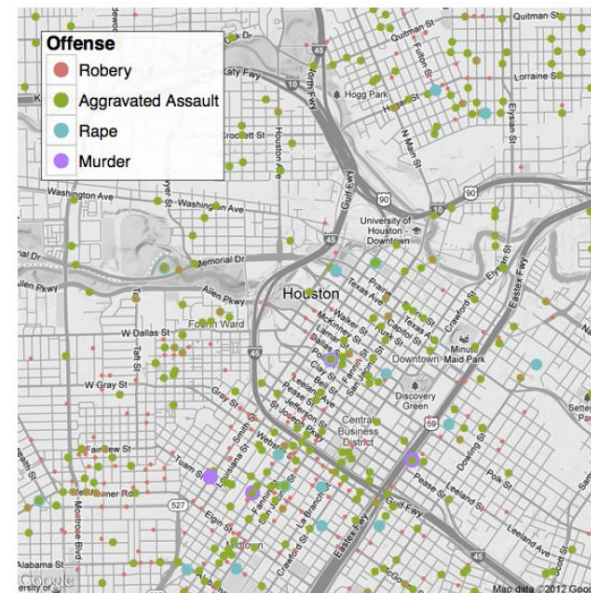
Examples:

- species range overlays



- Basemaps with ggmap

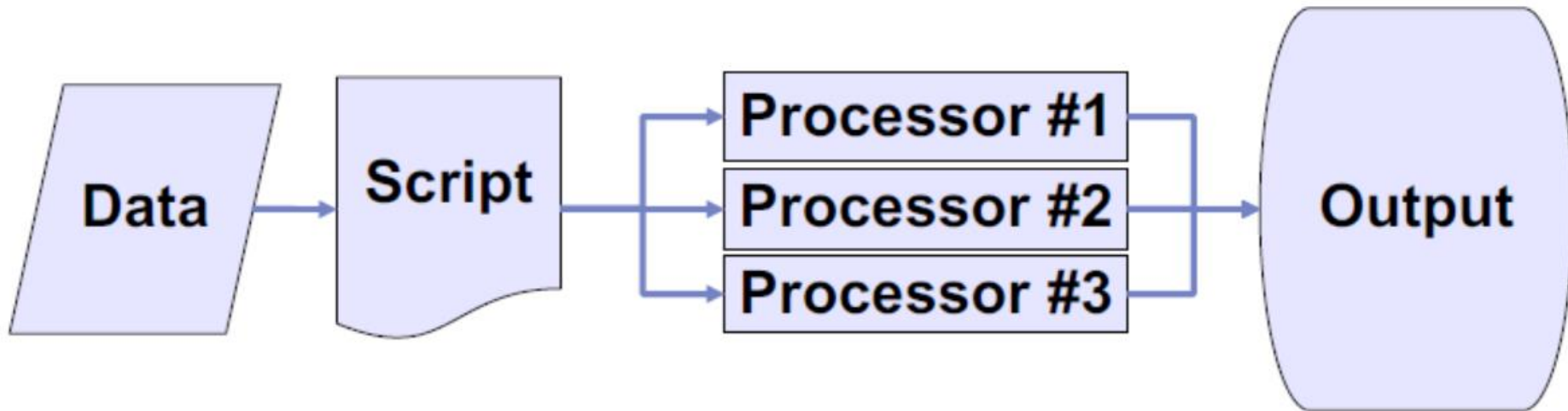
<http://www.nceas.ucsb.edu/>



<http://journal.r-project.org/archive/2013-1/kahle-wickham.pdf>

Parallel Processing

- For BIG jobs: Multi-core processors/ high performance computing with foreach



R Studio

RStudio interface showing a script, console, and a map plot.

Script (map.R):

```
43  
44  
45 map <- get_map(location='India', zoom=4)  
46  
47 ggmap(map) +  
48   geom_point(data=kim.points,  
49             aes(x=kim.points$longitude, y=kim.points$latitude), col="#ff000099", size=2)  
50  
51  
52 # This has the disadvantage of not letting us as easily see which points are more significant.  
52:95 (Top Level) ↕
```

