

Code management

1. Code management within files

- Coding style in R

2. Code management between files

- Project architecture
- Projects
- Git/github with Rstudio

Code management within files

R Style Guide: <https://google-styleguide.googlecode.com/svn/trunk/Rguide.xml>

Goal: improve readability

- Line length and indentation
 - line length max 80 characters
 - indentation: 2 spaces / no tabs
- spaces
 - around binary operators (“=”, “+”, “-”, “<=”, etc.)
 - after all commas
- braces
 - surround “else” with braces
 - open braces on same line, but with one space

```
for (i in 1:x) {
```
- comments
 - “# comment”, not “#comment”

Doesn't meet specs

```
tober2012.R x IncompleteExposureModelCode_Clean.R x IncompleteExposureModelCode_Bad.R x >>
Source on Save Run Source
#-- goal: plot bias in alpha, bias in beta as function of incomplete exposure rate --#
N<-100
S1<-10
Assumed.exposed<-1
True.exposed<-seq(from=.1,to=1,length.out=100)

True.mortrate<-1
Estimated.mortrate<-True.mortrate*True.exposed

#-- FOI = Beta*Y/N
#-- Beta = k * log(1-c)
#-- k = contact rate per unit time
#-- c = P(transmission|contact)
k<-.5
Y<-10 #-- Number of infectives
N<-100 #-- Pop size
True.P.transgivenesscontact<-c<-.99
True.beta<-k*log(1-c)
True.FOI<-True.beta*Y/N
#-- incomplete exposure changes N.
Estimated.FOI<-True.beta*Y/(N*True.exposed)

Estimated.P.transgivenesscontact<-(True.P.transgivenesscontact*(S1*True.exposed)/((N-S1)*True.exposed))
True.exposed.count<-True.exposed*N
svg(file="work/Kezia/Research/EcologyPapers/ClustersAssociations/Plots/RateEstimateBiasPlot_07Mar")

par(mfrow=c(1,1),cex.axis=.8,cex.lab=.8)
plot(Estimated.mortrate~True.exposed.count,type="l",lwd=2,ylim=c(0,2.5),xlab="Actual population size accessible via contact",
lines(Estimated.FOI~True.exposed.count,lty=2,lwd=2)
leg.text<-c("Estimated mortality rate",ylab="Estimated FOI",xlab="Actual mortality rate",xlab="Actual FOI")
legend("topright",leg.text,lty=c(1,2,1,2),bty="n",cex=.7,col=c("black","black","grey50","grey50"))
abline(h=True.FOI,lty=2,lwd=.5,col="grey50")
abline(h=True.mortrate,lty=1,lwd=.5,col="grey50")
mtext(side=1,line=1,outer=T,"Based on a scenario where the perceived population",cex=.85)
mtext(side=1,line=2,outer=T,"contains 100 individuals, 10 of which are infected",cex=.85)

dev.off()
```

Meets specs

```
sourcefuns.R x run.R x Grpsize_12October2012.R x IncompleteExposureModelCode_Clean.R x >>
Source on Save Run Source
1 #-- goal: plot bias in alpha, bias in beta as function of incomplete exposure rate --#
2 N <- 100
3 S1 <- 10
4 Assumed.exposed <- 1
5 True.exposed <- seq(from = .1, to = 1, length.out = 100)
6
7 True.mortrate <- 1
8 Estimated.mortrate <- True.mortrate*True.exposed
9
10 # FOI = Beta*Y/N
11 # Beta = k * log(1-c)
12 # k = contact rate per unit time
13 # c = P(transmission|contact)
14 k <- .5
15 Y <- 10 # Number of infectives
16 N <- 100 # Pop size
17 True.P.transgivenesscontact <- c <- .99
18 True.beta <- -k * log(1 - c)
19 True.FOI <- True.beta * Y / N
20 # incomplete exposure changes N.
21 Estimated.FOI <- True.beta * Y / (N * True.exposed)
22
23 Estimated.P.transgivenesscontact <- (True.P.transgivenesscontact * (S1 * True.exposed) *
24 ((N - S1) * True.exposed)
25 ) / (S1 * (N - S1))
26
27 True.exposed.count <- True.exposed * N
28 svg(file="./Plots/RateEstimateBiasPlot.svg", width = 10, height = 8, bg = "white")
29
30 par(mfrow = c(1, 1), cex.axis = .8, cex.lab = .8)
31 plot(Estimated.mortrate ~ True.exposed.count,
32 type = "l", lwd = 2, ylim = c(0, 2.5),
33 xlab = "Actual population size accessible via contact",
34 ylab = "Estimated rate assuming entire population available"
35 )
36 lines(Estimated.FOI ~ True.exposed.count, lty = 2, lwd = 2)
37 leg.text <- c("Estimated mortality rate",
38 ylab = "Estimated FOI", "Actual mortality rate", "Actual FOI"
39 )
40 legend("topright", leg.text, lty=c(1, 2, 1, 2), bty = "n", cex = .7,
41 col = c("black", "black", "grey50", "grey50"))
42 abline(h = True.FOI, lty = 2, lwd = .5, col = "grey50")
43 abline(h = True.mortrate, lty = 1, lwd = .5, col = "grey50")
44
1:1 (Top Level) R Script
```

Doesn't meet specs

Meets specs

Spacing

```
tober2012.R x IncompleteExposureModelCode_Clean.R x IncompleteExposureModelCode_Bad.R x >>
Source on Save Run Source
#-- goal: plot bias in alpha, bias in beta as function of incomplete exposure rate --#
N<-100
S1<-10
Assumed.exposed<-1
True.exposed<-seq(from=.1,to=1,length.out=100)

True.mortrate<-1
Estimated.mortrate<-True.mortrate*True.exposed

#-- FOI = Beta*Y/N
#-- Beta = k * log(1-c)
#-- k = contact rate per unit time
#-- c = P(transmission|contact)
k<-.5
Y<-10 #-- Number of infectives
N<-100 #-- Pop size
True.P.transgivencontact<-c<-.99
True.beta<-k*log(1-c)
True.FOI<-True.beta*Y/N
#-- incomplete exposure changes N.
Estimated.FOI<-True.beta*Y/(N*True.exposed)

Estimated.P.transgivencontact<-(True.P.transgivencontact*(S1*True.exposed)/((N-S1)*True.exposed))
True.exposed.count<-True.exposed*N
svg(file="work/Kezia/Research/EcologyPapers/ClustersAssociations/Plots/RateEstimateBiasPlot_07Mar")

par(mfrow=c(1,1),cex.axis=.8,cex.lab=.8)
plot(Estimated.mortrate~True.exposed.count,type="l",lwd=2,ylim=c(0,2.5),xlab="Actual population",
lines(Estimated.FOI~True.exposed.count,lty=2,lwd=2)
leg.text<-c("Estimated mortality rate",ylab="Estimated FOI","Actual mortality rate","Actual FOI")
legend("topright",leg.text,lty=c(1,2,1,2),bty="n",cex=.7,col=c("black","black","grey50","grey50"))
abline(h=True.FOI,lty=2,lwd=.5,col="grey50")
abline(h=True.mortrate,lty=1,lwd=.5,col="grey50")
mtext(side=1,line=1,outer=T,"Based on a scenario where the perceived population",cex=.85)
mtext(side=1,line=2,outer=T,"contains 100 individuals, 10 of which are infected",cex=.85)

dev.off()
```

```
sourcefuns.R x run.R x Grpsize_12October2012.R x IncompleteExposureModelCode_Clean.R x >>
Source on Save Run Source
1 #-- goal: plot bias in alpha, bias in beta as function of incomplete exposure rate --#
2 N <- 100
3 S1 <- 10
4 Assumed.exposed <- 1
5 True.exposed <- seq(from = .1, to = 1, length.out = 100)
6
7 True.mortrate <- 1
8 Estimated.mortrate <- True.mortrate*True.exposed
9
10 # FOI = Beta*Y/N
11 # Beta = k * log(1-c)
12 # k = contact rate per unit time
13 # c = P(transmission|contact)
14 k <- .5
15 Y <- 10 # Number of infectives
16 N <- 100 # Pop size
17 True.P.transgivencontact <- c <- .99
18 True.beta <- -k * log(1 - c)
19 True.FOI <- True.beta * Y / N
20 # incomplete exposure changes N.
21 Estimated.FOI <- True.beta * Y / (N * True.exposed)
22
23 Estimated.P.transgivencontact <- (True.P.transgivencontact * (S1 * True.exposed) *
24 ((N - S1) * True.exposed)
25 ) / (S1 * (N - S1))
26
27 True.exposed.count <- True.exposed * N
28
29 svg(file="./Plots/RateEstimateBiasPlot.svg", width = 10, height = 8, bg = "white")
30
31 par(mfrow = c(1, 1), cex.axis = .8, cex.lab = .8)
32 plot(Estimated.mortrate ~ True.exposed.count,
33 type = "l", lwd = 2, ylim = c(0, 2.5),
34 xlab = "Actual population size accessible via contact",
35 ylab = "Estimated rate assuming entire population available"
36 )
37 lines(Estimated.FOI ~ True.exposed.count, lty = 2, lwd = 2)
38 leg.text <- c("Estimated mortality rate",
39 ylab = "Estimated FOI","Actual mortality rate","Actual FOI")
40 legend("topright", leg.text, lty=c(1, 2, 1, 2), bty = "n", cex = .7,
41 col = c("black", "black", "grey50", "grey50"))
42 abline(h = True.FOI, lty = 2, lwd = .5, col = "grey50")
43 abline(h = True.mortrate, lty = 1, lwd = .5, col = "grey50")
44
45 1:1 (Top Level) >
```

