

Сергей Мастицкий

По материалам сайтов <http://rstudio.org> и <http://r-analytics.blogspot.com>

Работа с RStudio

12 марта

2012

RStudio представляет собой бесплатную интегрированную среду разработки (IDE) для R. Благодаря ряду своих особенностей этот активно развивающийся программный продукт делает работу с R очень удобной. Ниже представлено краткое руководство пользователя для версии программы v0.95 (авторский перевод документации с официального сайта проекта, с изменениями и дополнениями).

Перевод
официальной
документации

Содержание

1. Инсталляция	4
2. Работа с Консолью.....	4
2.1. Общая информация.....	4
2.2. Автоматическое завершение кода.....	4
2.3. Возврат к предыдущим командам.....	5
2.4. Заголовок окна Консоли	6
2.5. Горячие клавиши	6
3. Редактирование и выполнение кода.....	6
3.1. Общая информация.....	6
3.2. Управление файлами	7
3.3. Завершение кода.....	8
3.4. Поиск и замена частей кода	8
3.5. Автоматическое создание функций.....	8
3.6. Комментирование кода.....	9
3.7. Выполнение кода	9
3.8. Горячие клавиши	11
4. Навигация по коду.....	11
4.1. Общая информация.....	11
4.2. Переход к файлу или функции	11
4.3. Переход к определению функции	12
4.4. Навигация по скрипту	12
4.5. Двусторонняя навигация по коду.....	13
5. Использование истории команд.....	13
5.1. Общая информация.....	13

5.2. Обзор истории команд.....	14
5.3. Поиск в истории команд.....	14
5.4. Выполнение команд.....	15
6. Рабочие директории и рабочие пространства	16
6.1. Общая информация.....	16
6.2. Рабочая директория по умолчанию.....	16
6.3. Изменение рабочей директории	16
6.4. Загрузка и сохранение рабочих пространств.....	17
7. Проекты	18
7.1. Общая информация.....	18
7.2. Создание проектов	18
7.3. Работа с проектами	19

1. Инсталляция

Инсталлятор RStudio следует [скачать](#) с официального сайта проекта. RStudio можно установить как на персональный компьютер, так и на Linux-управляемый сервер. Ниже речь идет о варианте программы для персональных компьютеров.

Доступны инсталляторы RStudio для всех наиболее распространенных операционных систем, в частности Windows XP/Vista/7, Mac OS X 10.5, Debian 6+/Ubuntu 10.04 и Fedora 13. Выберите из списка соответствующий файл и скачайте его. Инсталляция происходит стандартно, без каких-либо особенностей.

