

# Introduction to Reproducible Research in R and R Studio.

Susan Johnston

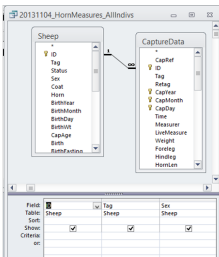
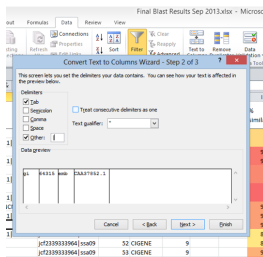
April 23, 2015

# What is Reproducible Research?

**Reproducibility** *is the ability of an entire experiment or study to be reproduced, either by the researcher or by someone else working independently, [and] is one of the main principles of the scientific method.*

Wikipedia

# Many of us are used to clicking, copying and pasting...



## Haggis population density in the Scottish Highlands

S Johnston, University of Edinburgh.

### Introduction.

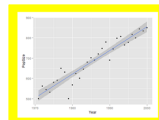


Figure 1: Linear regression of haggis population size and year.

The haggis is a common pest species in the Scottish Highlands. Haggis population densities were recorded annually from 1970 to 2000. We found that the haggis population size increased over this period by  $15.67 \text{ individuals year}^{-1}$  ( $P < 0.001$ , Figure 1).

- Can you repeat all of this again...
  - When new data is added to the dataset?
  - When a journal editor wants you to change a model parameter?
  - When you find an error, but not sure where you went wrong?
- ...and would you get the same results every time?

# Worst Case Scenario

## Retraction Watch

### Archive for the 'not reproducible' Category

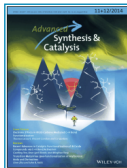
#### Molecular mixup burns chemistry paper

without comments

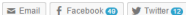
Chemists at Lanzhou University in China did the right thing last month, retracting a [paper](#) in *Advanced Synthesis & Catalysis* because of issues with a reagent that could only be corrected by changing "all the text and quantities."

When the scientists were adding what was labeled Reactant 1 to the mix, they believed it was  $\alpha$ -ethoxycarbonyl- $\alpha$ -azido-*N*-phenylacetamides. Unfortunately, what they were actually using was a decomposed version of the molecule, which threw everything off.

Here's the [notice](#) for "tert-Butyl Hydroperoxide and Tetrabutylammonium Iodide– Promoted Free Radical Cyclization of  $\alpha$ -Azido-*N*-arylamides": [Read the rest of this entry »](#)



Share this:



Written by Cat Ferguson  
April 14th, 2015 at 11:30 am

Posted in [Advanced Synthesis and Catalysis chemistry retractions china](#)  
[retractions doing the right thing freely available not reproducible wiley](#)

#### Two more retractions bring lab break-in biochemist up to eleven

without comments

[Karel Bezouška](#), the Czech biochemist who was caught on hidden camera breaking into a lab fridge to fake results, has [turned it up to eleven](#) with two new retractions.

Both retractions appeared in *Biochemical and Biophysical Research Communications*, one in October 2014 and one in January 2015. His story began two decades ago in 1994, when he published a paper in *Nature* that couldn't be reproduced, and was [eventually retracted in 2013](#).

The best part of the story, of course, is that when his university was attempting to recreate his experiments, Bezouška broke into a lab fridge to tamper with the experiments. Unbeknownst to him, he was caught on hidden camera. [Read the rest of this entry »](#)



Share this:

Tracking retractions as a window into the scientific process

#### Subscribe to Blog

Email Address

Subscribe

#### Pages

[How you can support](#)

[Retraction Watch](#)

[Meet the Retraction Watch staff](#)

[About Adam Marcus](#)

[About Ivan Oransky](#)

[The Center For Scientific Integrity](#)

[Board of Directors](#)

[The Retraction Watch FAQ, including comments policy](#)

[The Retraction Watch Transparency Index](#)

[The Retraction Watch Store](#)

[Upcoming Retraction Watch appearances](#)

[What people are saying about Retraction Watch](#)

Search for:

Search

Search

#### Recent Posts

[Beleaguered Forster turns down prestigious professorship, citing personal toll](#)

#### We're on Facebook



Like

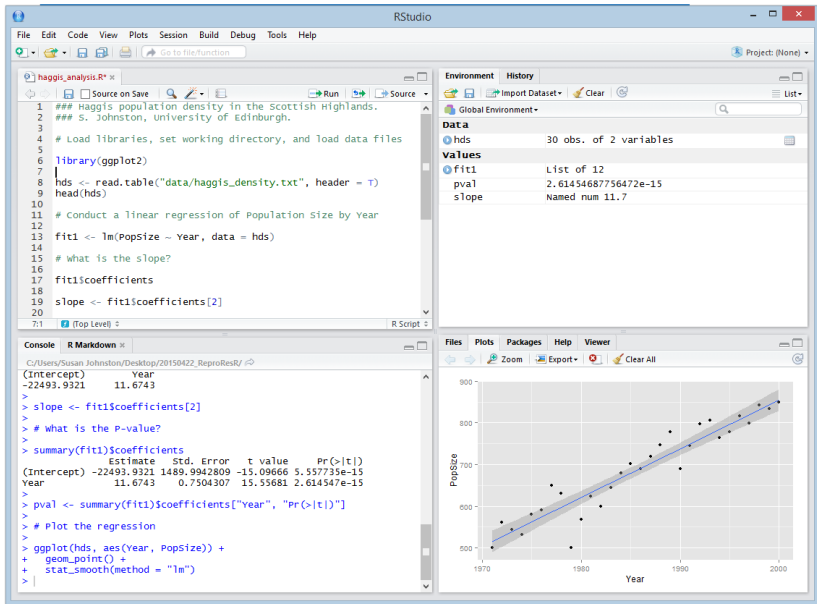
3,135

Retraction Watch

# Three rules for reproducibility

1. Avoid manual data manipulation steps.
  - ▶ Automate analyses using scripts
  - ▶ "This is R. There is no 'if'. Only 'how'."
2. Connect textual statements to underlying results.
  - ▶ Create documents where results and figures update automatically.
3. Version control all custom scripts and documents.
  - ▶ It should be possible to revert to a previous version of an analysis.

