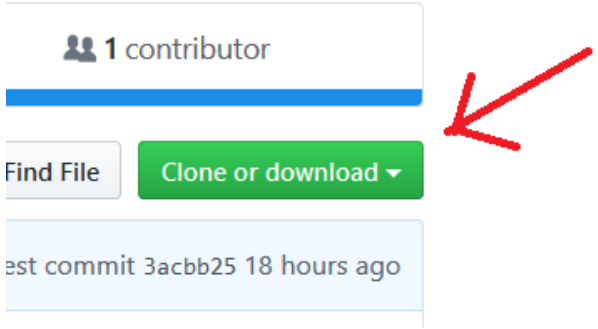
Session 1:
Introduction to R and Rstudio

Please download course materials at:
https://github.com/andrw-jns/intro_r



Andrew Jones | Ozayr Mohammed

andrew.jones40@nhs.net | ozayr.mohammed@nhs.net

Healthcare Analysts | The Strategy Unit

1-2-4

1. (Name)
2. What would be a good result today?

3:00

Our work

- Describe paediatric services across the region.
- Did such and such mental health intervention work?
- What influences demand for elective treatment?

Tools: R, Excel, SQL, QGIS….

Answers: Reports (Powerpoint and Word) and presentations

Assumptions about you

What is the nature of your work?

Reporting? Project based? Data reconciliation? Statistics? Collaboration?

Tools – Excel, SQL, Power BI? Tableau?

Answers – Reports? Presentations?

Will shape what you want from today.

Agenda

"Adjusting the seat and mirrors":