Doesn't meet specs

Meets specs

Spacing

Paths

```
N<-100
S1<-10
Assumed.exposed<-1
True.exposed<-seq(from=.1,to=1,length.out=100)
```

```
1 #-- goal: plot bias in alpha, bias in beta as function of incomplete exposure rate --#
2 N <- 100
3 S1 <- 10
4 Assumed.exposed <- 1
5 True.exposed <- seq(from = .1, to = 1, length.out = 100)
```

```
True.mortrate<-1
Estimated.mortrate<-True.mortrate*True.exposed
```

```
6
7 True.mortrate <- 1
8 Estimated.mortrate <- True.mortrate*True.exposed
```

```
-- FOI = Beta*Y/N
-- Beta = k * log(1-c)
-- k = contact rate per unit time
-- c = P(transmission|contact)
k<-.5
Y<-10 #-- Number of infectives
N<-100 #-- Pop size
True.P.transgivencontact<-c<-.99
True.beta<-k*log(1-c)
True.FOI<-True.beta*Y/N
#-- incomplete exposure changes N.
Estimated.FOI<-True.beta*Y/(N*True.exposed)
```

```
10 # FOI = Beta*Y/N
11 # Beta = k * log(1-c)
12 # k = contact rate per unit time
13 # c = P(transmission|contact)
14 k <- .5
15 Y <- 10 # Number of infectives
16 N <- 100 # Pop size
17 True.P.transgivencontact <- c <- .99
18 True.beta <- -k * log(1 - c)
19 True.FOI <- True.beta * Y / N
20 # incomplete exposure changes N.
21 Estimated.FOI <- True.beta * Y / (N * True.exposed)
```

```
Estimated.P.transgivencontact<-(True.P.transgivencontact*(S1*True.exposed)/((N-S1)*True.exposed))
True.exposed.count<-True.exposed*N
svg(file="work/Kezia/Research/EcologyPapers/ClustersAssociations/Plots/RateEstimateBiasPlot_07Ma
```

```
23 Estimated.P.transgivencontact <- (True.P.transgivencontact * (S1 * True.exposed) *
24 ((N - S1) * True.exposed)
25 ) / (S1 * (N - S1))
26 True.exposed.count <- True.exposed * N
27 svg(file="./Plots/RateEstimateBiasPlot.svg", width = 10, height = 8, bg = "white")
```

```
par(mfrow=c(1,1),cex.axis=.8,cex.lab=.8)
plot(Estimated.mortrate~True.exposed.count,type="l",lwd=2,ylim=c(0,2.5),xlab="Actual population size accessible via contact",
lines(Estimated.FOI~True.exposed.count,lty=2,lwd=2)
leg.text<-c("Estimated mortality rate",ylab="Estimated FOI","Actual mortality rate","Actual FOI"
legend("topright",leg.text,lty=c(1,2,1,2),bty="n",cex=.7,col=c("black","black","grey50","grey50")
abline(h=True.FOI,lty=2,lwd=.5,col="grey50")
abline(h=True.mortrate,lty=1,lwd=.5,col="grey50")
mtext(side=1,line=1,outer=T,"Based on a scenario where the perceived population",cex=.85)
mtext(side=1,line=2,outer=T,"contains 100 individuals, 10 of which are infected",cex=.85)

dev.off()
```

```
28
29 par(mfrow = c(1, 1), cex.axis = .8, cex.lab = .8)
30 plot(Estimated.mortrate ~ True.exposed.count,
31 type = "l", lwd = 2, ylim = c(0, 2.5),
32 xlab = "Actual population size accessible via contact",
33 ylab = "Estimated rate assuming entire population available"
34 )
35 lines(Estimated.FOI ~ True.exposed.count, lty = 2, lwd = 2)
36 leg.text <- c("Estimated mortality rate",
37 ylab = "Estimated FOI", "Actual mortality rate", "Actual FOI"
38 )
39 legend("topright", leg.text, lty=c(1, 2, 1, 2), bty = "n", cex = .7,
40 col = c("black", "black", "grey50", "grey50"))
41 abline(h = True.FOI, lty = 2, lwd = .5, col = "grey50")
42 abline(h = True.mortrate, lty = 1, lwd = .5, col = "grey50")
```

1:1 (Top Level)

R Script

Doesn't meet specs

Meets specs

Spacing

Paths

Wrap lines

```
tober2012.R x IncompleteExposureModelCode_Clean.R x IncompleteExposureModelCode_Bad.R x >>
Source on Save Run Source
#-- goal: plot bias in alpha, bias in beta as function of incomplete exposure rate --#
N<-100
S1<-10
Assumed.exposed<-1
True.exposed<-seq(from=.1,to=1,length.out=100)

True.mortrate<-1
Estimated.mortrate<-True.mortrate*True.exposed

#-- FOI = Beta*Y/N
#-- Beta = k * log(1-c)
#-- k = contact rate per unit time
#-- c = P(transmission|contact)
k<-.5
Y<-10 #-- Number of infectives
N<-100 #-- Pop size
True.P.transgivencontact<-c<-.99
True.beta<-k*log(1-c)
True.FOI<-True.beta*Y/N
#-- incomplete exposure changes N.
Estimated.FOI<-True.beta*Y/(N*True.exposed)

Estimated.P.transgivencontact<-((True.P.transgivencontact*(S1*True.exposed)/((N-S1)*True.exposed))
True.exposed.count<-True.exposed*N
svg(file="work/Kezia/Research/EcologyPapers/ClustersAssociations/Plots/RateEstimateBiasPlot_07Ma

par(mfrow=c(1,1),cex.axis=.8,cex.lab=.8)
plot(Estimated.mortrate~True.exposed.count,type="l",lwd=2,ylim=c(0,2.5),xlab="Actual population
lines(Estimated.FOI~True.exposed.count,lty=2,lwd=2)
leg.text<-c("Estimated mortality rate",ylab="Estimated FOI","Actual mortality rate","Actual FOI"
legend("topright",leg.text,lty=c(1,2,1,2),bty="n",cex=.7,col=c("black","black","grey50","grey50"
abline(h=True.FOI,lty=2,lwd=.5,col="grey50")
abline(h=True.mortrate,lty=1,lwd=.5,col="grey50")
mtext(side=1,line=1,outer=T,"Based on a scenario where the perceived population",cex=.85)
mtext(side=1,line=2,outer=T,"contains 100 individuals, 10 of which are infected",cex=.85)

dev.off()
```

```
sourcefuns.R x run.R x Grpsize_12October2012.R x IncompleteExposureModelCode_Clean.R x >>
Source on Save Run Source
1 #-- goal: plot bias in alpha, bias in beta as function of incomplete exposure rate --#
2 N <- 100
3 S1 <- 10
4 Assumed.exposed <- 1
5 True.exposed <- seq(from = .1, to = 1, length.out = 100)
6
7 True.mortrate <- 1
8 Estimated.mortrate <- True.mortrate*True.exposed
9
10 # FOI = Beta*Y/N
11 # Beta = k * log(1-c)
12 # k = contact rate per unit time
13 # c = P(transmission|contact)
14 k <- .5
15 Y <- 10 # Number of infectives
16 N <- 100 # Pop size
17 True.P.transgivencontact <- c <- .99
18 True.beta <- -k * log(1 - c)
19 True.FOI <- True.beta * Y / N
20 # incomplete exposure changes N.
21 Estimated.FOI <- True.beta * Y / (N * True.exposed)
22
23 Estimated.P.transgivencontact <- (True.P.transgivencontact * (S1 * True.exposed) *
24 ((N - S1) * True.exposed)
25 ) / (S1 * (N - S1))
26 True.exposed.count <- True.exposed * N
27 svg(file="./Plots/RateEstimateBiasPlot.svg", width = 10, height = 8, bg = "white")
28
29 par(mfrow = c(1, 1), cex.axis = .8, cex.lab = .8)
30 plot(Estimated.mortrate ~ True.exposed.count,
31 type = "l", lwd = 2, ylim = c(0, 2.5),
32 xlab = "Actual population size accessible via contact",
33 ylab = "Estimated rate assuming entire population available"
34 )
35 lines(Estimated.FOI ~ True.exposed.count, lty = 2, lwd = 2)
36 leg.text <- c("Estimated mortality rate",
37 ylab = "Estimated FOI", "Actual mortality rate", "Actual FOI"
38 )
39 legend("topright", leg.text, lty=c(1, 2, 1, 2), bty = "n", cex = .7,
40 col = c("black", "black", "grey50", "grey50"))
41 abline(h = True.FOI, lty = 2, lwd = .5, col = "grey50")
42 abline(h = True.mortrate, lty = 1, lwd = .5, col = "grey50")
43
44 1:1 (Top Level) >
```

Code management within files

```
CalculateSampleCovariance <- function(x, y, verbose = TRUE) {  
  # Computes the sample covariance between two vectors.  
  #  
  # Args:  
  #   x: One of two vectors whose sample covariance is to be calculated.  
  #   y: The other vector. x and y must have the same length, greater than one,  
  #       with no missing values.  
  #   verbose: If TRUE, prints sample covariance; if not, not. Default is TRUE.  
  #  
  # Returns:  
  #   The sample covariance between x and y.  
  n <- length(x)  
  # Error handling  
  if (n <= 1 || n != length(y)) {  
    stop("Arguments x and y have different lengths: ",  
         length(x), " and ", length(y), ".")  
  }  
  if (TRUE %in% is.na(x) || TRUE %in% is.na(y)) {  
    stop(" Arguments x and y must not have missing values.")  
  }  
  covariance <- var(x, y)  
  if (verbose)  
    cat("Covariance = ", round(covariance, 4), ".\n", sep = "")  
  return(covariance)  
}
```