# The R Studio Environment: <http://www.rstudio.com>



# Reproducible Research in Studio

1. Creating a Portable Project (.Rproj)
2. Dynamic report writing with R Markdown and `knitr`
3. Version control using `git`

# Reproducible Research in Studio.

1. Creating a Portable Project (.Rproj)
2. Dynamic report writing with R Markdown and `knitr`
3. Version control using `git`



# Bad habits can hinder portability.

<https://support.rstudio.com/hc/en-us/articles/200526207-Using-Projects>



**Hadley Wickham** @hadleywickham · Jan 27

Never plan on sharing your code? Using `setwd()` means that old code will break if you ever reorganise your directories



7



6



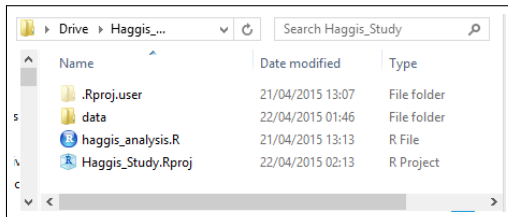
```
setwd("C:/Users/susjoh/Desktop/SalmoAnalysis")
setwd("C:/Users/Susan Johnston/Desktop/SalmoAnalysis")
setwd("C:/Users/Susan Johnston/Drive/SalmoAnalysis")

source("../../OvisAnalysis/GWASplotfunctions.R")
```

An analysis should be contained within a directory, and it should be easy to move it or pass on to someone new.

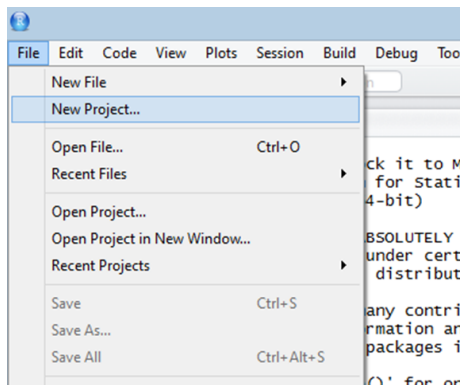
# Solution: using R Studio Projects.

<https://support.rstudio.com/hc/en-us/articles/200526207-Using-Projects>

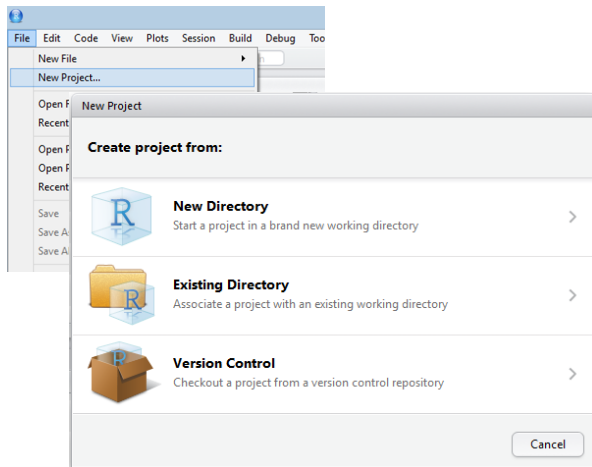


- ▶ Establishes a directory with associated .Rproj file.
- ▶ Automatically sets the working directory.
- ▶ Can save and source .Rprofile, .Rhistory, .Rdata files.
- ▶ Allows version control within R Studio.

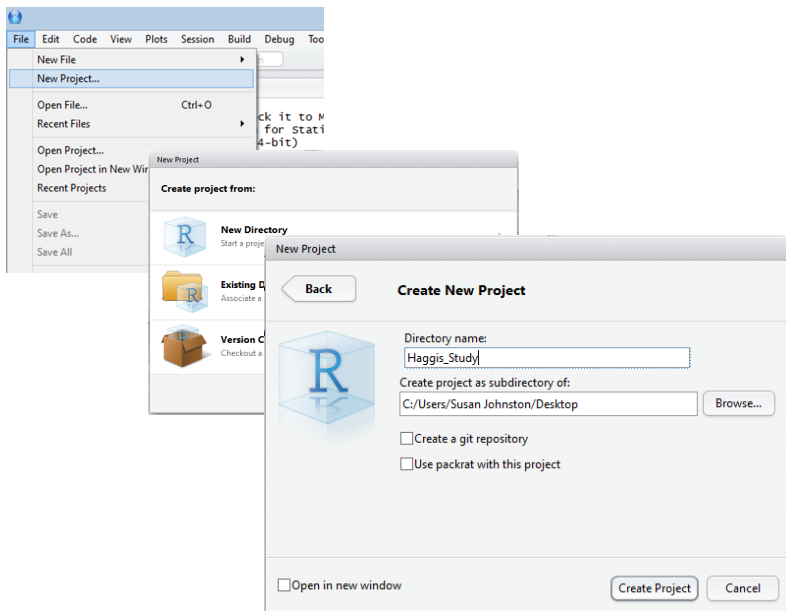
# Creating a Portable Project (.Rproj)



# Creating a Portable Project (.Rproj)



# Creating a Portable Project (.Rproj)



# Creating a Portable Project (.Rproj)

The screenshot displays the RStudio interface for a project named "Haggis\_Study" located at `C:/Users/Susan Johnston/Desktop/Haggis_Study`. The interface is divided into several panes:

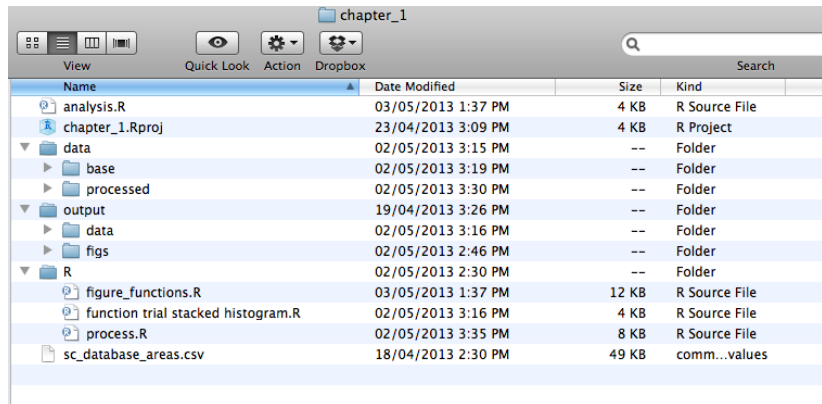
- Source Editor:** Contains an R script with the following code:

```
1 ### Haggis population density in the Scottish Highlands  
2 ### S. Johnston, University of Edinburgh.  
3  
4 # Load libraries, set working directory, and load data  
5  
6 library(ggplot2)  
7  
8 hds <- read.table("data/haggis_density.txt", header = TRUE)  
9 head(hds)  
10  
11 # Conduct a linear regression of Population Size vs. Area  
12  
13
```
- Environment:** Shows the "Global Environment" with the message "Environment is empty".
- Console:** Shows the current working directory as `C:/Users/Susan Johnston/Desktop/Haggis_Study/` and a prompt `> |`.
- Files:** Displays the project's file structure:

	Name	Size	Modified
	..		
	data		
	haggis_analysis.R	665 B	Apr 21, 2015, 1:13 PM
	Haggis_Study.Rproj	218 B	Apr 21, 2015, 1:07 PM

# Maintaining a portable R Project.

<http://nicercode.github.io/blog/2013-05-17-organising-my-project/>



Name	Date Modified	Size	Kind
analysis.R	03/05/2013 1:37 PM	4 KB	R Source File
chapter_1.Rproj	23/04/2013 3:09 PM	4 KB	R Project
data	02/05/2013 3:15 PM	--	Folder
base	02/05/2013 3:19 PM	--	Folder
processed	02/05/2013 3:30 PM	--	Folder
output	19/04/2013 3:26 PM	--	Folder
data	02/05/2013 3:16 PM	--	Folder
figs	02/05/2013 2:46 PM	--	Folder
R	02/05/2013 2:30 PM	--	Folder
figure_functions.R	03/05/2013 1:37 PM	12 KB	R Source File
function trial stacked histogram.R	02/05/2013 3:16 PM	4 KB	R Source File
process.R	02/05/2013 3:35 PM	8 KB	R Source File
sc_database_areas.csv	18/04/2013 2:30 PM	49 KB	comm...values

All data, scripts and output should be kept within the project directory.

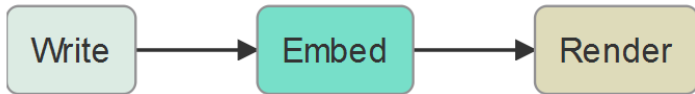
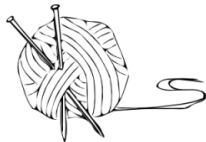
# Reproducible Research in Studio.

1. Creating a Portable Project (`.Rproj`)
2. Dynamic report writing with R Markdown and `knitr`
3. Version control using `git`



# knitr

Elegant, flexible and fast  
dynamic report generation with R



The knitr package allows R code and document templates to be compiled into a single report containing text, results and figures.



## Write reports directly in R

```
1 ## Example of using "R Markdown" and 'knitr'
2
3 This is an R Markdown Document. Markdown is a
4 simple formatting syntax for authoring
5 documents. For example, text can be emphasise
6 d with italics and bold.
7
8 wrap code to be evaluated in R in an R
9 "chunk"
10
11 ```{r}
12 nrow(cars)|
13 ```
14
15 Code can be evaluated in-line using backticks
16 . For example, we can write 1 + 1 = `r 1+1`.
17
18 Plots will also be printed.
19
20 ```{r fig.width = 3, fig.height = 3}
21 library(ggplot2)
22 ggplot(cars, aes(speed, dist)) +
23   geom_point()
24 ```
```

# knitr

Elegant, flexible and fast  
dynamic report generation with R



## Write reports directly in R

example.html | Open in Browser | Publish | Find

### Example of using "R Markdown" and `knitr`

This is an R Markdown Document. Markdown is a simple formatting syntax for authoring documents. For example, text can be emphasised with *italics* and **bold**.

Wrap code to be evaluated in R in an R "chunk"

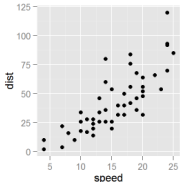
```
nrow(cars)
```

```
## [1] 50
```

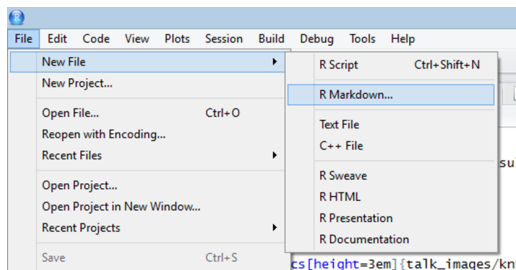
Code can be evaluated in-line using backticks. For example, we can write `1 + 1 = 2`.

Plots will also be printed.

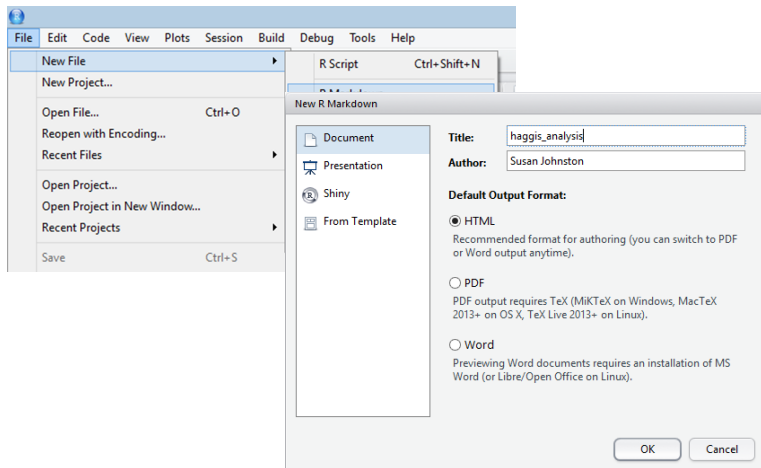
```
library(ggplot2)
ggplot(cars, aes(speed, dist)) +
  geom_point()
```