Console:

```
> cannot open: HTTP status was '0 (null)'  
>  
> ggmap(map) +  
+   geom_point(data=kim.points, position=position_jitter(width=1, height=1),  
+             aes(x=kim.points$longitude, y=kim.points$latitude), col="#ff000099", size=2)  
Warning message:  
Removed 3 rows containing missing values (geom_point).  
>  
>  
> map <- get_map(location='India', zoom=4)  
Map from URL : http://maps.googleapis.com/maps/api/staticmap?  
center=India&zoom=4&size=%20640x640&scale=%202&maptype=terrain&sensor=false  
Google Maps API Terms of Service : http://developers.google.com/maps/terms  
Information from URL : http://maps.googleapis.com/maps/api/geocode/json?address=India&sensor=false  
Google Maps API Terms of Service : http://developers.google.com/maps/terms  
>  
> ggmap(map) +  
+   geom_point(data=kim.points,  
+             aes(x=kim.points$longitude, y=kim.points$latitude), col="#ff000099", size=2)  
Warning message:  
Removed 3 rows containing missing values (geom_point).  
>  
>  
> # This has the disadvantage of not letting us as easily see which points are more significant.  
>
```

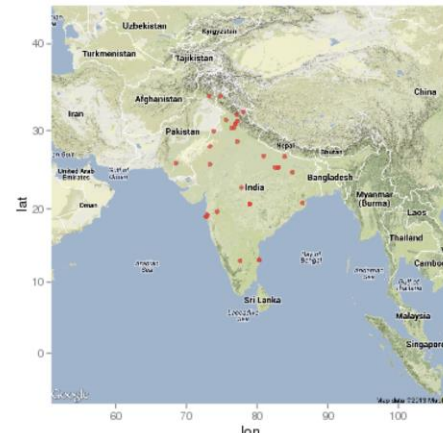
Workspace:

Data	Observations
afghanistan	77 obs. of 6 variables
india	363 obs. of 6 variables
kim.points	38 obs. of 5 variables
mdat	363 obs. of 6 variables
pakistan	126 obs. of 6 variables

Values:

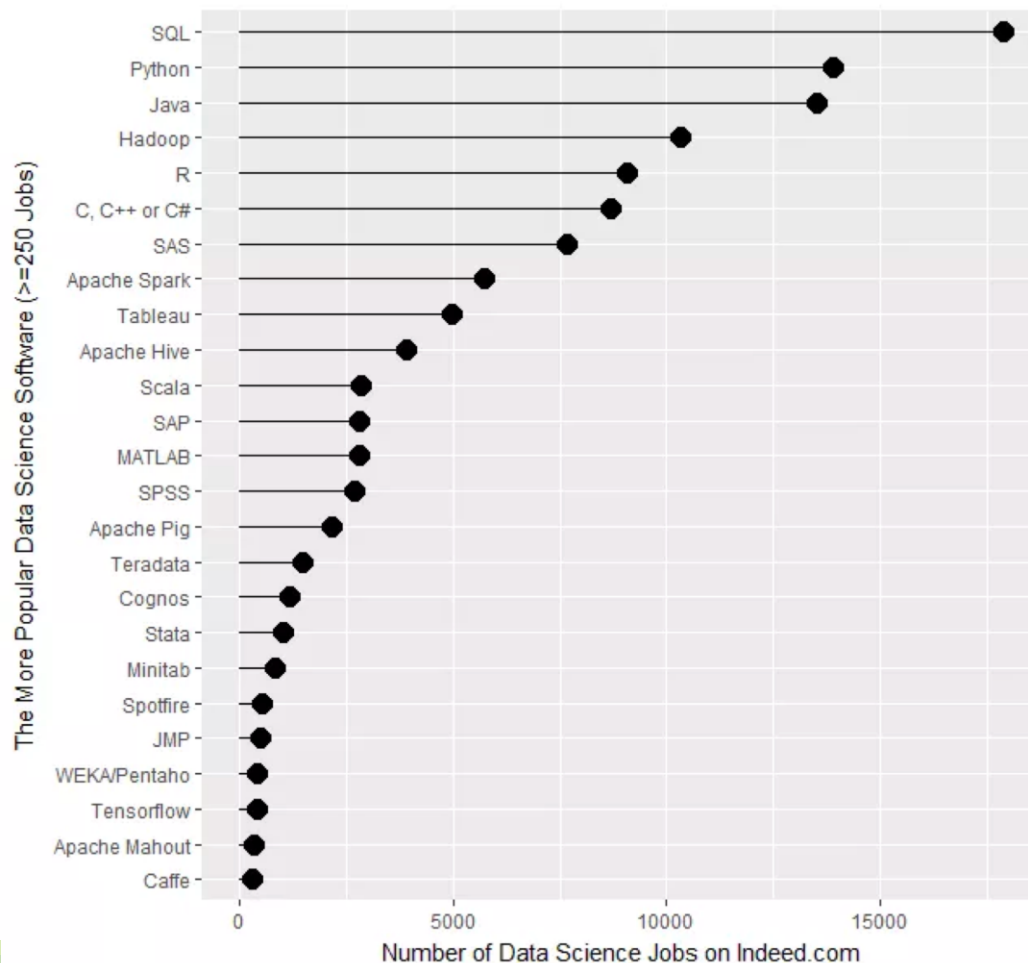
Variable	Value
map	ggmap[1638400]
names	character[2284]