Code management within files

Function headers
should contain

1. description
2. list of arguments
3. list of returns



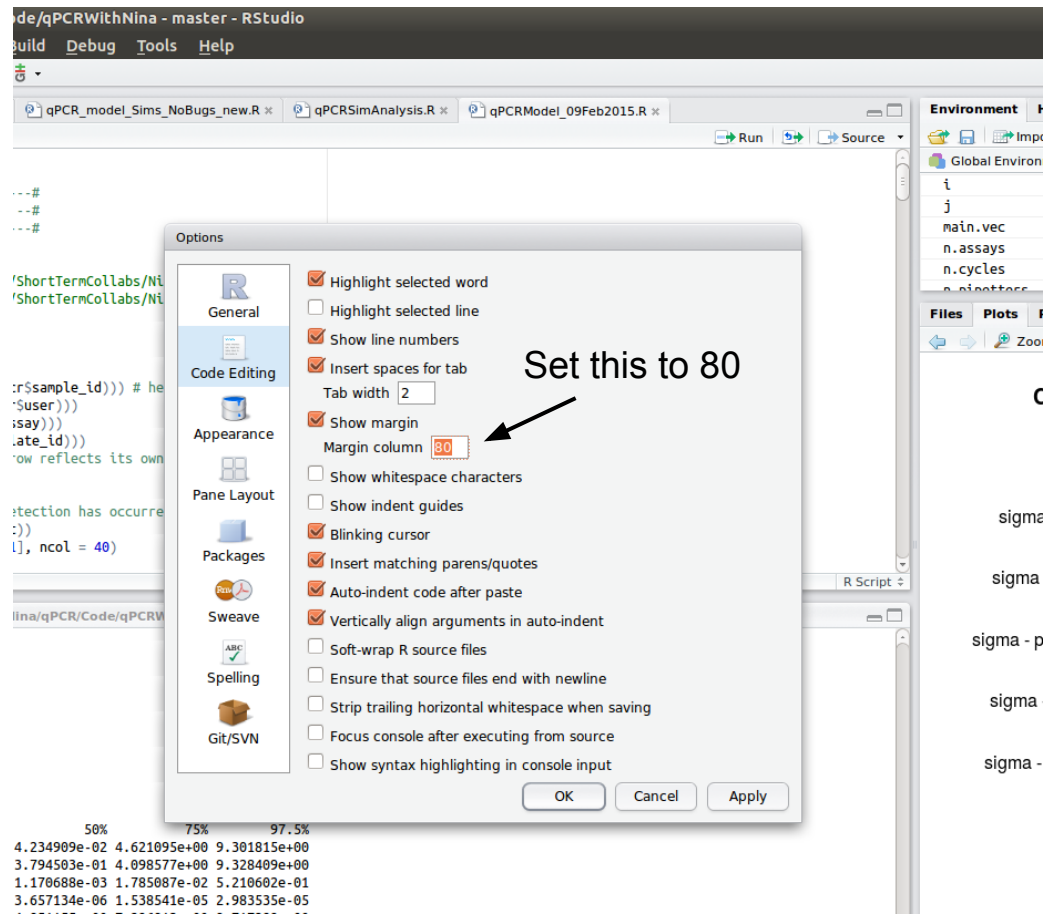
```
CalculateSampleCovariance <- function(x, y, verbose = TRUE) {  
  # Computes the sample covariance between two vectors.  
  #  
  # Args:  
  # x: One of two vectors whose sample covariance is to be calculated.  
  # y: The other vector. x and y must have the same length, greater than one,  
  #     with no missing values.  
  # verbose: If TRUE, prints sample covariance; if not, not. Default is TRUE.  
  #  
  # Returns:  
  # The sample covariance between x and y.  
  
  n <- length(x)  
  
  # Error handling  
  if (n <= 1 || n != length(y)) {  
    stop("Arguments x and y have different lengths: ",  
         length(x), " and ", length(y), ".")  
  }  
  
  if (TRUE %in% is.na(x) || TRUE %in% is.na(y)) {  
    stop("Arguments x and y must not have missing values.")  
  }  
  
  covariance <- var(x, y)  
  
  if (verbose)  
    cat("Covariance = ", round(covariance, 4), ".\n", sep = "")  
  
  return(covariance)  
}
```


Code management within files

Place 80-character line in RStudio

Tools -> Global options

- Code editing tab
- Set Margin Column box to 80



Code management between files:

Projects

- Project: All files that go with one analysis
- Store all project files in a single repository
- Projects consist of
 - Data
 - Code
 - Documentation
 - Figures
- Mix of R files, other files

Code management systems

- SVN (“Subversion”)
- Github
 - Advantage: well-known
 - Disadvantage: limited private repos
 - 5 free private repos for students: <https://education.github.com/>
- Bitbucket
 - Advantage: free private repositories

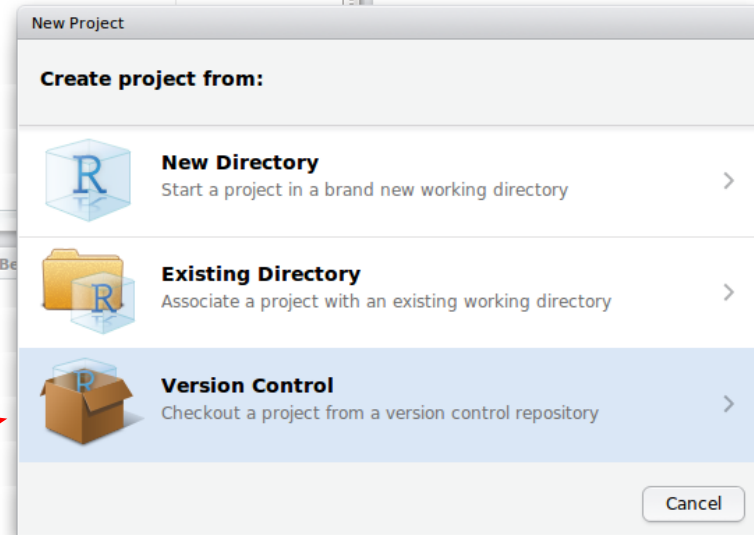
Github / RStudio Link

1. Build a new project

- Launch RStudio
- File -> New Project

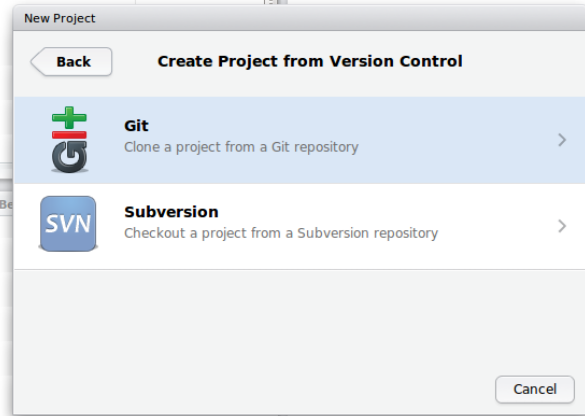
2. Select Version Control

- lets you link project to git or SVN



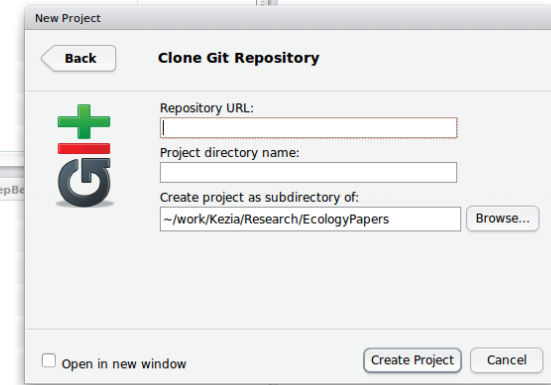
Github / RStudio Link

3. Select Git



4. Link to github repo

(“repo” = “repository”; analogous to “project”)



Github / RStudio Link

GitHub - Build software better, together. - Mozilla Firefox

Managing a statistical a... x Inbox (336) - kezia... x Data - Google Drive x CompiledData_09Fe... x Code Management i... x GitHub - Build softw... x

GitHub, Inc. (US) | https://github.com

Search GitHub

Explore Features Enterprise Blog

Sign up Sign in

Build software better, together.

Powerful collaboration, code review, and code management for open source and private projects. Need private repositories? [Upgraded plans start at \\$7/mo.](#)

Pick a username

Your email

Create a password

Use at least one lowercase letter, one numeral, and seven characters.

[Sign up for GitHub](#)

By clicking "Sign up for GitHub", you agree to our [terms of service](#) and [privacy policy](#). We will send you account related emails occasionally.

Introducing a faster, more flexible GitHub Enterprise.

Now available on AWS.

[Tell me more](#)

Github / RStudio Link

“Repositories” tab

The screenshot shows the GitHub profile page for user kmanlove. The browser address bar displays the URL <https://github.com/kmanlove?tab=repositories>. The page features a profile picture of a man in a winter hat, the name **Kezia Manlove** (kmanlove), and location information (Bozeman). The 'Repositories' tab is selected, showing a list of repositories. A red arrow points to the 'Repositories' tab, and another red arrow points to the 'New' button in the top right corner of the repository list.