Intro + Setup + Import data

Graphics with ggplot2

Break 15 mins →

Data wrangling with dplyr

--------- Lunch ----------

← *Demo*

Naming objects | Joins

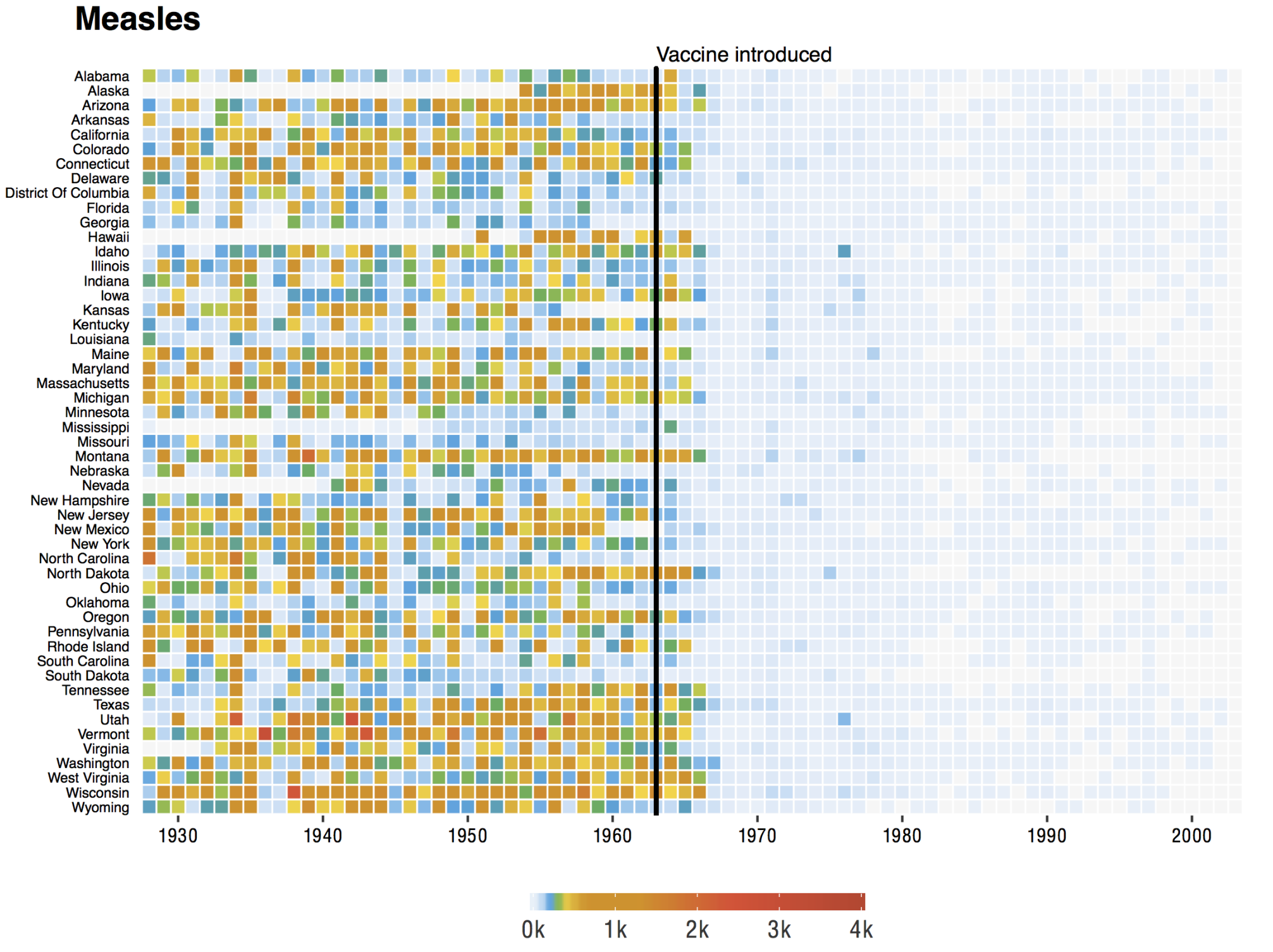
R Markdown

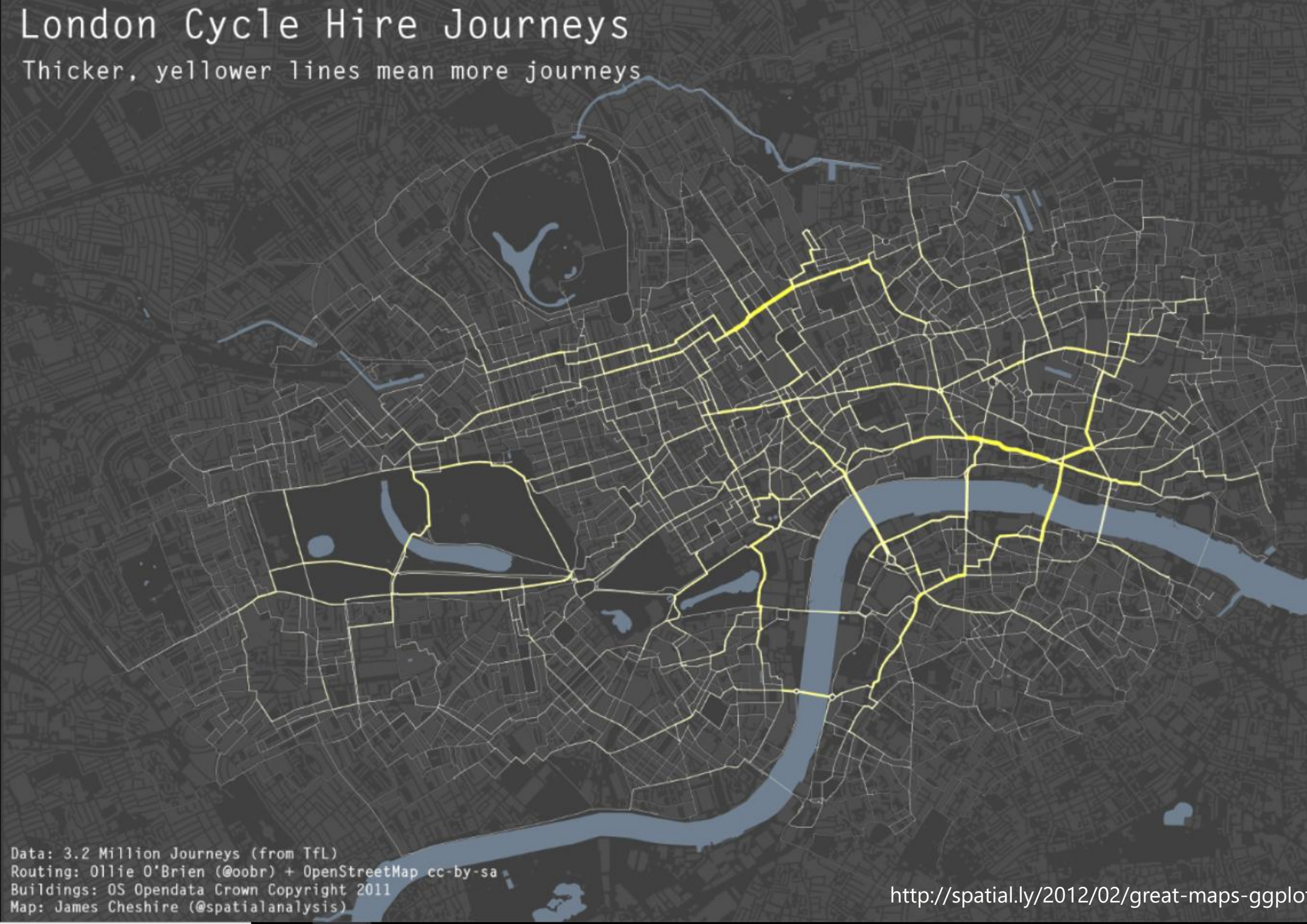
Course Aims

1. To show you some of the possibilities:

elegant

Graphics


Measles
Vaccine introduced
Alabama
Alaska
Arizona
Arkansas
California
Colorado
Connecticut
Delaware
District Of Columbia
Florida
Georgia
Hawaii
Idaho
Illinois
Indiana
Iowa
Kansas
Kentucky
Louisiana
Maine
Maryland
Massachusetts
Michigan
Minnesota
Mississippi
Missouri
Montana
Nebraska
Nevada
New Hampshire
New Jersey
New Mexico
New York
North Carolina
North Dakota
Ohio
Oklahoma
Oregon
Pennsylvania
Rhode Island
South Carolina
South Dakota
Tennessee
Texas
Utah
Vermont
Virginia
Washington
West Virginia
Wisconsin
Wyoming
1930
1940
1950
1960
1970
1980
1990
2000
0k
1k
2k
3k
4k
https://benjaminlmoore.wordpress.com/tag/ggplot2/


London Cycle Hire Journeys
Thicker, yellower lines mean more journeys
Data: 3.2 Million Journeys (from TfL)
Routing: Ollie O'Brien (@oobr) + OpenStreetMap cc-by-sa
Buildings: OS Opendata Crown Copyright 2011
Map: James Cheshire (@spatialanalysis)
http://spatial.ly/2012/02/great-maps-ggplo

Statistics
& Machine Learning
(example later)

Collaboration & Reproducibility

R Markdown

Automated reports

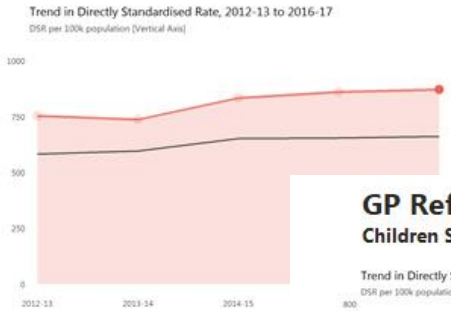
Inpatient Summary

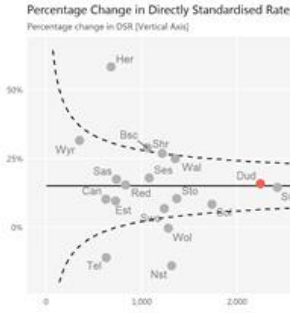
Opportunity	Admissions	2016-17 Spend (000s)	Rate	Rate of Change
ACS Acute	2,750	£3,445	High	High
ACS Chronic	2,680	£4,856	Low	-
ACS Vaccine	1,770	£5,085	High	-
Alcohol (wholly)	1,220	£1,724	High	High
Alcohol (partially - chronic)	4,970	£7,775	High	Low
Alcohol (partially - acute)	1,930	£2,506	High	High
End of Life Care (3-14 days)	320	£1,030	High	-
End of Life Care (0-2 days)	200	£305	High	-
Falls	3,620	£8,873	High	-
Frail Elderly (occasional)	1,090	£2,349	High	-
Frail Elderly (usual)				
Medically Unexplained				
Medicines - Explicit				
Medicines - Implicit AntiDiab				
Medicines - Implicit Benzo				
Medicines - Implicit Diuretics				
Medicines - Implicit NSAIDs				
Obesity (largely)				
Obesity (marginal)				
Obesity (somewhat)				
PLCV Cosmetic				
PLCV Alternative				
PLCV Ineffective				
PLCV Risks				
Mental Health Admissions from ED				
Self-harm				
Smoking (large)				
Smoking (somewhat)				
Zero Length of Stay (adult)				
Zero Length of Stay (child)				

Notes: Rate and rate of change comp

Frail Elderly Admissions

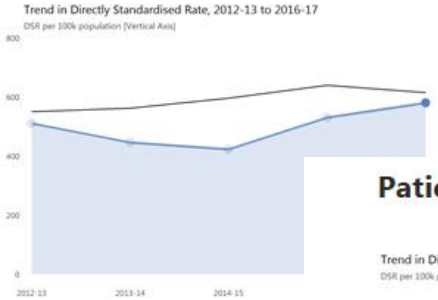
Could Usually be Managed in a Non-Acute Setting