# Creating an R Markdown Script (.Rmd).



# Creating an R Markdown Script (.Rmd).



# A Quick Start Guide

<http://nicercode.github.io/guides/reports/>

## 1. Type report text into .Rmd file

```
Lorem ipsum dolor sit amet, consectetur adipiscing elit.
```

## 2. Enclose code to be evaluated in chunks

```
```{r}  
model1 <- lm(speed ~ dist, data = cars)  
```
```

## 3. Evaluate code inline

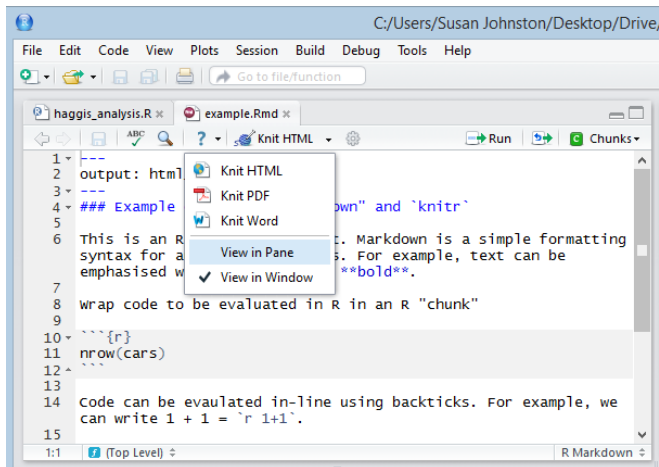
```
The slope of the model is `r coefficients(model1)[2]`
```

*The slope of the model is 0.16557*

## 4. Compile report as .html, .pdf or .doc

# A Quick Start Guide

<http://nicercode.github.io/guides/reports/>

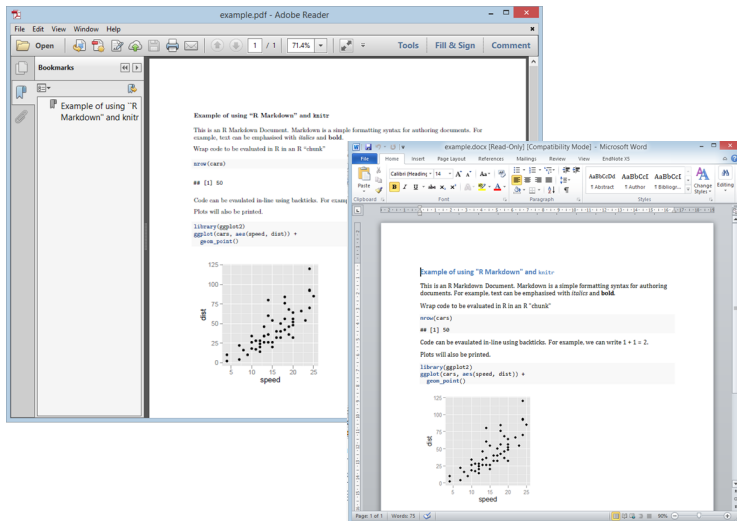


NB. PDF and Word docs require additional software.

<http://rmarkdown.rstudio.com/?version=0.98.1103&mode=desktop>

# A Quick Start Guide

<http://nicercode.github.io/guides/reports/>



<http://rmarkdown.rstudio.com/?version=0.98.1103&mode=desktop>



# Advanced Tips

- ▶ Control how chunks are reported and evaluated


```
```{r echo = F, warning = F, fig.width = 3}  
model1 <- lm(speed ~ dist, data = cars)  
plot(model1)  
```
```

- ▶ `spin()`: compile .R files using `#'` and `#+`  
<http://deanattali.com/2015/03/24/knitr-best-hidden-gem-spin/>
- ▶  $\text{\LaTeX}$  documents, Presentations, Shiny, etc.

# Reproducible Research in Studio.

1. Creating a Portable Project (.Rproj)
2. Dynamic report writing with R Markdown and `knitr`
3. Version control using `git`

# Version Control can revert a document to a previous version.



**WIKIPEDIA**  
The Free Encyclopedia

[Main page](#)  
[Contents](#)  
[Featured content](#)  
[Current events](#)  
[Random article](#)  
[Donate to Wikipedia](#)  
[Wikipedia store](#)

**Interaction**  
[Help](#)  
[About Wikipedia](#)  
[Community portal](#)  
[Recent changes](#)  
[Contact page](#)

**Tools**  
[What links here](#)  
[Related changes](#)  
[Upload file](#)  
[Special pages](#)  
[Permanent link](#)

ArticleTalk

ReadView sourceView history

Search

## Batman


From Wikipedia, the free encyclopedia

*This article is about the superhero character. For other uses, see [Batman \(disambiguation\)](#).*

**Batman** is a fictional [superhero](#) appearing in [American comic books](#) published by [DC Comics](#), as well as its associated media. The character was created by artist [Bob Kane](#) and writer [Bill Finger](#), and first appeared in *[Detective Comics](#)* #27 (May 1939). Originally named "the Bat-Man," the character is also referred to by such epithets as "the Caped Crusader,"<sup>[5]</sup> "the Dark Knight,"<sup>[5]</sup> and "the World's Greatest Detective."<sup>[5]</sup>

Batman is the [secret identity](#) of **Bruce Wayne**, an American billionaire, industrialist, and [philanthropist](#). Having witnessed the murder of his parents as a child, he swore revenge on criminals, an oath tempered with a sense of justice. Wayne trains himself both physically and intellectually and dons a [bat-themed costume](#) to fight crime.<sup>[6]</sup> Batman operates in the fictional [Gotham City](#), assisted by various supporting characters including his butler [Alfred Pennyworth](#), his crime-fighting

### Batman



Version Control can revert a document to a previous version.