Contributions Repositories Public activity Edit profile

Find a repository... Search All Public Private Sources Forks Mirrors New

SheepBehavior R ★ 0 0
bighorn behavioral analyses from Hells Canyon
Updated an hour ago

ParticipationInComputing R ★ 0 0
Updated 7 days ago

BighornIPM R ★ 0 0
Updated 7 days ago

BighornCrossScaleMathModel R ★ 0 0
Updated 7 days ago

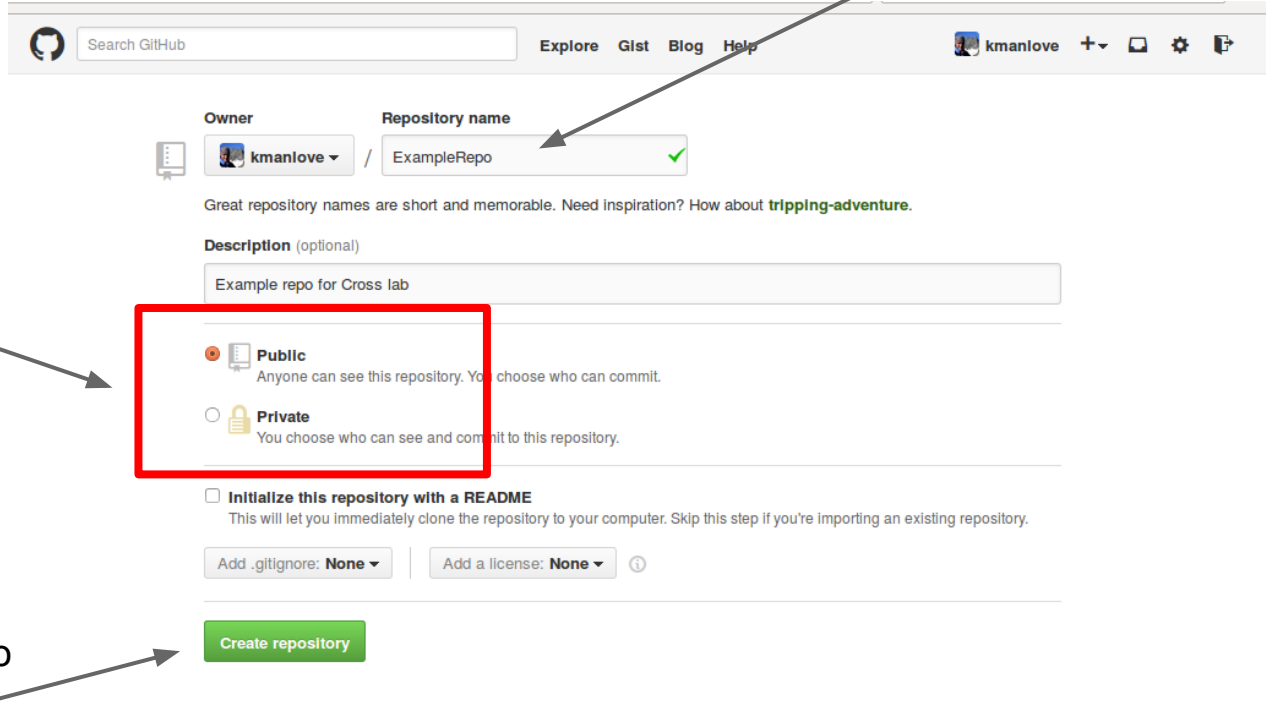
WomenInStatsComputing R ★ 0 0
Updated 18 days ago

BB14OutbreakCode R ★ 0 0

Select “New”

Github / RStudio Link

Fill in project name



The screenshot shows the GitHub 'Create repository' interface. At the top, there's a search bar and navigation links. The main form has two sections: 'Owner' and 'Repository name'. The 'Repository name' field contains 'ExampleRepo' and has a green checkmark. Below this is a hint: 'Great repository names are short and memorable. Need inspiration? How about [tripping-adventure](#).' The 'Description' field is optional and contains 'Example repo for Cross lab'. The 'Visibility' section is highlighted with a red box and contains two options: 'Public' (selected) and 'Private'. Below this is a checkbox for 'Initialize this repository with a README'. At the bottom, there are dropdowns for '.gitignore' and 'License', both set to 'None'. A green 'Create repository' button is at the bottom right. Annotations with arrows point to the 'Repository name' field, the 'Public' option, and the 'Create repository' button.

Owner: kmanlove

Repository name: ExampleRepo ✓

Great repository names are short and memorable. Need inspiration? How about [tripping-adventure](#).

Description (optional): Example repo for Cross lab

☒ **Public**
Anyone can see this repository. You choose who can commit.

☐ **Private**
You choose who can see and commit to this repository.

☐ **Initialize this repository with a README**
This will let you immediately clone the repository to your computer. Skip this step if you're importing an existing repository.

Add .gitignore: **None** | Add a license: **None** ⓘ

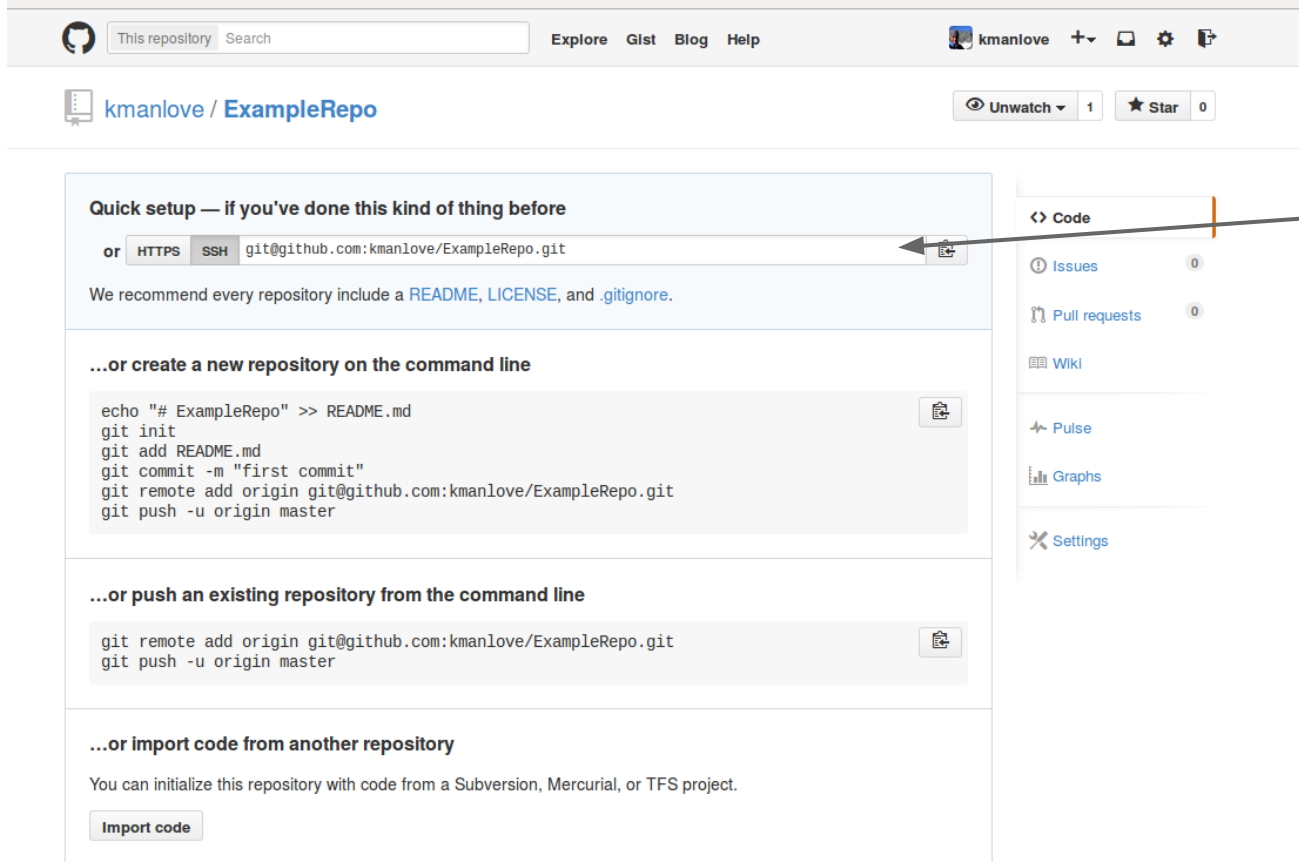
Create repository

Defaults to
"Public"

Click here to
finalize



Github / RStudio Link



The screenshot shows the GitHub interface for a repository named 'ExampleRepo' by user 'kmanlove'. The top navigation bar includes links for 'Explore', 'Gist', 'Blog', and 'Help'. The repository header shows 'Unwatch' (1), 'Star' (0), and a 'Code' button. The main content area is titled 'Quick setup — if you've done this kind of thing before' and offers two options: 'HTTPS' and 'SSH'. The 'SSH' option is selected, showing the URL 'git@github.com:kmanlove/ExampleRepo.git'. Below this, there are instructions for creating a new repository or pushing an existing one from the command line, each with a copy icon. The sidebar on the right contains links for 'Issues', 'Pull requests', 'Wiki', 'Pulse', 'Graphs', and 'Settings'.

Quick setup — if you've done this kind of thing before

or **HTTPS** **SSH** `git@github.com:kmanlove/ExampleRepo.git`

We recommend every repository include a [README](#), [LICENSE](#), and [.gitignore](#).

...or create a new repository on the command line

```
echo "# ExampleRepo" >> README.md
git init
git add README.md
git commit -m "first commit"
git remote add origin git@github.com:kmanlove/ExampleRepo.git
git push -u origin master
```

...or push an existing repository from the command line

```
git remote add origin git@github.com:kmanlove/ExampleRepo.git
git push -u origin master
```

...or import code from another repository

You can initialize this repository with code from a Subversion, Mercurial, or TFS project.