Plots:



Who uses R?

- “Data Science” Jobs on indeed.com



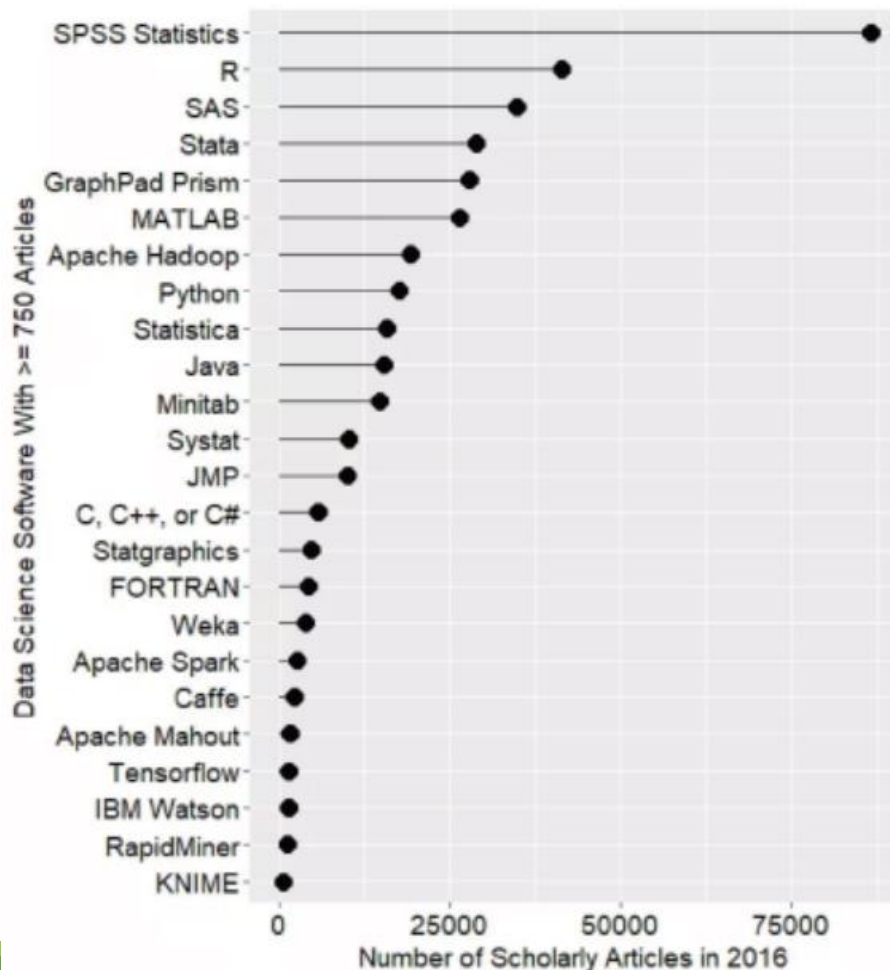
The number of data science jobs for the more popular software (those with 250 jobs or more, 2/2017).

Feb 2017:

<http://r4stats.com/articles/popularity/>

Who uses R?

- Scholarly articles by software package



Number of scholarly articles found in 2016 for each software package used as a topic or tool of analysis.

<http://r4stats.com/articles/popularity/>

240 books on R since 2000



- R is a great language to learn and it can take you far...
- But others (python, etc.) are useful too. Once you learn one, it is much faster to learn others

Roll Call

- Name?
- Undergraduate or graduate?
- Major?
- Why are you taking this class?
- Any programming and GIS background?

Installing R and RStudio

Install

1. **R** (<https://cran.revolutionanalytics.com/>)
2. **RStudio** (<https://www.rstudio.com>)