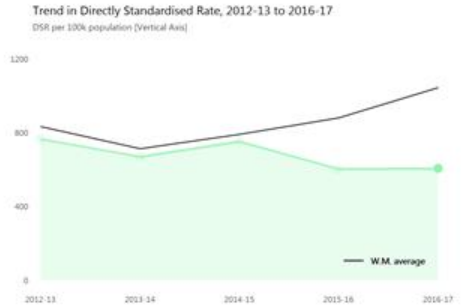


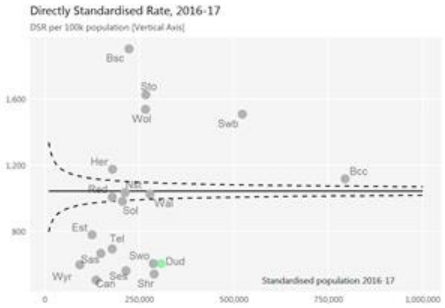
GP Referred First Outpatient Attendances

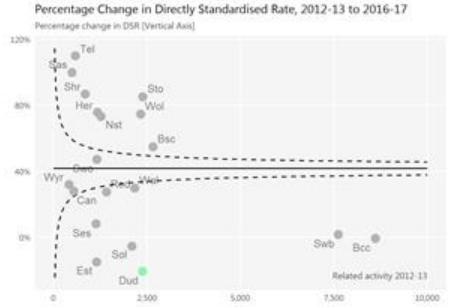
Children Surgical Specialties



Patients Leaving ED Before Being Seen



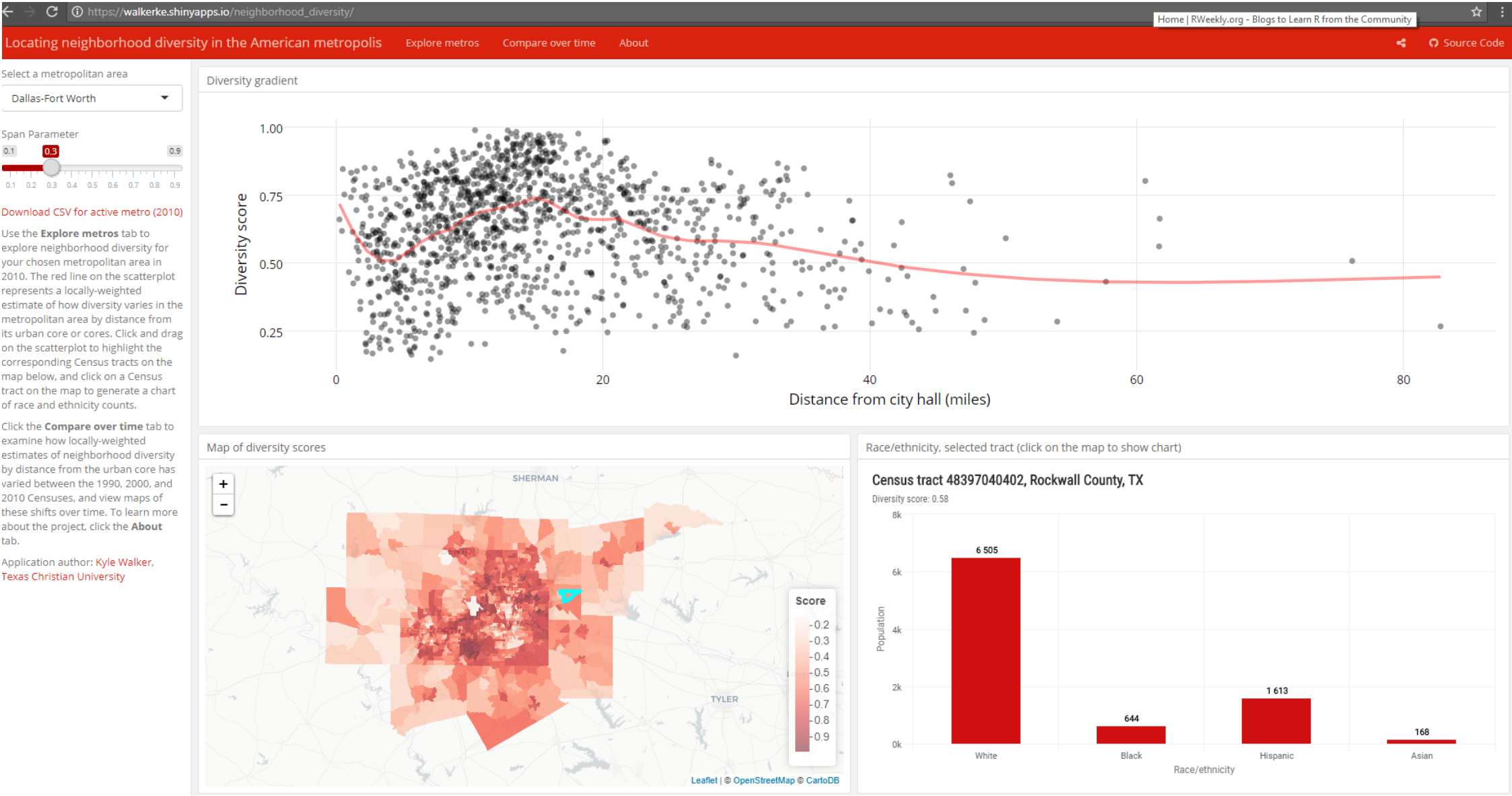




1,860 attendances	**£0.1M** spent
2.1% of all attendances	**£70** per unit activity