[Import code](#)

Code

Issues 0

Pull requests 0

Wiki

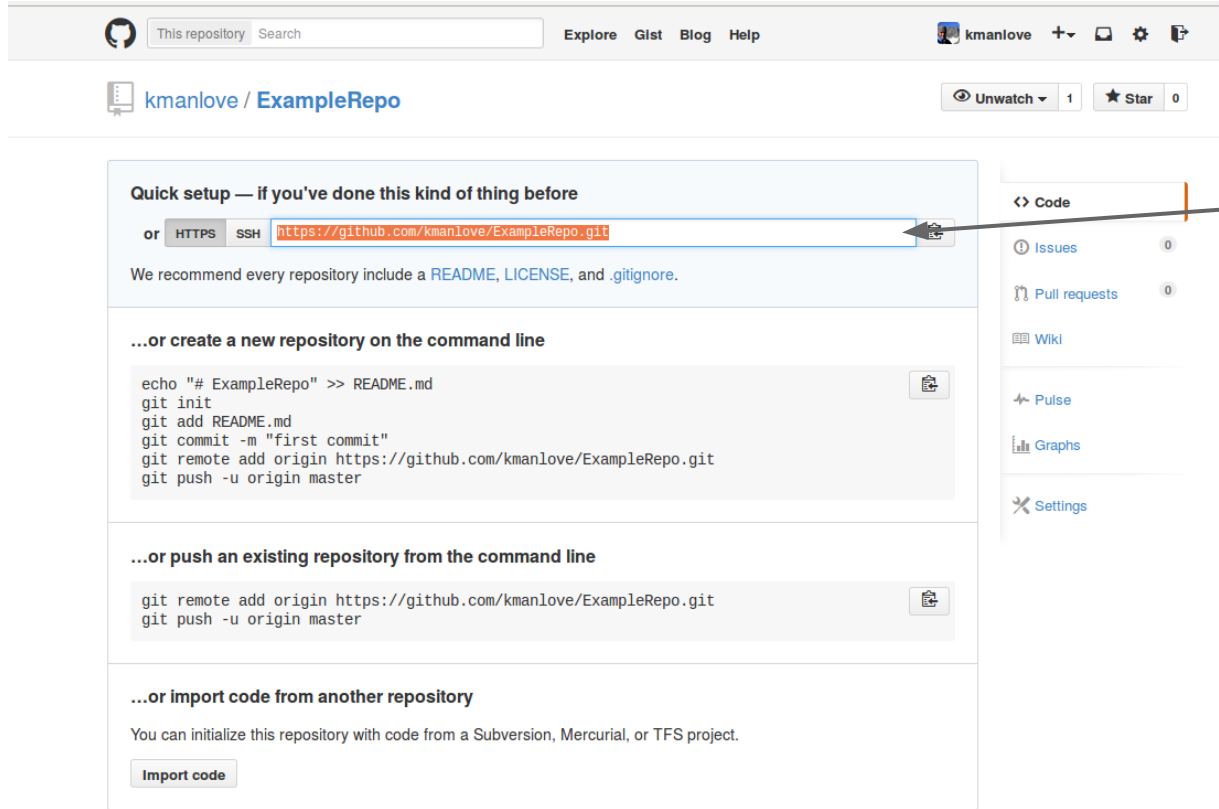
Pulse

Graphs

Settings

This link connects repo to local directory and R project

Github / RStudio Link



The screenshot shows the GitHub repository page for 'kmanlove / ExampleRepo'. The 'Quick setup' section is highlighted, showing the 'HTTPS' option selected. The URL 'https://github.com/kmanlove/ExampleRepo.git' is highlighted in the input field. A black arrow points from the 'Code' button in the right sidebar to the highlighted URL. Below the URL, there are instructions for creating a new repository on the command line and pushing an existing repository from the command line. At the bottom, there is a 'ProTip!' section with a lightbulb icon and the text: 'Use the URL for this page when adding GitHub as a remote.'

Quick setup — if you've done this kind of thing before

or ☐ HTTPS ☒ SSH

We recommend every repository include a [README](#), [LICENSE](#), and [.gitignore](#).

...or create a new repository on the command line

```
echo "# ExampleRepo" >> README.md
git init
git add README.md
git commit -m "first commit"
git remote add origin https://github.com/kmanlove/ExampleRepo.git
git push -u origin master
```

...or push an existing repository from the command line

```
git remote add origin https://github.com/kmanlove/ExampleRepo.git
git push -u origin master
```

...or import code from another repository

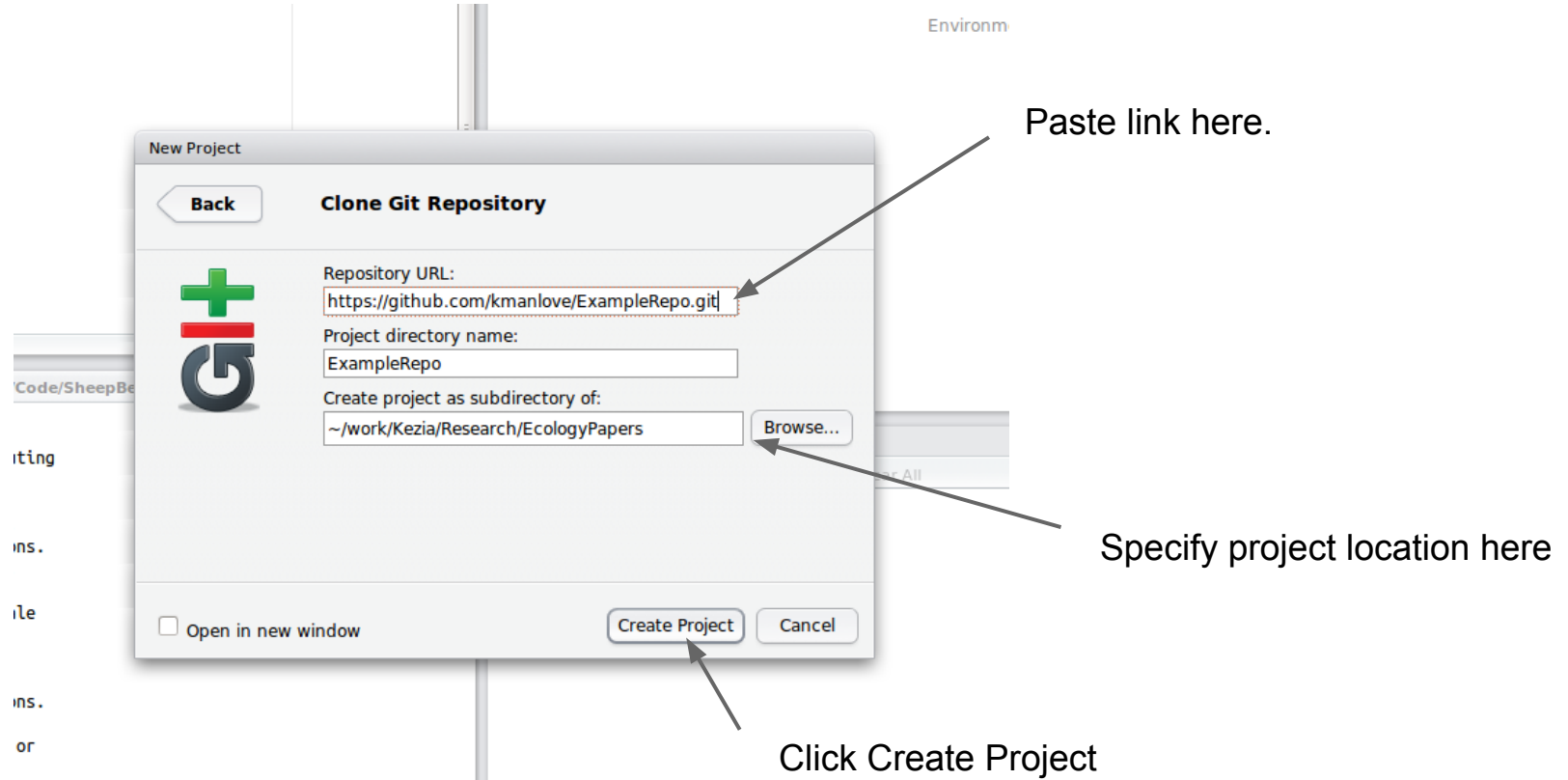
You can initialize this repository with code from a Subversion, Mercurial, or TFS project.

[Import code](#)

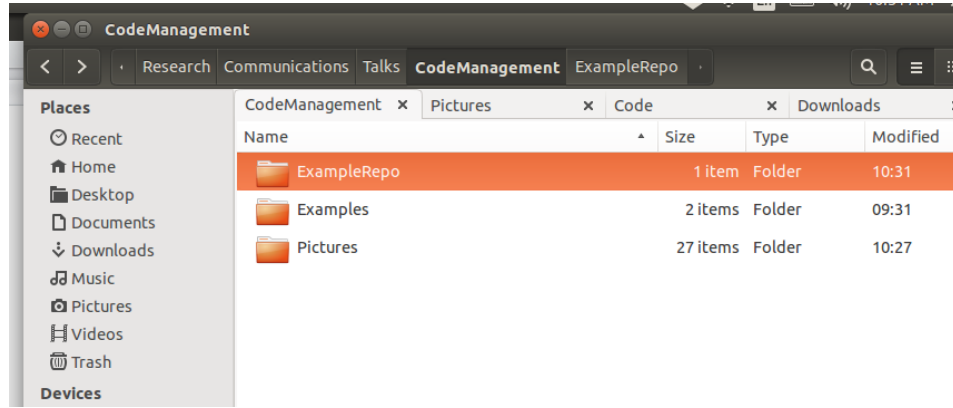
ProTip! Use the URL for this page when adding GitHub as a remote.