[Recent changes](#) | [minor edit](#), → = [section edit](#), ← = [automatic edit summary](#)

([newest](#) | [oldest](#)) View ([newer 50](#) | [older 50](#)) ([20](#) | [50](#) | [100](#) | [250](#) | [500](#))


[Compare selected revisions](#)

- ([cur](#) | [prev](#)) • [14:23, 15 April 2015 Kinfol1993 \(talk | contribs\) m . . \(118,704 bytes\) \(+10\)](#) . .  
([→Personality](#))
- ([cur](#) | [prev](#)) • [14:20, 15 April 2015 Kinfol1993 \(talk | contribs\) m . . \(118,694 bytes\) \(+5\)](#) . .  
([→Personality](#))
- ([cur](#) | [prev](#)) • [14:20, 15 April 2015 Kinfol1993 \(talk | contribs\) m . . \(118,689 bytes\) \(+1\)](#) . .  
([→Personality](#))
- ([cur](#) | [prev](#)) • [14:19, 15 April 2015 Kinfol1993 \(talk | contribs\) . . \(118,688 bytes\) \(+1,063\)](#) . .  
([→Personality: More info on Miller's Batman, original personality](#))
- ([cur](#) | [prev](#)) • [13:25, 15 April 2015 Kinfol1993 \(talk | contribs\) m . . \(117,625 bytes\) \(+34\)](#) . .  
*(Category:fictional antiheroes per citations at list of fictional antiheroes)*
- ([cur](#) | [prev](#)) • [10:58, 15 April 2015 ProspectofArts \(talk | contribs\) . . \(117,591 bytes\) \(+4\)](#)
- ([cur](#) | [prev](#)) • [10:57, 15 April 2015 ProspectofArts \(talk | contribs\) . . \(117,587 bytes\) \(+67\)](#)
- ([cur](#) | [prev](#)) • [15:22, 9 April 2015 JosephSpiral \(talk | contribs\) . . \(117,520 bytes\) \(+2\)](#)
- ([cur](#) | [prev](#)) • [11:24, 7 April 2015 Brandmeister \(talk | contribs\) . . \(117,518 bytes\) \(-2,880\)](#) . . *(more concise plot per tag)*
- ([cur](#) | [prev](#)) • [23:42, 4 April 2015 Samhan Dobo \(talk | contribs\) . . \(120,398 bytes\) \(+4\)](#) . . ([→Skills and abilities](#))
- ([cur](#) | [prev](#)) • [14:25, 3 April 2015 JosephSpiral \(talk | contribs\) . . \(120,394 bytes\) \(0\)](#) . . ([→Romantic relationships](#))
- ([cur](#) | [prev](#)) • [14:24, 3 April 2015 JosephSpiral \(talk | contribs\) . . \(120,394 bytes\) \(+21\)](#) . . ([→Romantic](#)

Version Control can revert a document to a previous version.

[illegible]

# Version Control can revert a document to a previous version.



**WIKIPEDIA**  
The Free Encyclopedia

[Main page](#)  
[Contents](#)  
[Featured content](#)  
[Current events](#)  
[Random article](#)  
[Donate to Wikipedia](#)  
[Wikipedia store](#)

**Interaction**  
[Help](#)  
[About Wikipedia](#)  
[Community portal](#)  
[Recent changes](#)  
[Contact page](#)

**Tools**  
[What links here](#)  
[Related changes](#)  
[Upload file](#)  
[Special pages](#)  
[Permanent link](#)

Create account Log in

ArticleTalk

ReadView sourceView history

## Batman


From Wikipedia, the free encyclopedia

*This article is about the superhero character. For other uses, see [Batman \(disambiguation\)](#).*

**Batman** is a fictional [superhero](#) appearing in [American comic books](#) published by [DC Comics](#), as well as its associated media. The character was created by artist [Bob Kane](#) and writer [Bill Finger](#), and first appeared in *Detective Comics* #27 (May 1939). Originally named "the Bat-Man," the character is also referred to by such epithets as "the Caped Crusader,"<sup>[5]</sup> "the Dark Knight,"<sup>[5]</sup> and "the World's Greatest Detective."<sup>[5]</sup>

Batman is the [secret identity](#) of **Bruce Wayne**, an American billionaire, industrialist, and [philanthropist](#). Having witnessed the murder of his parents as a child, he swore revenge on criminals, an oath tempered with a sense of justice. Wayne trains himself both physically and intellectually and dons a [bat-themed costume](#) to fight crime.<sup>[6]</sup> Batman operates in the fictional [Gotham City](#), assisted by various supporting characters including his butler [Alfred Pennyworth](#), his crime-fighting

**Batman**



# Version Control Using git.

<https://support.rstudio.com/hc/en-us/articles/200532077-Version-Control-with-Git-and-SVN>

A banner for the Git website. On the left, the Git logo (a red diamond with a white 'g') is followed by the text 'git --local-branching-on-the-cheap'. Below this, two paragraphs describe Git as a free and open source distributed version control system designed for speed and efficiency, and as easy to learn with a tiny footprint and lightning fast performance. It mentions that Git outclasses SCM tools like Subversion, CVS, Perforce, and ClearCase. At the bottom left, there is a GitHub logo and the text 'Learn Git in your browser for free with Try Git.' On the right side of the banner, there is a search bar with the placeholder text 'Search entire site...'. Below the search bar is a 3D diagram showing several stacks of papers (representing code repositories) connected by colored lines (red, blue, yellow) in a branching structure, illustrating the concept of local branching.

**git** --local-branching-on-the-cheap

Git is a **free and open source** distributed version control system designed to handle everything from small to very large projects with speed and efficiency.

Git is **easy to learn** and has a **tiny footprint with lightning fast performance**. It outclasses SCM tools like Subversion, CVS, Perforce, and ClearCase with features like **cheap local branching**, convenient staging areas, and **multiple workflows**.

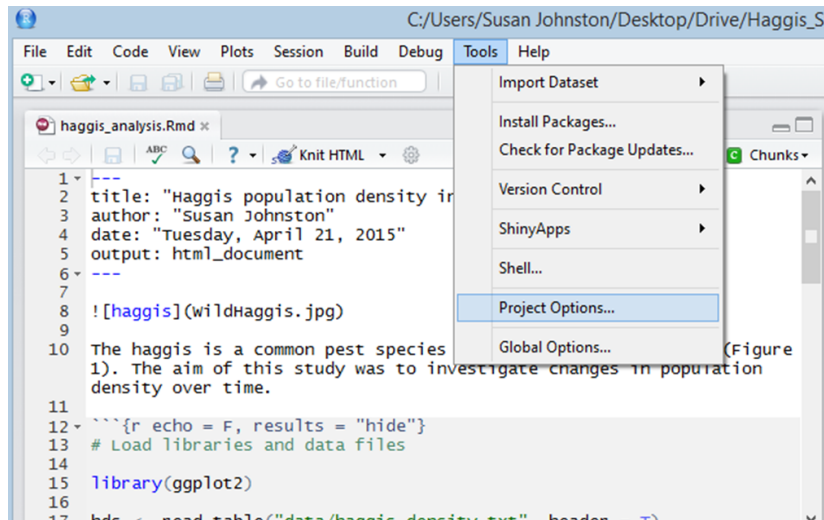
 **Learn Git in your browser for free with Try Git.**

Search entire site...