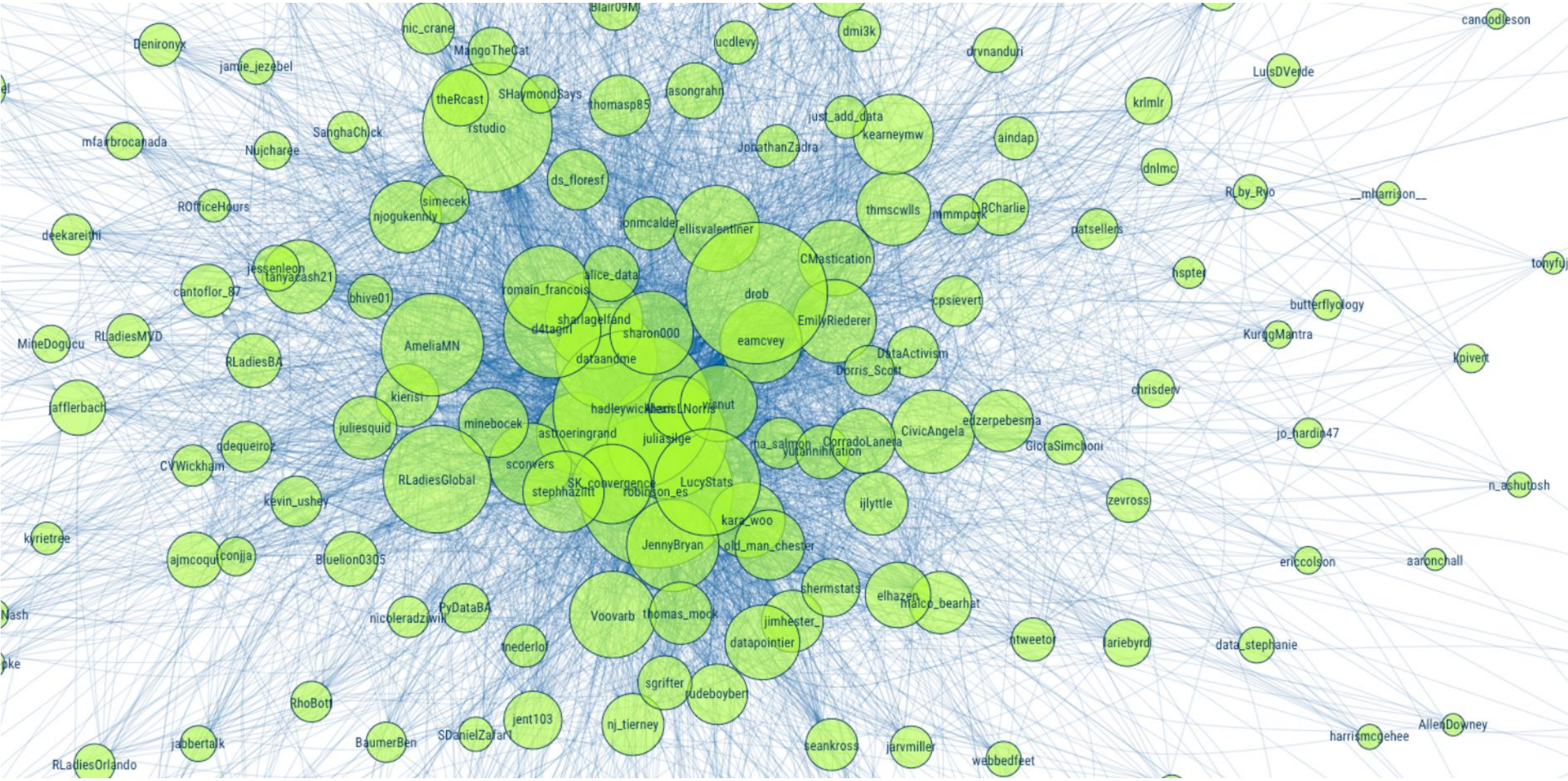
1800 unique graphics
100 unique tables
built into a
PowerPoint
from R

(Interactive) Dashboards



https://walkerke.shinyapps.io/neighborhood_diversity/

R to SQL connection



Community

https://github.com/mkearney/rstudioconf_tweets

stackoverflow Questions Developer Jobs Tags Users Search...

I have a data.frame (df) with 17 rows and 40 columns. I would like to plot all those columns like this:

0

```
windows()
plot(NULL,xlim=c(0,17),ylim=c(5000,90000),xaxt='n',xlab="", ylab="")
points(df$c1,type="b",pch=15,col="gold3")
points(df$c2,type="b",pch=15,col="gold3")
.
.
points(df$c40,type="b",pch=15,col="gold3")
```

I would like to create a loop inside the plot to not have to write all the lines for the 40 columns. I tried different things without success. Thanks in advance!

1 Answer

active oldest votes

Here is an example using standard `plot` and `points` as well as a `ggplot2` example.

This answer is useful

```
ta.frame(x=1:10,
       y1=rnorm(10),
       y2=rnorm(10),
       y3=rnorm(10))
plot(df$x, df$y1)
# points(df$x, df$y2)
# points(df$x, df$y3)
for(i in 3:4) {
  points(df$x, df[[i]])
}


library(reshape2)
library(ggplot2)
melt_df <- melt(df, 'x')
ggplot(melt_df, aes(x, value)) +
  geom_point()
```

share improve this answer

answered 1 hour ago

drmariod

3 503 ● 19 ● 53

Inclusivity

Home Moments Search Twitter

Tweets 3,551 | Following 2,410 | Followers 4,957 | Likes 4,959 | Lists 3

R-Ladies Global
@RLadiesGlobal

Promoting gender diversity in the #rstats community via meetups, mentorship & global collaboration! 60+ groups worldwide. #RLadies

The World
rladies.org
Joined August 2016
237 Photos and videos

Tweets | Tweets & replies | Media

Pinned Tweet

R-Ladies Global @RLadiesGlobal · 30 Aug 2016
Interested in starting an #RLadies meetup in your city? We'd love to help! Send us an email at info@rladies.org to get in touch. #rstats

17 80 117

R-Ladies Global Retweeted

Forwards @R_Forwards · 3h
Our advice to a black woman who has told her mentor "the tech industry is not for people like me" - what would you say?

Course Aims

1. To show you some of the possibilities.
2. To give you a feel for how R works.

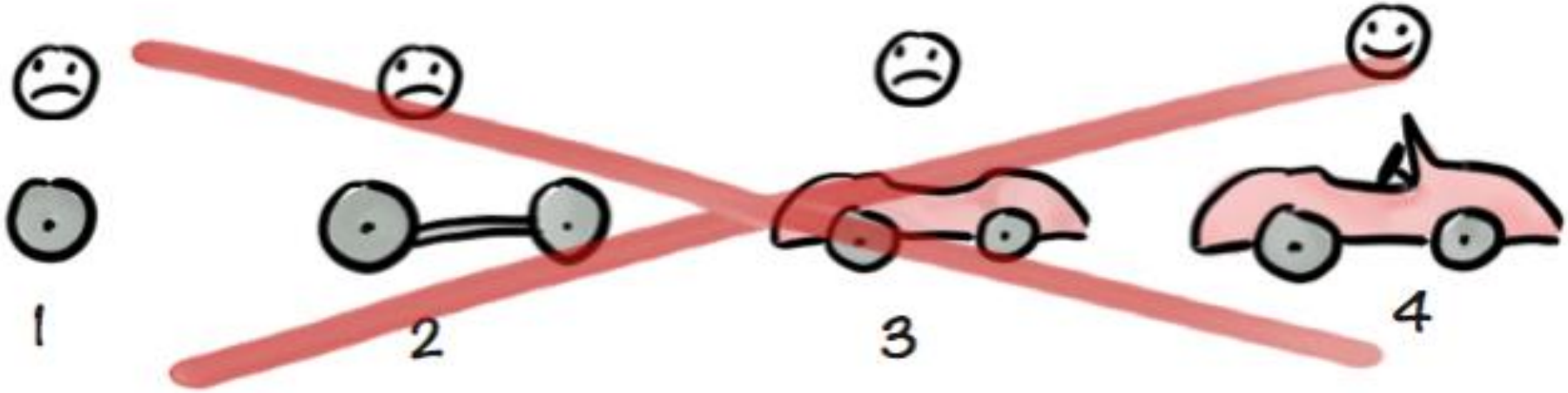
Course Aims

1. To show you some of the possibilities.
2. To give you a feel for how R works.
3. To show you enough for you to begin teaching yourself