1. Switch from SSH to HTTPS
2. Copy link

Github / RStudio Link



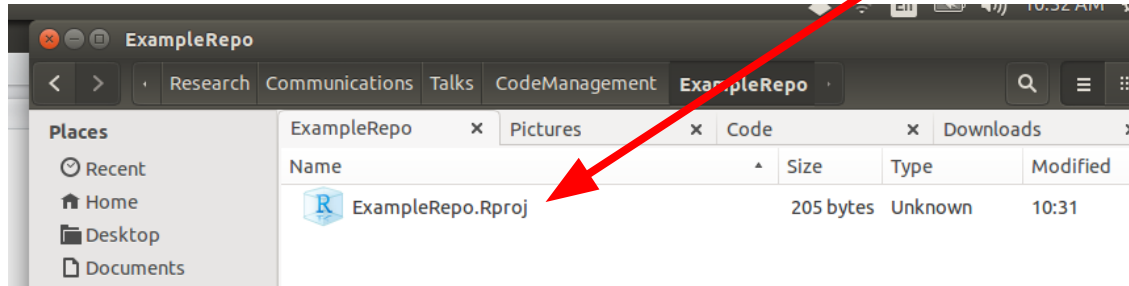
Github / RStudio Link



RStudio builds project folder

(DON'T BUILD the ExampleRepo folder manually)

RProject now sits in ExampleRepo folder



- Double-click to launch in Rstudio
- Folder where project sits is default working directory

Project structure

Prioritize mobility

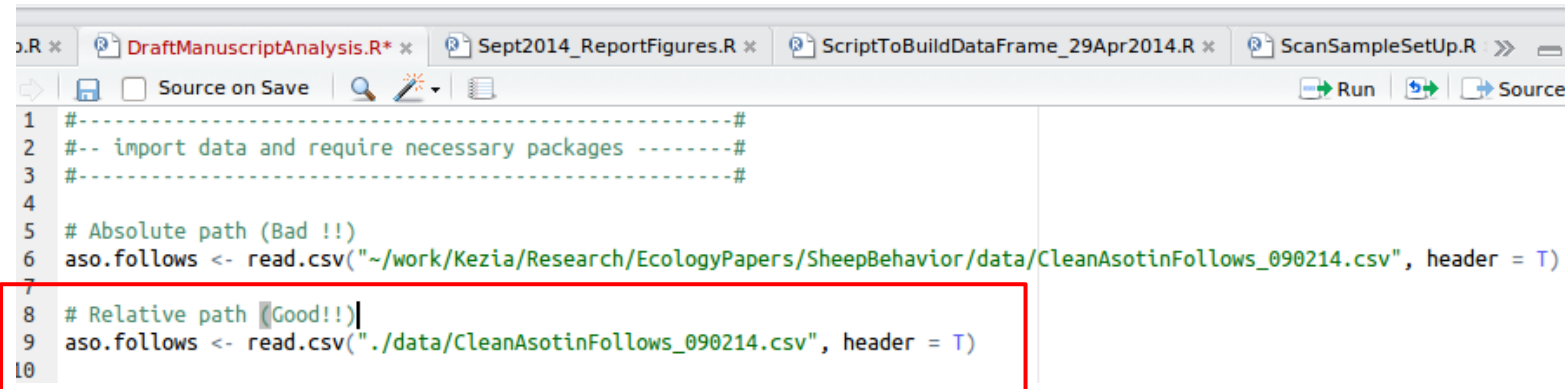
1. Reproducible research
2. Organized projects are easier to move to servers

Project structure

1. Keep paths RELATIVE

Project structure

1. Keep paths RELATIVE



```
1 #-----#
2 #-- import data and require necessary packages -----#
3 #-----#
4
5 # Absolute path (Bad !!)
6 aso.follows <- read.csv("~/work/Kezia/Research/EcologyPapers/SheepBehavior/data/CleanAsotinFollows_090214.csv", header = T)
7
8 # Relative path (Good!!)
9 aso.follows <- read.csv("./data/CleanAsotinFollows_090214.csv", header = T)
10
```

Project structure

1. Keep paths RELATIVE
2. Build consistent set of internal files
 - data
 - documentation
 - R
 - figures

Project structure

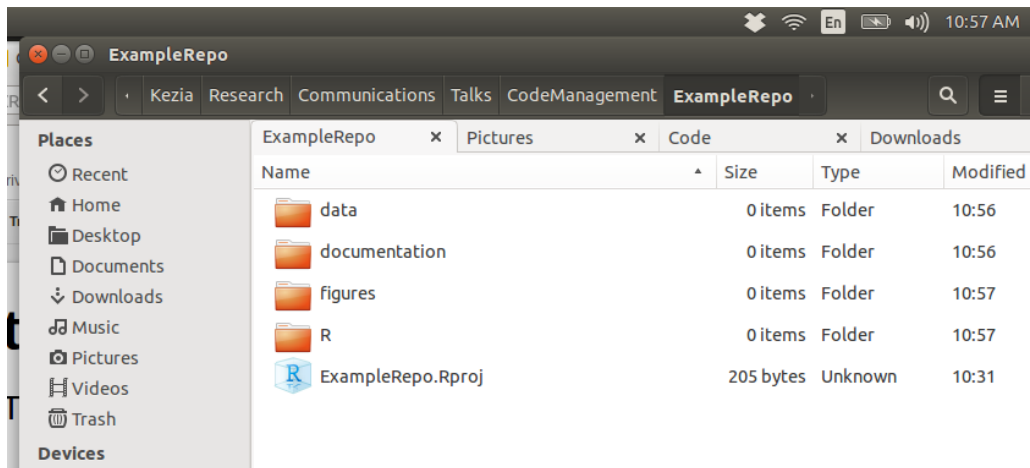
1. Keep paths RELATIVE
2. Build consistent set of internal files

data

documentation

R

figures



Project structure

1. Keep paths RELATIVE
2. Build consistent set of internal files

data

documentation

R

figures

3. Build code in logical batches

load.R ----- pulls in datasets

clean.R ----- contains all data processing

sourcefuns.R -- loads all handwritten analysis functions

run.R ----- runs analysis functions / produces output

Project structure

1. Keep paths RELATIVE
2. Build consistent set of internal files

data

documentation

R

figures

3. Build code in logical batches

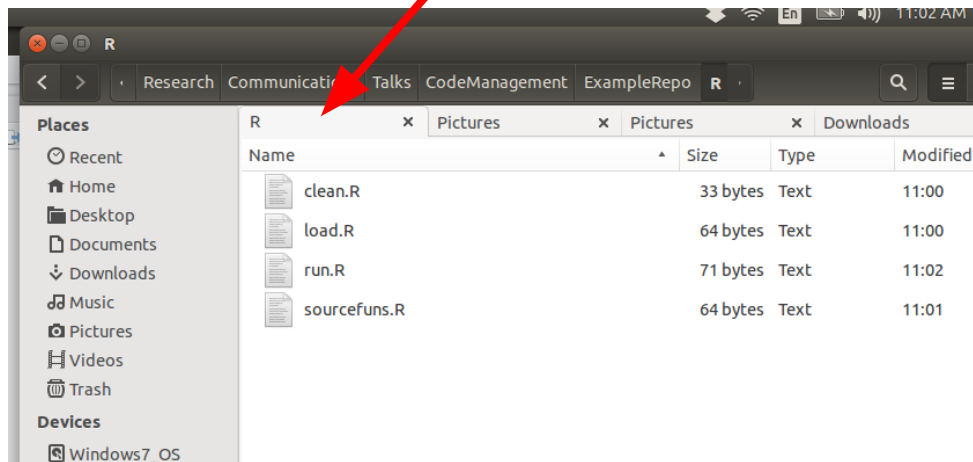
load.R ----- pulls in datasets

clean.R ----- contains all data processing

sourcefuns.R -- loads all handwritten analysis functions

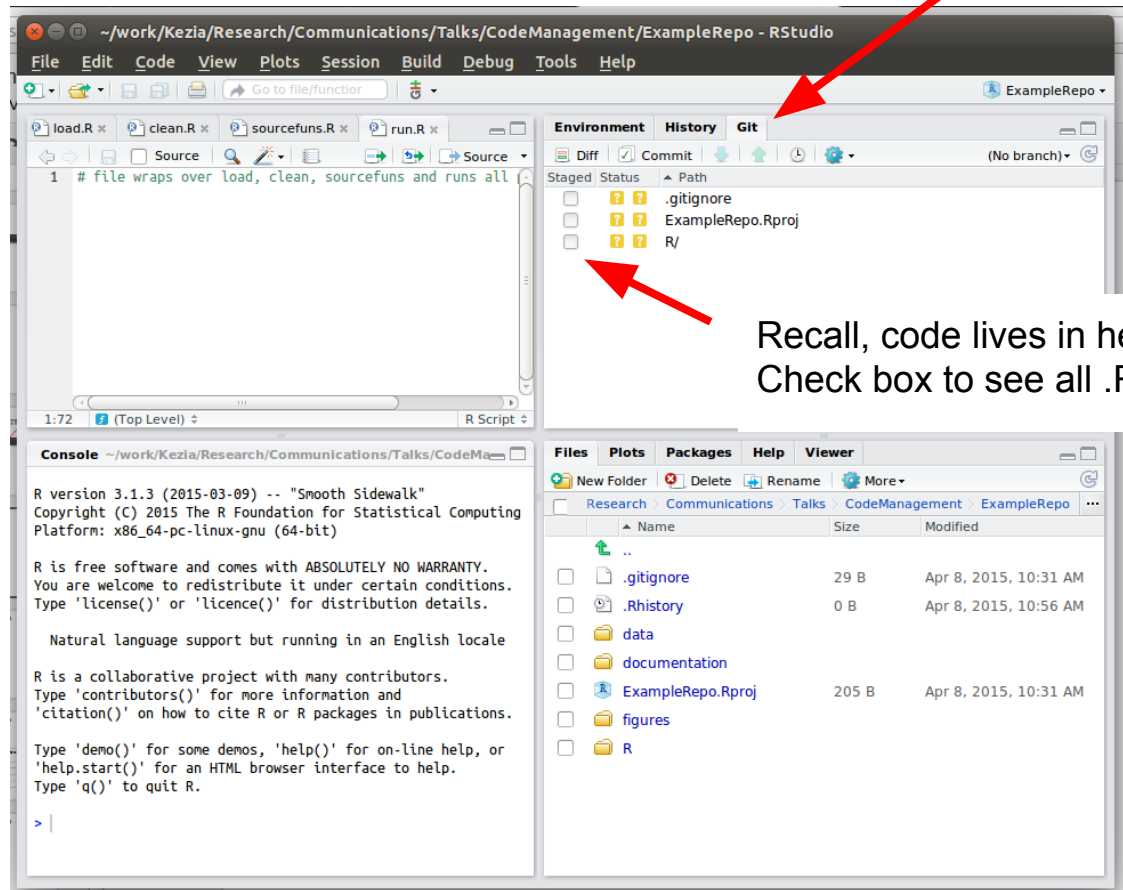
run.R ----- runs analysis functions / produces output