Git can be installed on all platforms, and can be used to implement version control within an R Studio Project.

<http://git-scm.com/downloads>

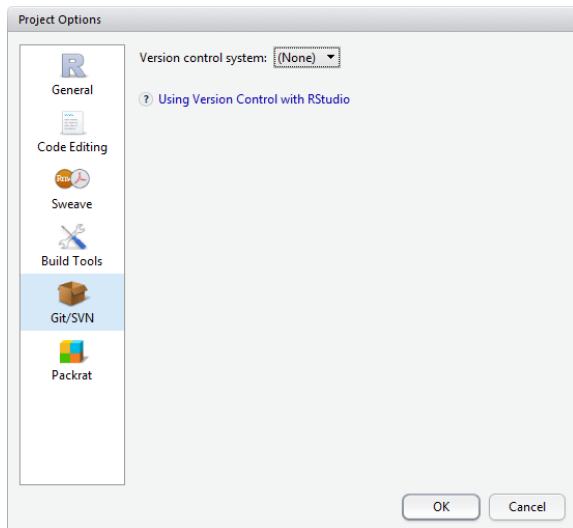
# Version Control in R Studio



Tools > Project Options allows setup of git version control.

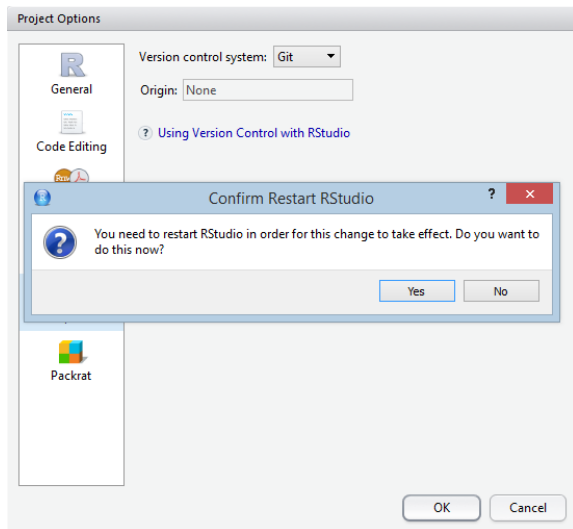


# Version Control in R Studio



Select git as a version control system

# Version Control in R Studio



Select git as a version control system

# Version Control in R Studio

The screenshot displays the RStudio IDE interface for a project named 'Haggis\_Study'. The main editor window shows an R Markdown file 'haggis\_analysis.Rmd' with the following content:

```
1 title: "Haggis population density in the Scottish Highlands"
2 author: "Susan Johnston"
3 date: "Tuesday, April 21, 2015"
4 output: html_document
5 ---
6 {r echo = F, results = "hide"}
7 # Load libraries and data files
8 library(ggplot2)
9
10 hds <- read.table("data/haggis_density.txt", header = T)
11
12 # conduct a linear regression of Population Size by Year
13
14 fit1 <- lm(PopSize ~ Year, data = hds)
15
```

The Console window at the bottom left shows the R version and copyright information:

```
R version 3.1.1 (2014-07-10) -- "Sock it to Me"
Copyright (C) 2014 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

> |
```

The Environment pane on the right shows the project files:

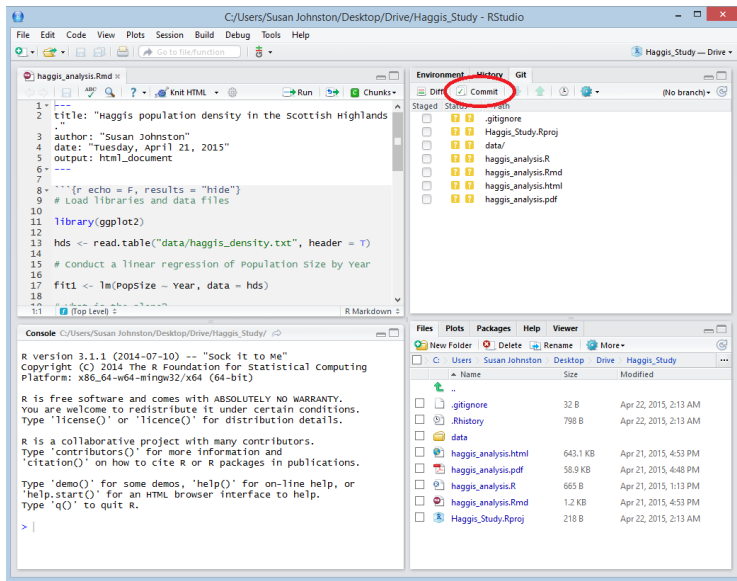
| Staged                   | Status | Path                 |
|--------------------------|--------|----------------------|
| <input type="checkbox"/> |        | .gitignore           |
| <input type="checkbox"/> |        | Haggis_Study.Rproj   |
| <input type="checkbox"/> |        | data/                |
| <input type="checkbox"/> |        | haggis_analysis.R    |
| <input type="checkbox"/> |        | haggis_analysis.Rmd  |
| <input type="checkbox"/> |        | haggis_analysis.html |
| <input type="checkbox"/> |        | haggis_analysis.pdf  |

The Files pane at the bottom right shows the project structure:

| Name                 | Size     | Modified              |
|----------------------|----------|-----------------------|
| .gitignore           | 32 B     | Apr 22, 2015, 2:13 AM |
| .Rhistory            | 798 B    | Apr 22, 2015, 2:13 AM |
| data                 |          |                       |
| haggis_analysis.html | 643.1 KB | Apr 21, 2015, 4:53 PM |
| haggis_analysis.pdf  | 58.9 KB  | Apr 21, 2015, 4:48 PM |
| haggis_analysis.R    | 665 B    | Apr 21, 2015, 1:13 PM |
| haggis_analysis.Rmd  | 1.2 KB   | Apr 21, 2015, 4:53 PM |
| Haggis_Study.Rproj   | 218 B    | Apr 22, 2015, 2:13 AM |