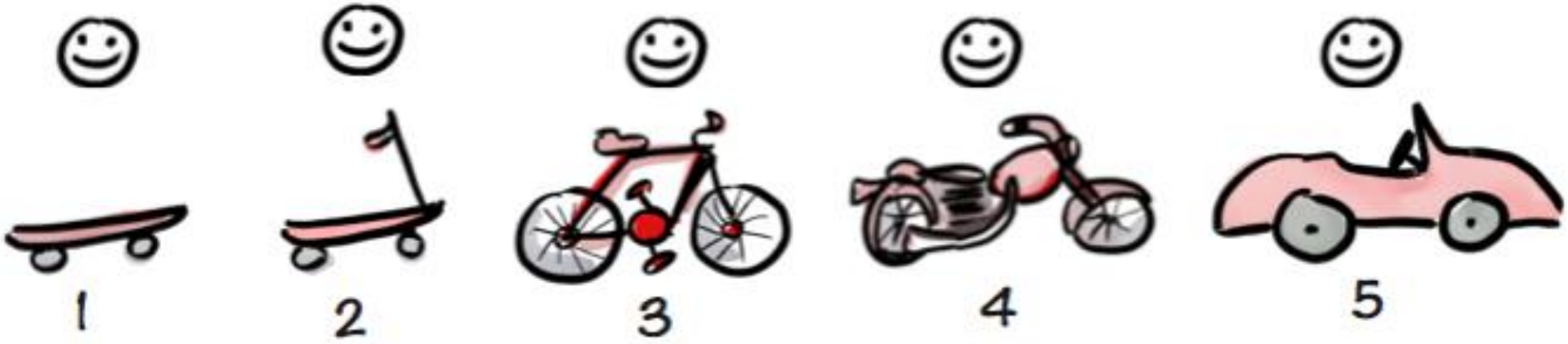
(Excellent free resources available)

Course philosophy

Not like this....



Like this!



Henrik Kniberg

https://blog.crisp.se/2016/01/25/henrikkniberg/making-sense-of-mvp

Course philosophy

Relaxed and informal.

(Slides and code are available online)

The truth; but it can’t be the whole truth…

too much to cover in a day).

Let’s begin…

R *vs.* RStudio

R is a programming language

RStudio is a software application with tools to improve your programming experience

Credit: http://www.rollcentre.com/project--onuk-sazan-

R


RStudio
Photo by Roberto Nickson (@g) on Unsplash

RStudio

Many excellent features to help you with your analyses.

Never again have to think about R and RStudio as separate:
Opening R-Studio opens an R session.

Analogy from the book Modern Dive: www.moderndive.com

Open RStudio

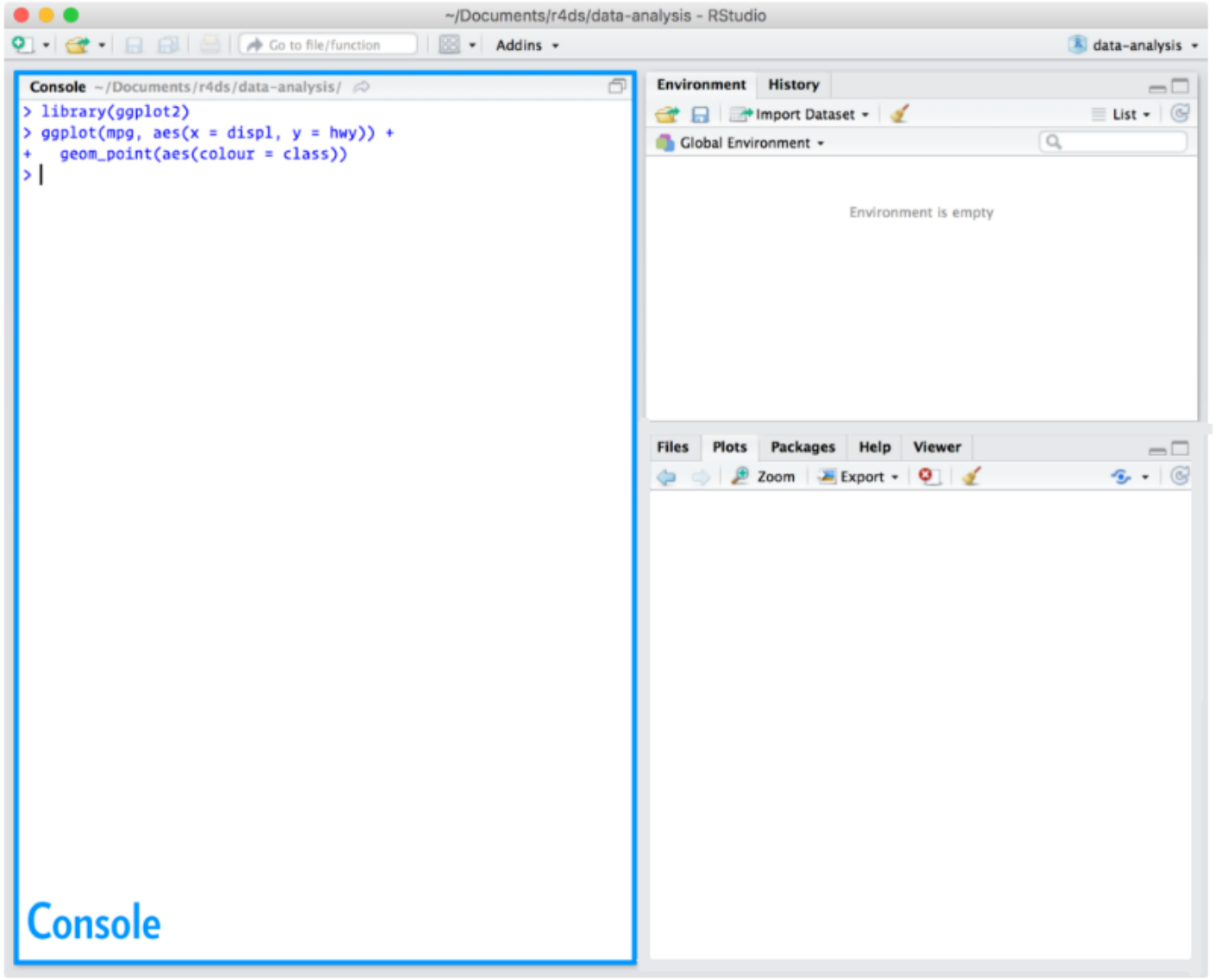
Opening R-Studio opens an R session.

The **Console** is your window to R.

You can code directly in the console...

```
pi*2      [Enter]
37/12     [Enter]
```

... but there is a better way...