These are R files, so they sit
in R folder of project



Git from RStudio

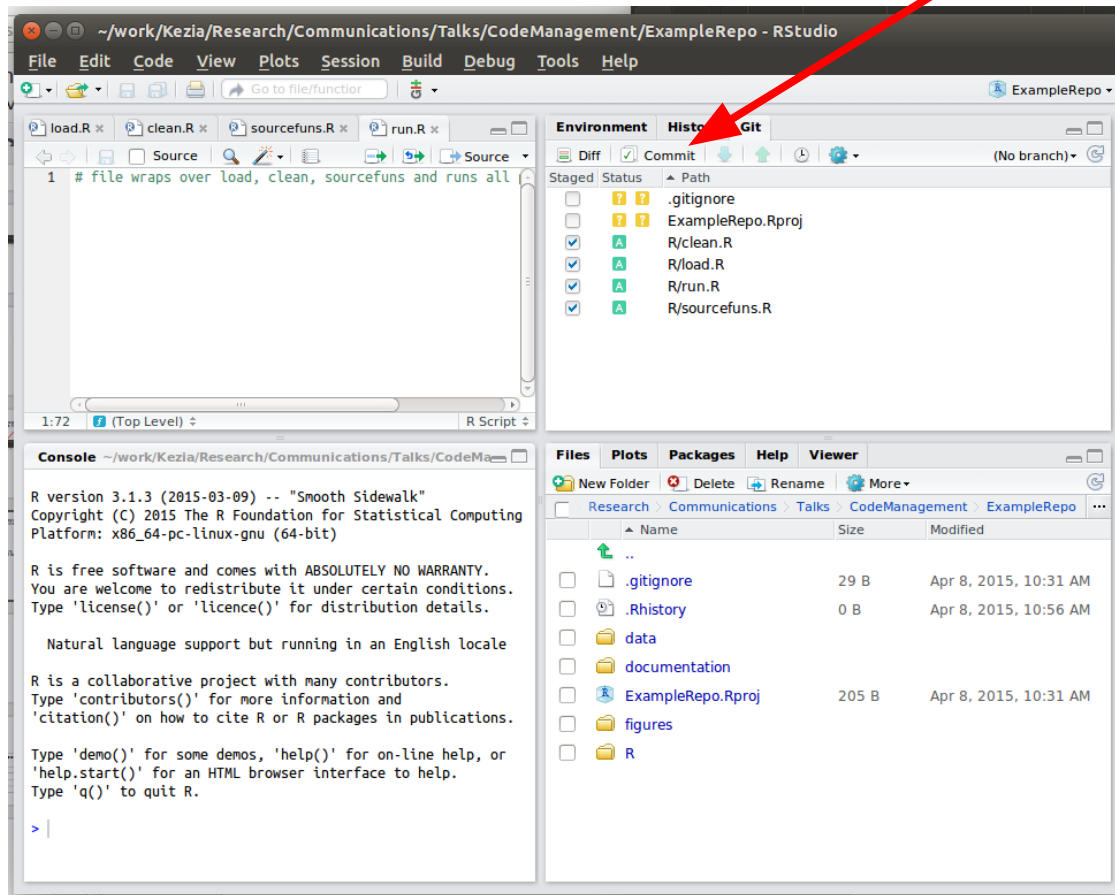
communicate with git through
Git tab



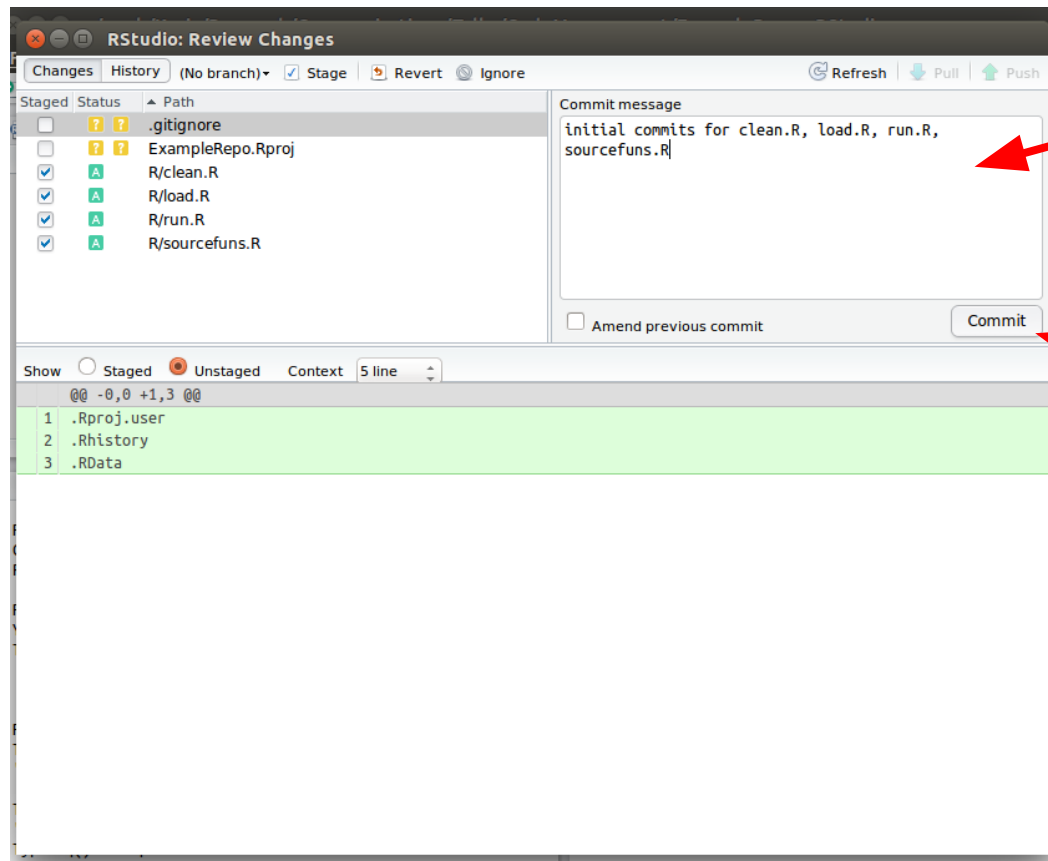
Recall, code lives in here.
Check box to see all .R files

Git from RStudio

Click “Commit” to save code revisions for selected files.



Git from RStudio



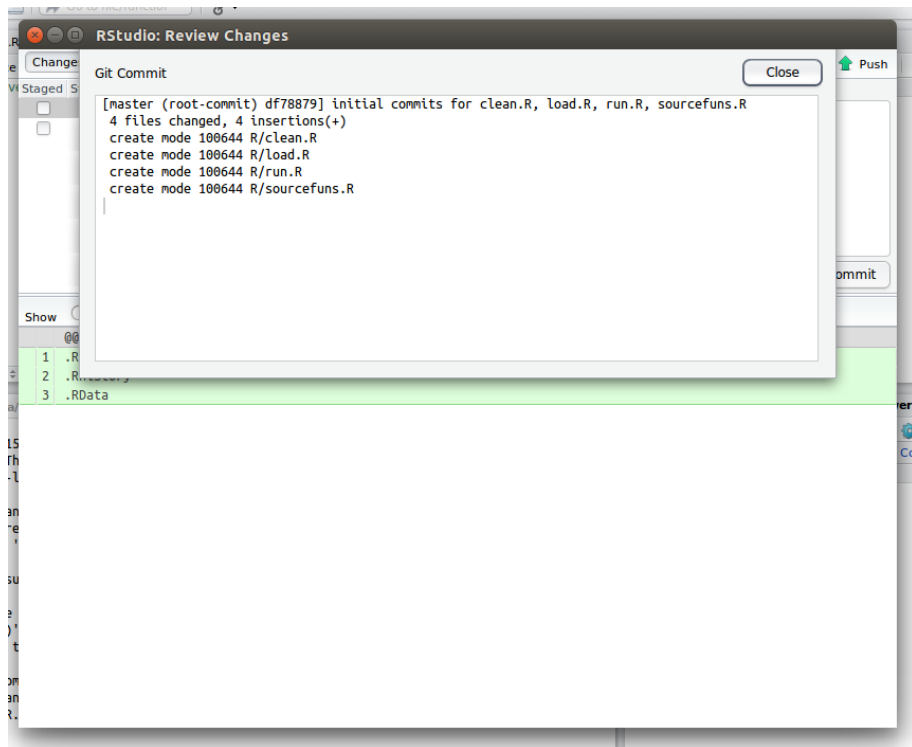
Type comments on revisions here.

Click "commit" to save changes and message

Committing saves changes locally (NOT to github).