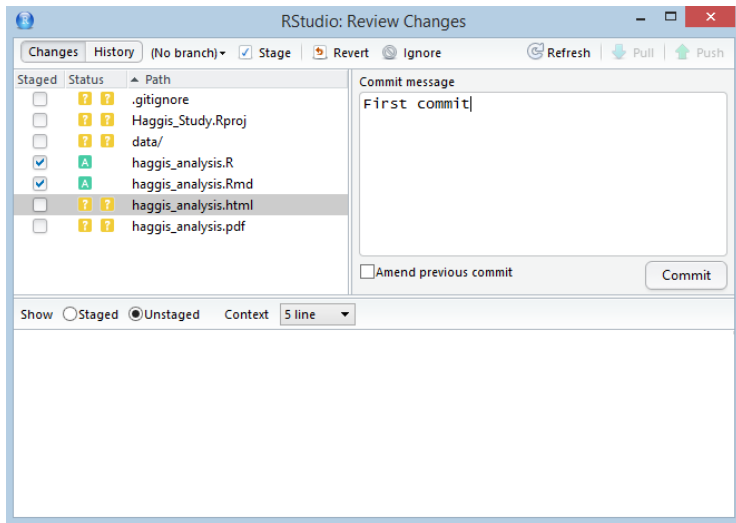
git information will appear in the top-right frame.

# Version Control in R Studio



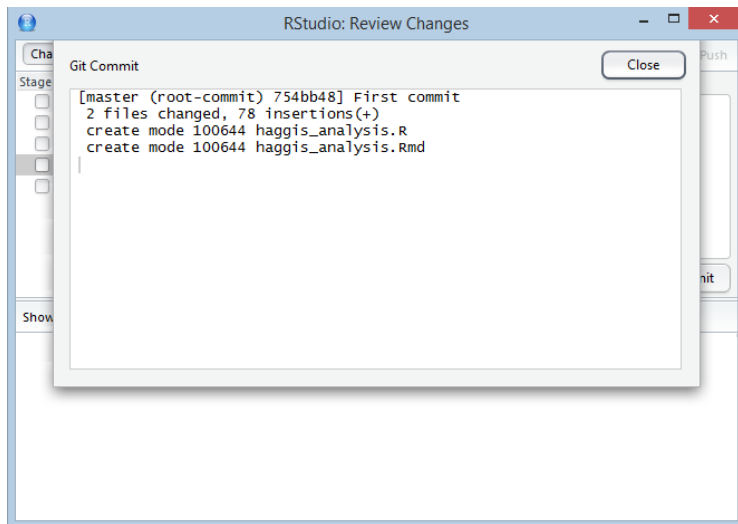
git information will appear in the top-right frame.

# Version Control in R Studio



Select files to version control, write a meaningful commit message  
>Commit

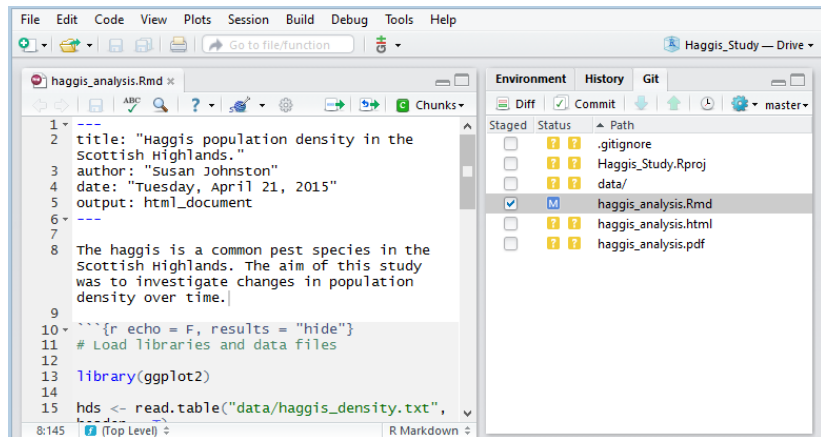
# Version Control in R Studio



Select files to version control, write a meaningful commit message

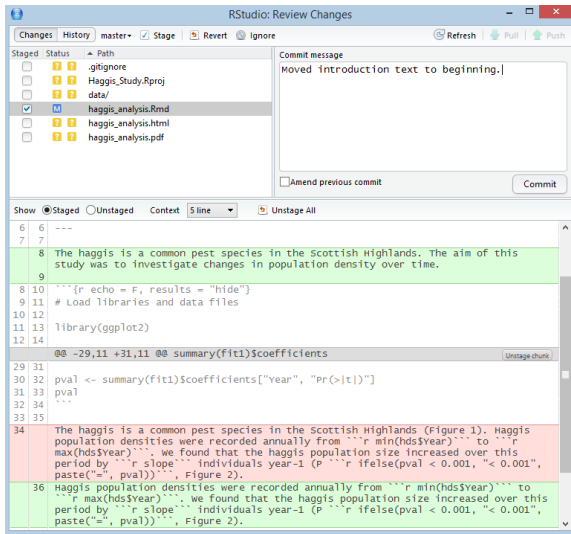
>Commit

# Version Control in R Studio



After modifying the file, repeat the process.

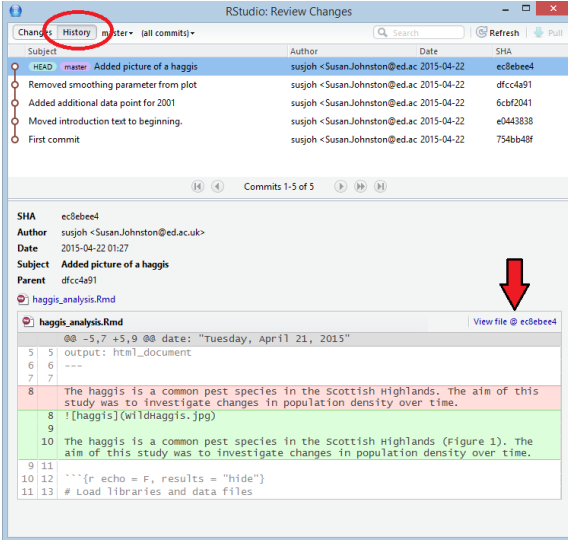
# Version Control in R Studio



After modifying the file, repeat the process.