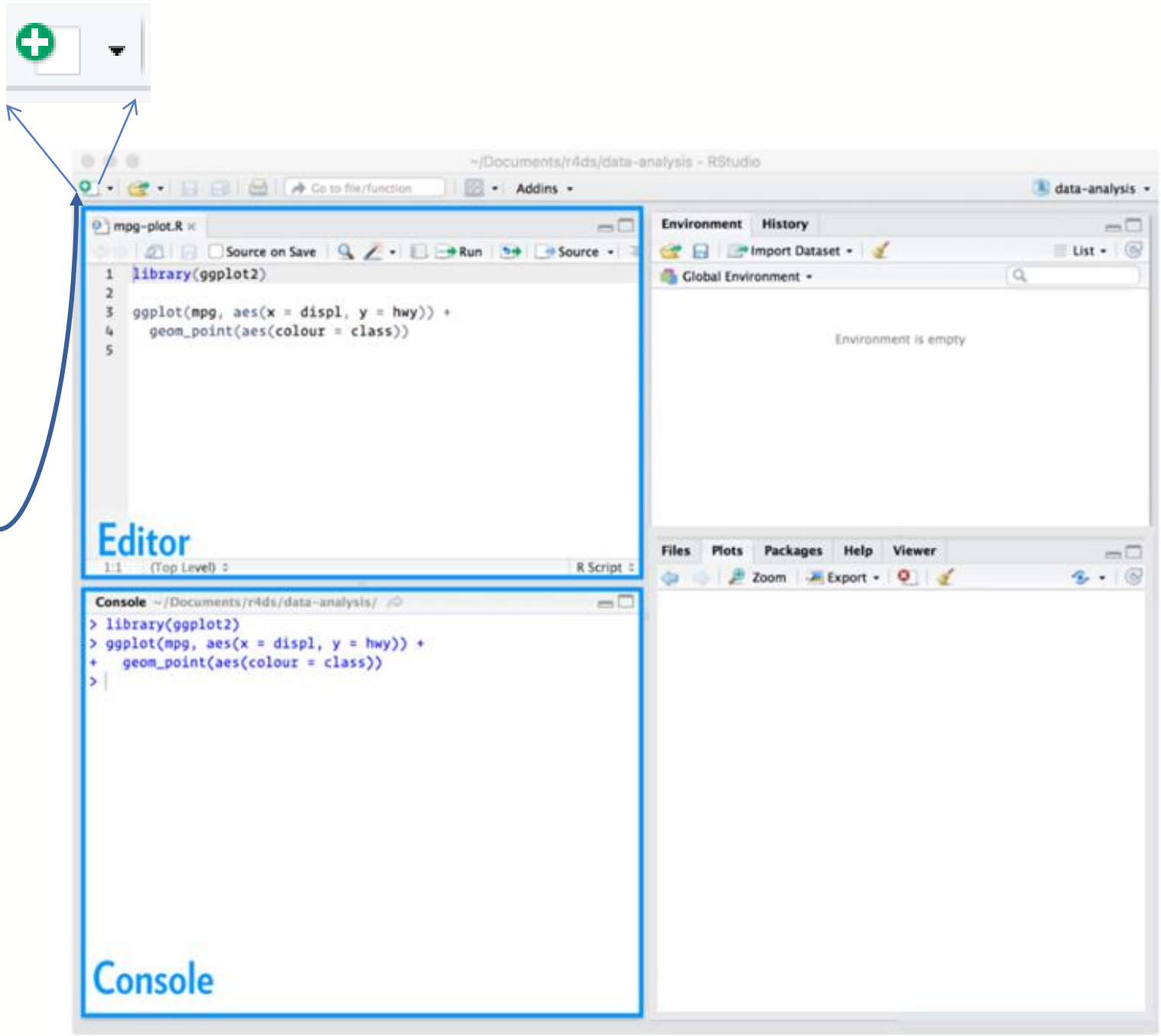


The Editor.

If you don't see the **Editor** pane, click:

And choose "R Script" from the drop-down.

Or, shortcut: Ctrl+ Shift+ N



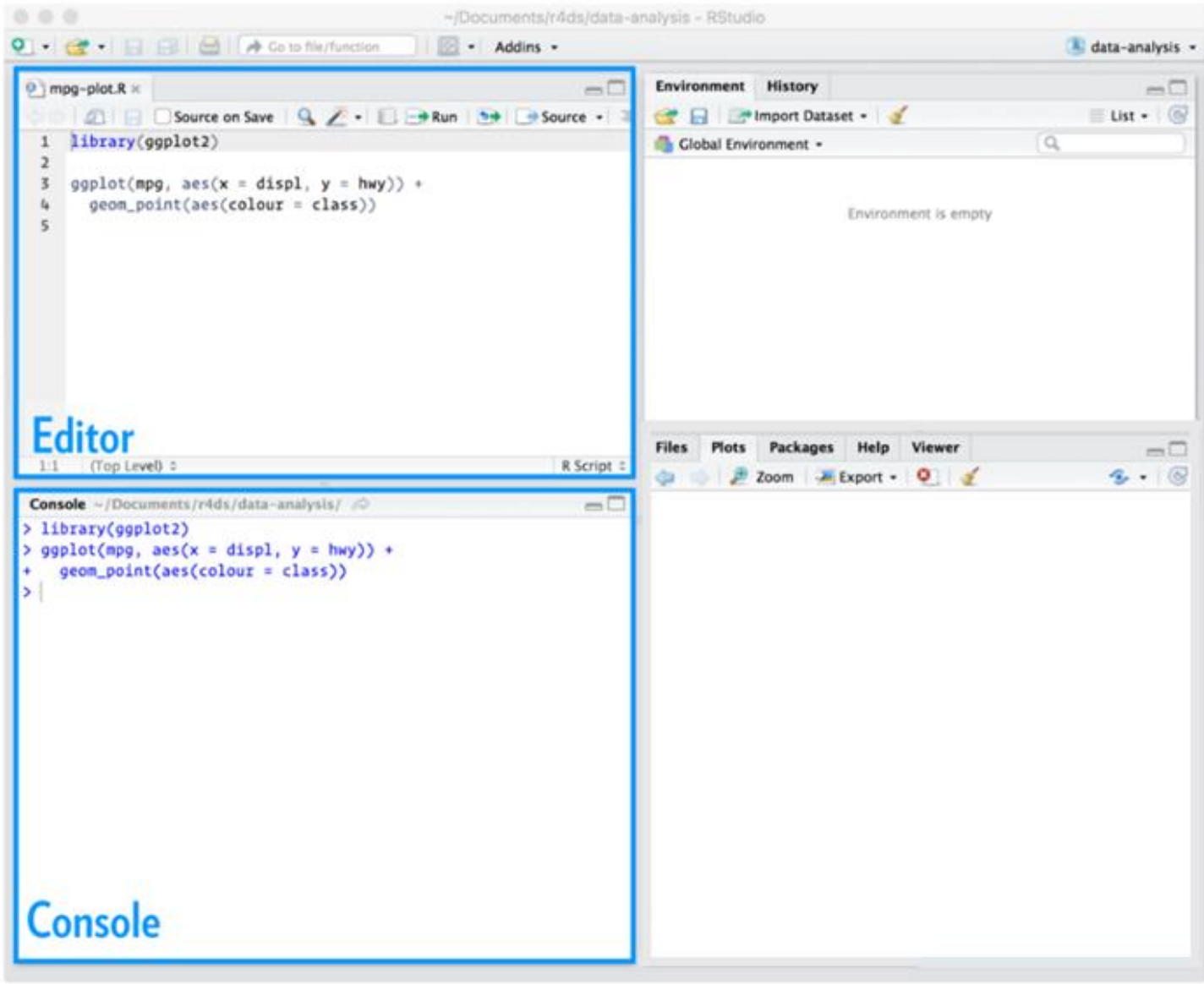
The **Editor** is just like any other text editor: you can copy, paste, and save text.