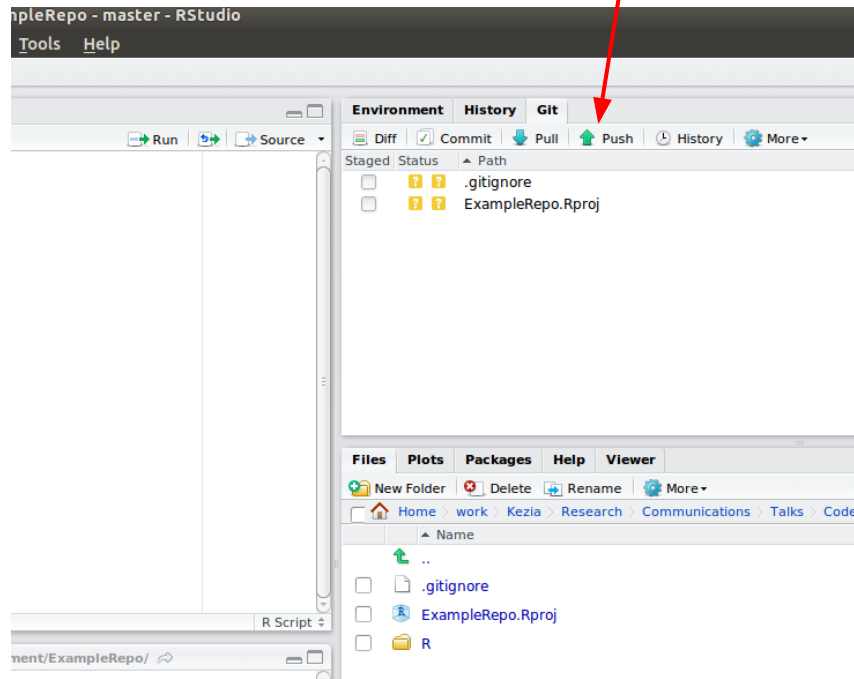
Git from RStudio

Post-committing dialogue



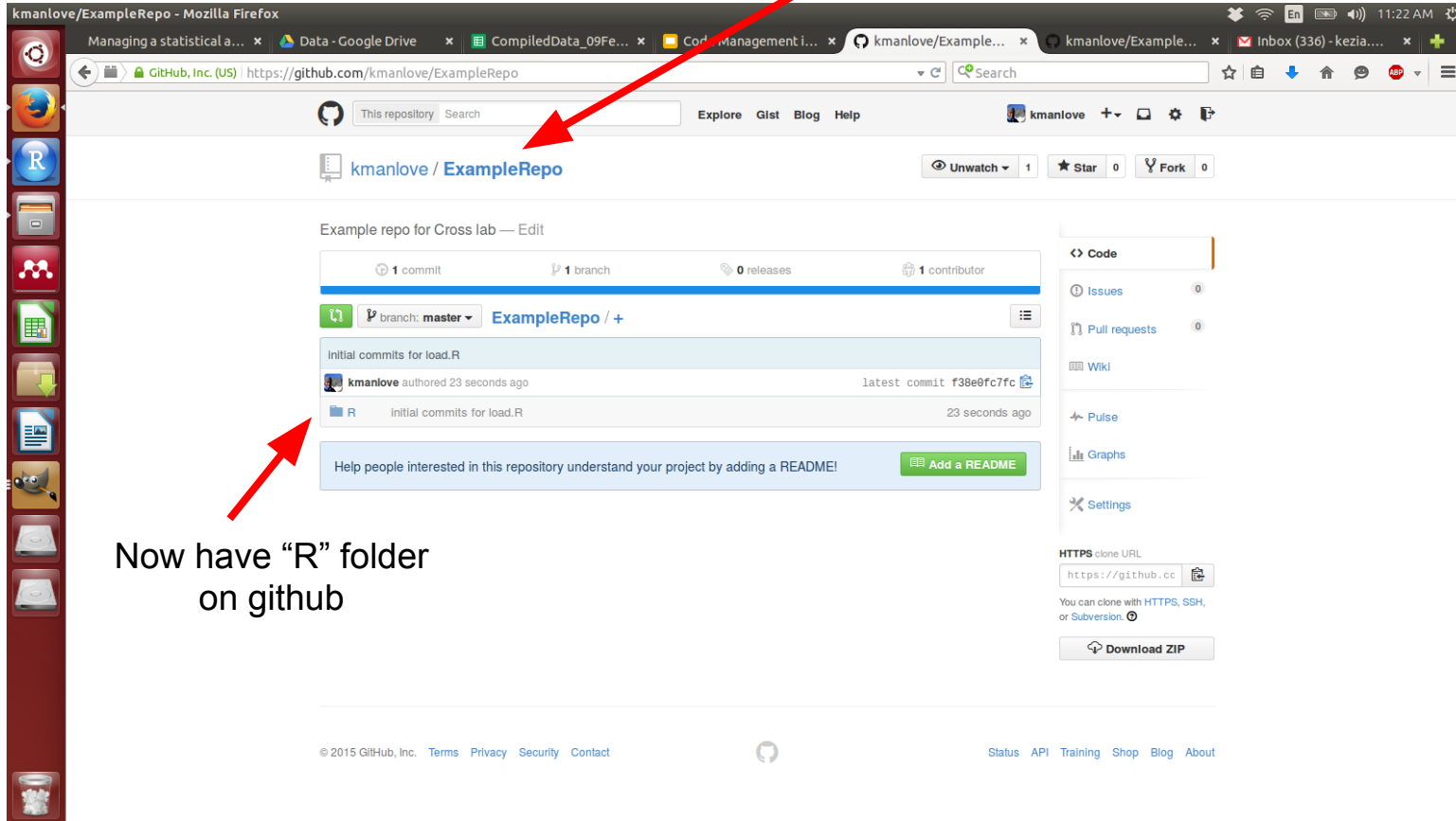
“Push” pushes local commits to github

“Push” option now active in Git window. Click it.



Git from RStudio

Check to be sure you're
in the right repo



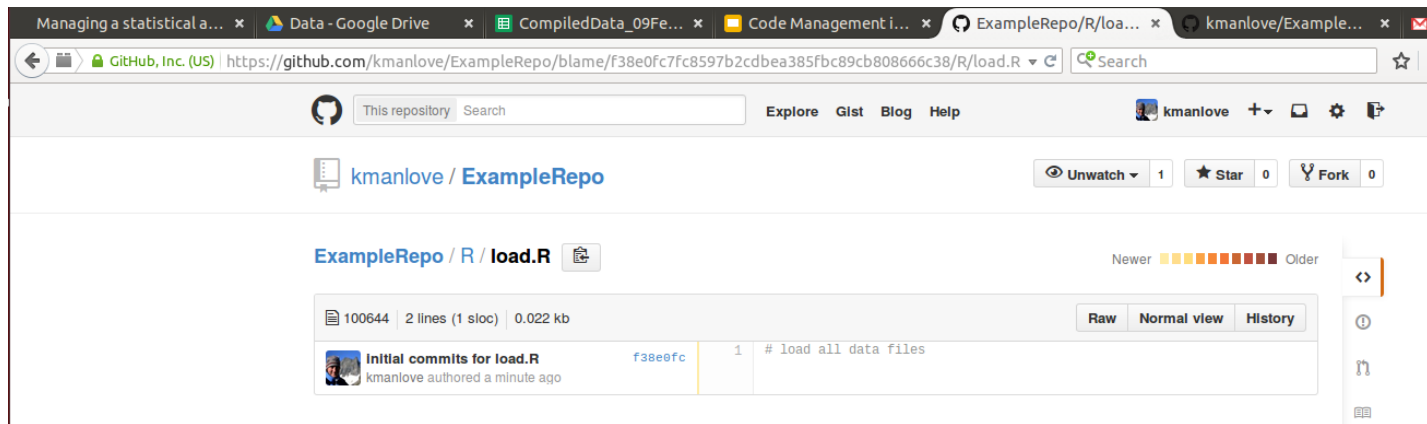
The screenshot shows a web browser window with the URL `https://github.com/kmanlove/ExampleRepo`. The repository page for `kmanlove/ExampleRepo` is displayed. The repository has 1 commit, 1 branch (master), 0 releases, and 1 contributor. The commit history shows an initial commit for `load.R` by `kmanlove` 23 seconds ago. The sidebar on the right contains links for Code, Issues, Pull requests, Wiki, Pulse, Graphs, and Settings. The bottom of the page shows the GitHub footer with copyright information and links for Terms, Privacy, Security, and Contact.

Now have "R" folder
on github

Git from RStudio

To see full commit history for file on github

1. Click R folder
2. Click on <file name> ,
3. Click Blame



Code management suggestions

- Set project structure you like
 - write it down
 - apply to all (future) projects

Code management suggestions

- Set project structure you like
 - write it down
 - apply to all (future) projects

Example

All projects have these subdirectories	All R files have these scripts
data (contains all requisite csvs)	load.R
R (contains all R code)	clean.R
documentation (contains ms/presentations)	sourcefuns.R
figures	run.R

Code management suggestions

- Set project structure you like
 - write it down
 - apply to all (future) projects
- Compartmentalize code
 - Easier to annotate
 - More transferable / reproducible

Code management suggestions

- Set project structure you like
 - write it down
 - apply to all (future) projects
- Compartmentalize code
 - Easier to annotate
 - More transferable / reproducible
- Modify style to match style guide
 - Pick a few elements to begin (recommend: spacing and indentation)

Code management suggestions

- Set project structure you like
 - write it down
 - apply to all (future) projects
- Compartmentalize code
 - Easier to annotate
 - More transferable / reproducible
- Modify style to match style guide
 - Pick a few elements to begin (recommend: spacing and indentation)
- Incorporate git / github / bitbucket into future workflow

Useful links

Overall

- <http://www.r-statistics.com/2010/09/managing-a-statistical-analysis-project-guidelines-and-best-practices/#more-556>

R style guides

- Google: <https://google-styleguide.googlecode.com/svn/trunk/Rguide.xml#spacing>
- Wickham : <http://adv-r.had.co.nz/Style.html>

Project structure

- RStudio: <http://nicercode.github.io/blog/2013-04-05-projects/>
- Stack: <http://stackoverflow.com/questions/1429907/workflow-for-statistical-analysis-and-report-writing/1434424#1434424>

Git/Github + Rstudio

- <https://support.rstudio.com/hc/en-us/articles/200532077-Version-Control-with-Git-and-SVN>
- <http://christianlemp.com/blog/2014/02/13/How-I-Manage-Data-Projects-with-RStudio-and-Git-Part-2.html>