# Version Control in R Studio



The screenshot shows the RStudio 'Review Changes' interface. At the top, the 'History' tab is selected and circled in red. Below it is a table of commit history:

| Subject                               | Author                        | Date       | SHA      |
|---------------------------------------|-------------------------------|------------|----------|
| HEAD master Added picture of a haggis | susjoh <Susan.Johnston@ed.ac> | 2015-04-22 | ec8ebee4 |
| Removed smoothing parameter from plot | susjoh <Susan.Johnston@ed.ac> | 2015-04-22 | dfcc4a91 |
| Added additional data point for 2001  | susjoh <Susan.Johnston@ed.ac> | 2015-04-22 | 6cbf2041 |
| Moved introduction text to beginning. | susjoh <Susan.Johnston@ed.ac> | 2015-04-22 | e0443838 |
| First commit                          | susjoh <Susan.Johnston@ed.ac> | 2015-04-22 | 754bb48f |

Below the table, the selected commit (SHA: ec8ebee4) is detailed:

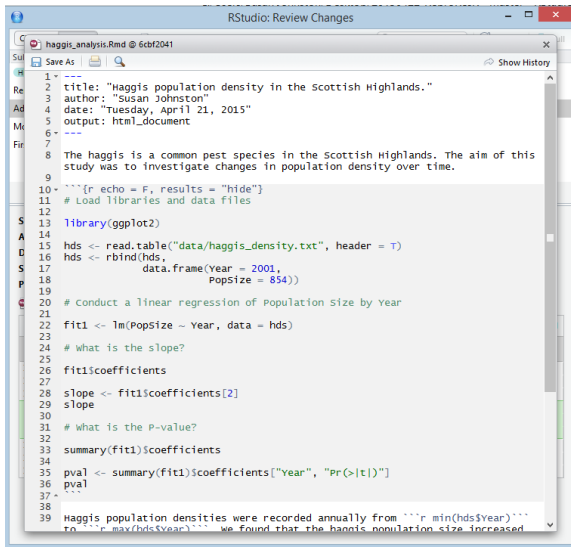
**SHA** ec8ebee4  
**Author** susjoh <Susan.Johnston@ed.ac.uk>  
**Date** 2015-04-22 01:27  
**Subject** Added picture of a haggis  
**Parent** dfcc4a91  
**File** haggis\_analysis.Rmd

A red arrow points to the 'View file @ ec8ebee4' link. The file content is shown in a diff view:

```
@@ -5,7 +5,9 @@ date: "Tuesday, April 21, 2015"
5 5 output: html_document
6 6 ---
7 7
8 8 The haggis is a common pest species in the Scottish Highlands. The aim of this
9 9 study was to investigate changes in population density over time.
10 10 ![[haggis]](wildhaggis.jpg)
11 11
12 12 The haggis is a common pest species in the Scottish Highlands (Figure 1). The
13 13 aim of this study was to investigate changes in population density over time.
14 14
15 15 ```{r echo = F, results = "hide"}
16 16 # Load libraries and data files
```

Previous versions can be viewed and restored from the History tab.

# Version Control in R Studio

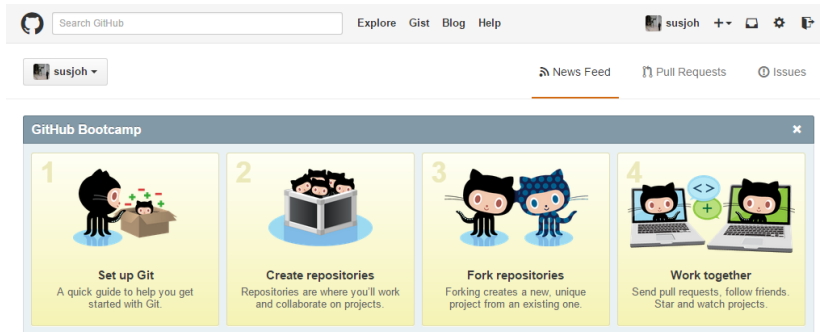


The screenshot shows the RStudio interface. The main window is titled 'RStudio: Review Changes'. Below the title bar, there's a file explorer showing 'haggis\_analysis.Rmd @ 6cbf2041'. The script editor displays R code for a document. The code includes a title, author, date, and output format. It then loads the 'ggplot2' library, reads a data file 'haggis\_density.txt', and performs a linear regression of population size by year. The code also calculates the slope and p-value for the regression. The final output is a text block summarizing the findings.

```
1 ---
2 title: "Haggis population density in the Scottish Highlands."
3 author: "Susan Johnston"
4 date: "Tuesday, April 21, 2015"
5 output: html_document
6 ---
7
8 The haggis is a common pest species in the Scottish Highlands. The aim of this
9 study was to investigate changes in population density over time.
10
11 ```{r echo = F, results = "hide"}
12 # Load libraries and data files
13 library(ggplot2)
14
15 hds <- read.table("data/haggis_density.txt", header = T)
16 hds <- rbind(hds,
17             data.frame(Year = 2001,
18                       PopSize = 854))
19
20 # Conduct a linear regression of Population Size by Year
21
22 fit1 <- lm(PopSize ~ Year, data = hds)
23
24 # What is the slope?
25
26 fit1$coefficients
27
28 slope <- fit1$coefficients[2]
29 slope
30
31 # What is the P-value?
32
33 summary(fit1)$coefficients
34
35 pval <- summary(fit1)$coefficients["Year", "Pr(>|t|)"]
36 pval
37 ```
38
39 Haggis population densities were recorded annually from ```r min(hds$Year)```
40 to ```r max(hds$Year)``` we found that the haggis population size increased
```

Previous versions can be viewed and restored from the History tab.

# Advanced Steps: Github



- ▶ Forking projects
- ▶ All scripts are backed up online
- ▶ Facilitates collaboration and working on different computers

# Online Resources

- ▶ RStudio: Idiot-proof guides and cheat sheets  
<http://www.rstudio.com/>
- ▶ Nice R Code: How-tos and advice on good coding practice  
<http://nicercode.github.io/guide.html>
- ▶ Ten Simple Rules for Reproducible Computational Research  
<http://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1003285>
- ▶ Yihui Xie's blog (knitr) <http://yihui.name/en/categories/>
- ▶ R Bloggers: <http://www.r-bloggers.com/>
- ▶ StackOverflow questions on R and knitr  
<http://stackoverflow.com/questions/tagged/r+knitr>