More forgiving

R Syntax highlighting

Autocomplete
(use it!)

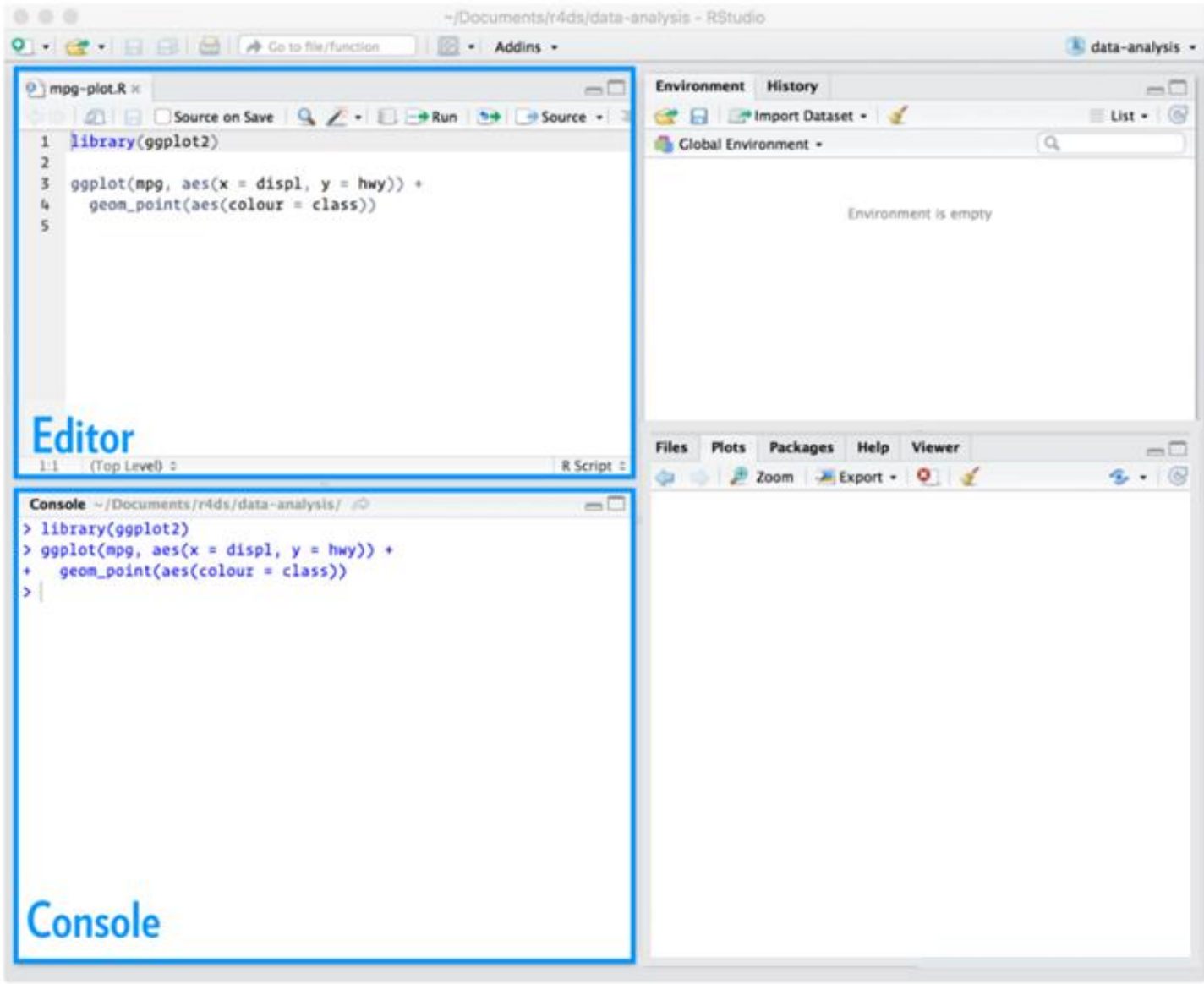
Ctrl + Enter
(sends line of code to Console)



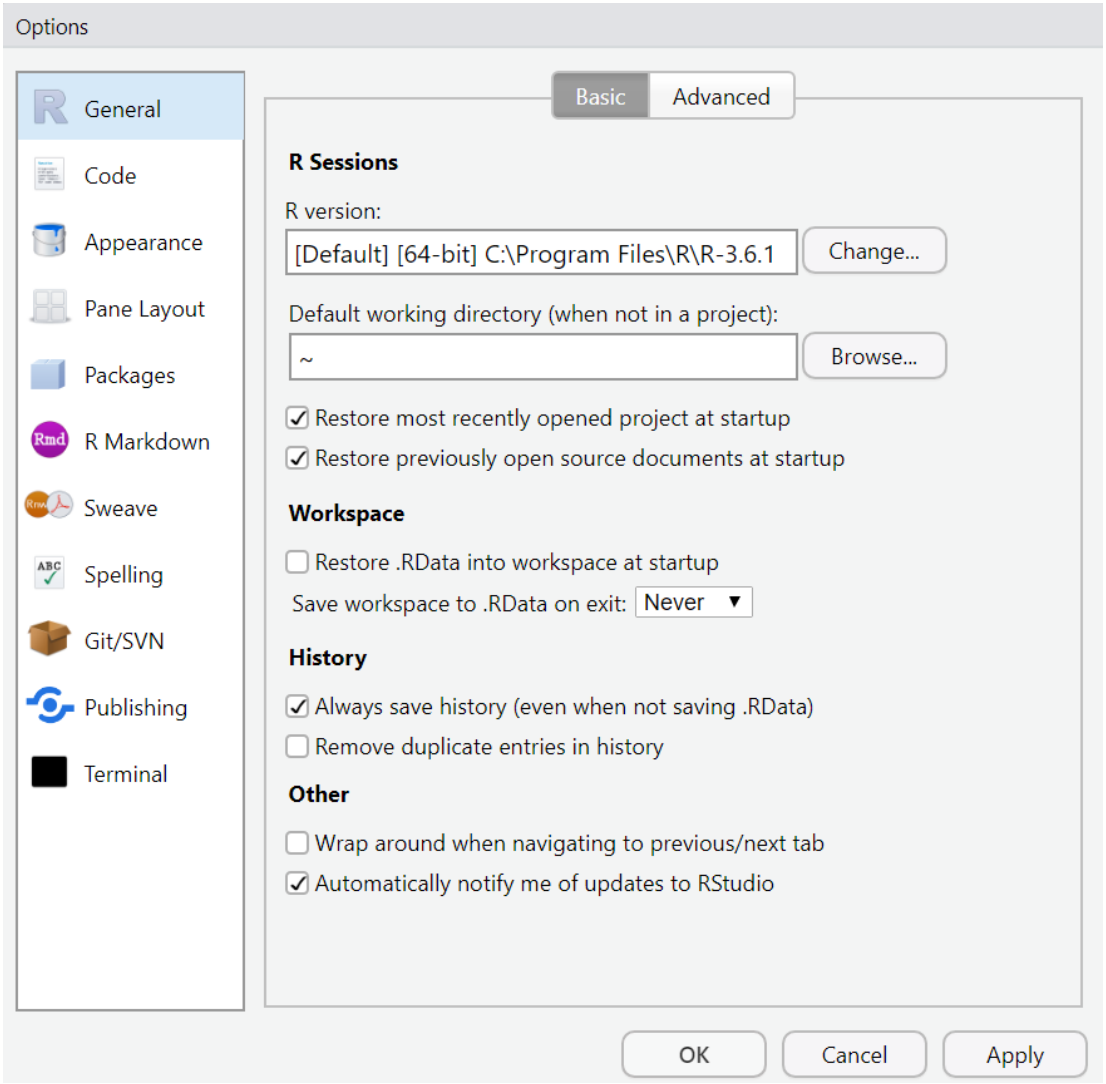
Comment code with a #

E.g:

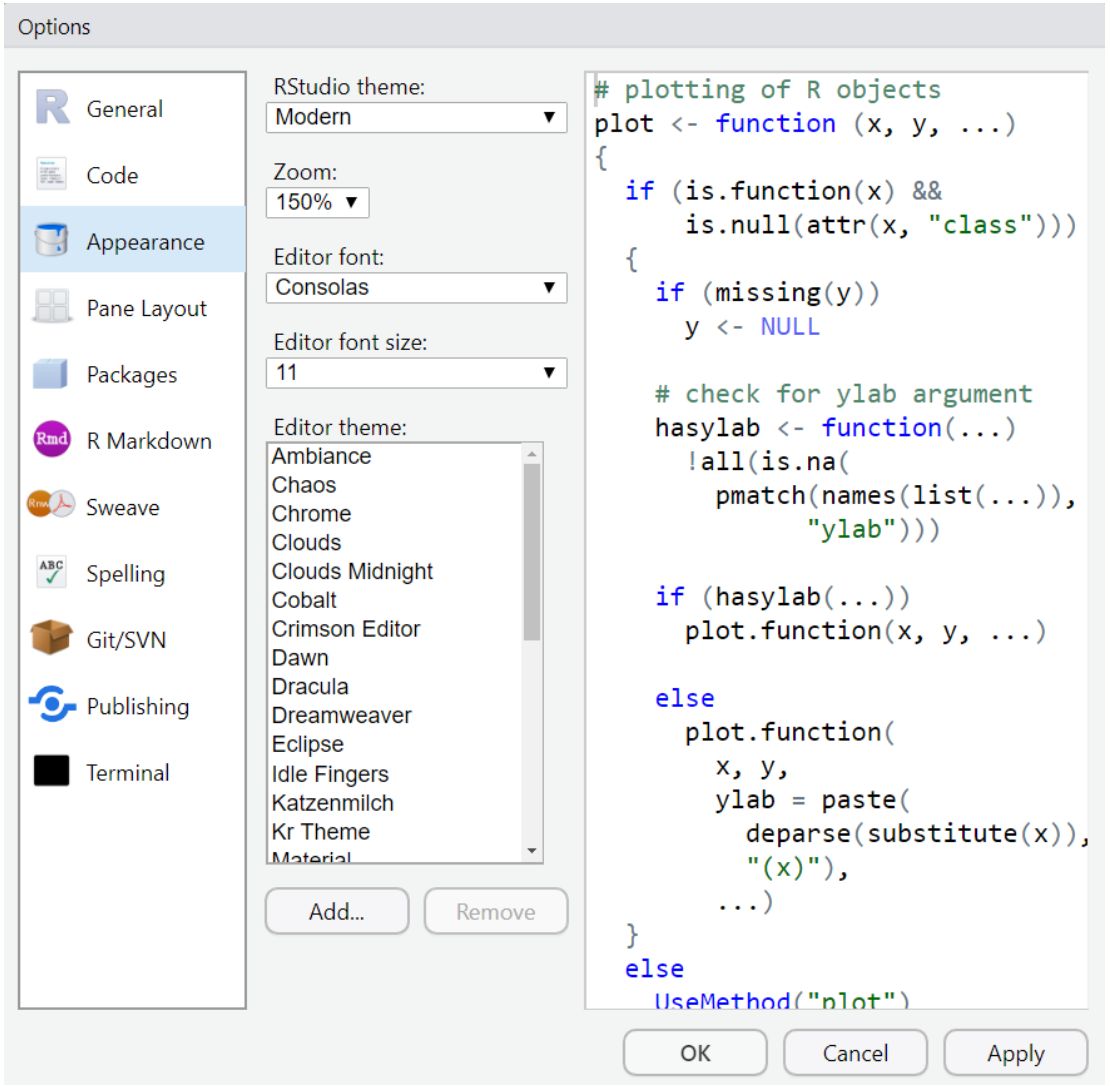
```
# this was a
bad idea
```



Tools → Global Options



Tools → Global Options



Packages

Packages

R packages are like apps for your phone:

Extend the capabilities of the basic or “base" R with extra functions, datasets, documentation.

Packages

Download App

Open the App

Happens just once

Every new session

Download (`install.packages`)

"Open" `library()`

Phone detail from photo by Rahul Chakraborty on Unsplash

Packages

```
install.packages("tidyverse")
```

will download a package to your personal library. Then:

```
library(tidyverse)
```

tells R to load the tidyverse package from your personal library.
(Needed for every new session)

CRAN repository

14,000+ packages. Free. Peer reviewed.

(Manifold possibilities) eg. interactive graphics and dashboards, machine learning, mine twitter data, create PowerPoint docs, maps…

Other ways to get packages (eg. GitHub, …)


Packages:
tidyverse

What is the tidyverse?

The tidyverse package collects (some of) the most popular R packages into one.

All have the same underlying principles:

Provide simple tools (with consistent structure) which may be used together to help solve complex problems.

What is the tidyverse?

During the workshop we will use the ggplot2, dplyr, and readr packages. These are bundled up in the tidyverse package. We load it by running:

```
library(tidyverse)
```

This work is licensed as

End