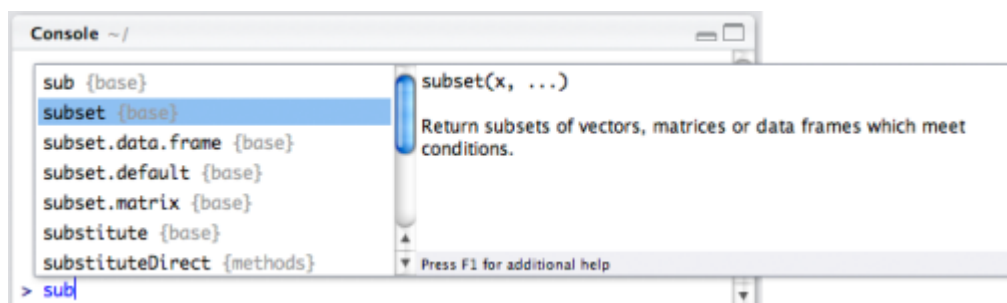
2. Работа с Консолью

2.1. Общая информация

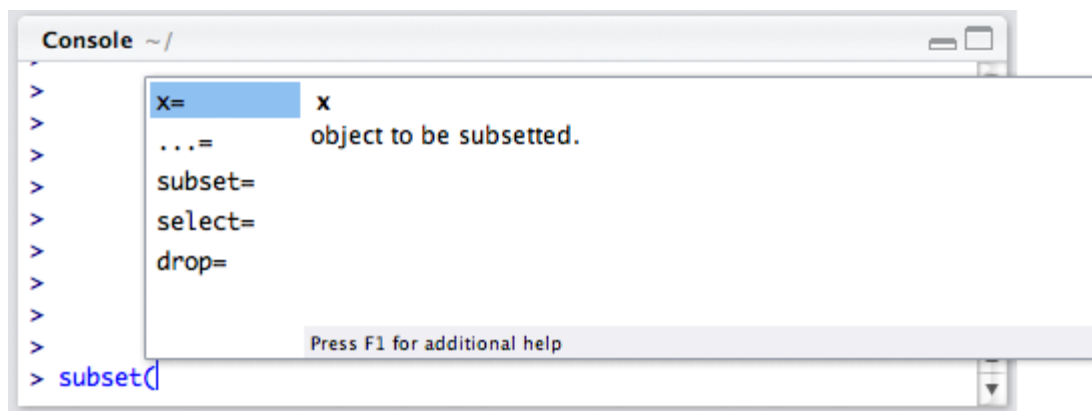
Консоль RStudio (Console) предоставляет целый ряд опций, делающих работу с R простой и продуктивной. Освоение этих опций, наряду с возможностями, доступными в панелях Source (Редактор кода) и History (История), может с лихвой воздать затраченное на обучение время.

2.2. Автоматическое завершение кода

RStudio поддерживает автоматическое завершение кода при помощи клавиши Tab. Например, если в рабочем пространстве имеется объект с именем `pollResults`, то можно набрать на клавиатуре `poll`, нажать Tab, и RStudio автоматически завершит название этого объекта. Аналогично можно получать подсказки по функциям при введении их имен. Например, введя название функции `sub` и нажав на Tab, получим следующее:



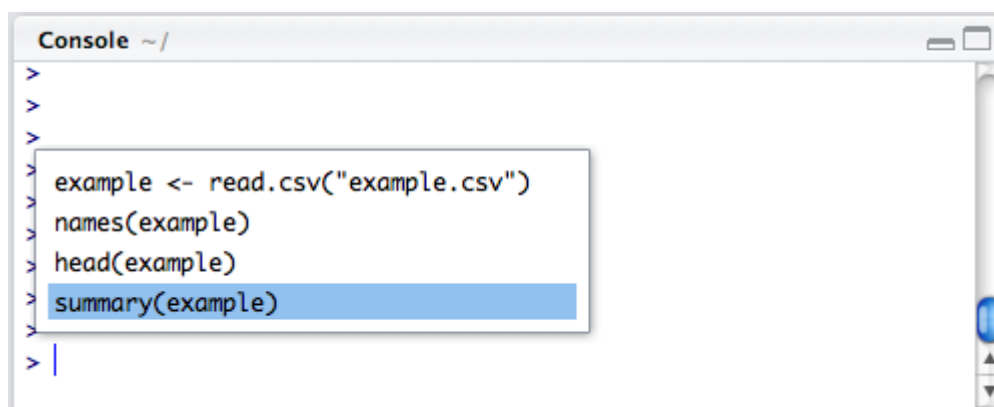
Завершение кода работает также для аргументов функций; например, если ввести `subset (` и нажать `Tab`, то получим:



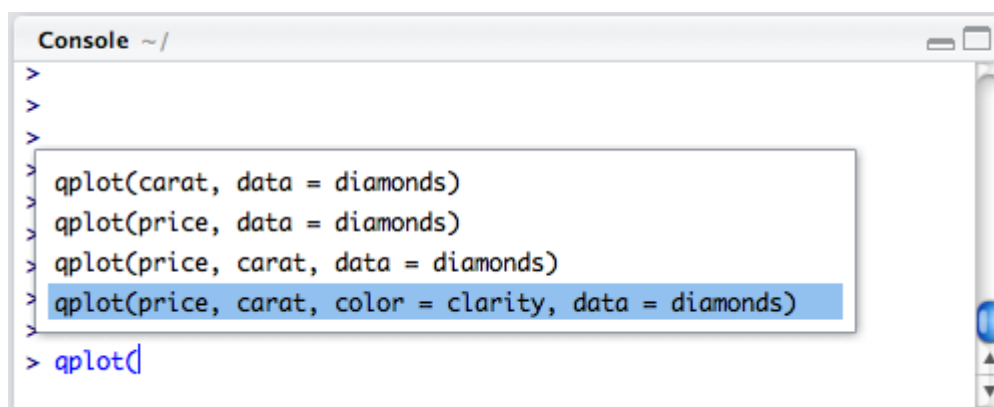
2.3. Возврат к предыдущим командам

При работе с R постоянно возникает необходимость выполнить заново ту или иную команду, которая уже была выполнена ранее. Как и стандартная R-консоль, консоль RStudio поддерживает возможность навигации по ранее выполненным командам, используя клавиши со стрелками.

Для просмотра списка недавно выполненных команд и для выбора определенной команды из этого списка можно использовать также сочетание клавиш `Ctrl+Вверх` (`Command-Вверх` на Mac-компьютерах):



Такое же сочетание клавиш подходит и для быстрого поиска ранее вызванных функций, в имени которых имеется определенный префикс. Например, для поиска функций, в имени которых есть `plot`, следует просто ввести `plot` и нажать `Ctrl+Вверх`:



```
>  
>  
>  
>  
> qplot(carat, data = diamonds)  
> qplot(price, data = diamonds)  
> qplot(price, carat, data = diamonds)  
> qplot(price, carat, color = clarity, data = diamonds)  
> qplot(|
```

2.4. Заголовок окна Консоли

В заголовке окна Консоли имеется несколько дополнительных опций:

- Отображение рабочей директории.
- Возможность прервать текущие вычисления.
- Сворачивание и разворачивание окна Консоли по отношению к панели Редактора кода (Source) (с использованием кнопок в правом верхнем углу или двойного клика по заголовку окна).

2.5. Горячие клавиши

Ctrl+1 - перемещает курсор в окно Источника кода

Ctrl+2 - перемещает курсор в Консоль

Ctrl+L - Очищает окно Консоли от текста

Esc - прерывает вычисления

3. Редактирование и выполнение кода

3.1. Общая информация

Редактор кода RStudio включает ряд опций для продуктивной работы, в частности подсветку кода, автоматическое завершение кода, одновременное редактирование нескольких файлов, поиск и замену определенных частей кода.

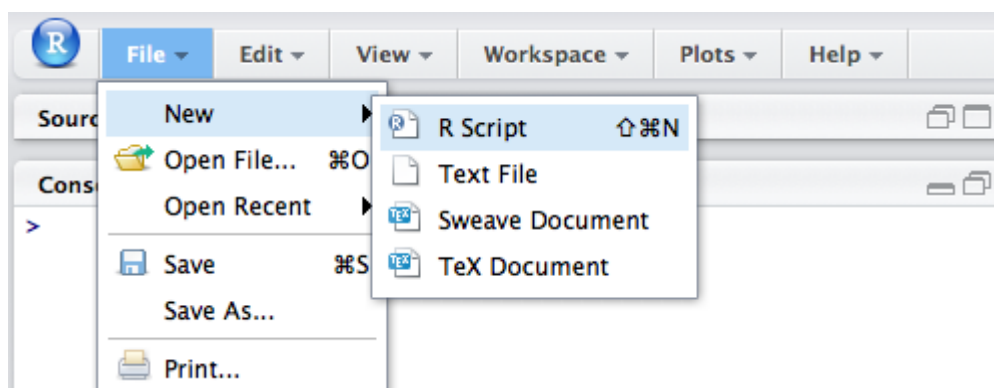
Кроме того, в RStudio имеются гибкие возможности по выполнению кода непосредственно из окна редактора. Для многих пользователей это является предпочтительным способом работы с R. Выполнение команд из окна Редактора кода вместо командной строки Консоли облегчает воспроизведение одних и тех же команд и позволяет "упаковать" такие команды в одну функцию для последующего использования.

3.2. Управление файлами

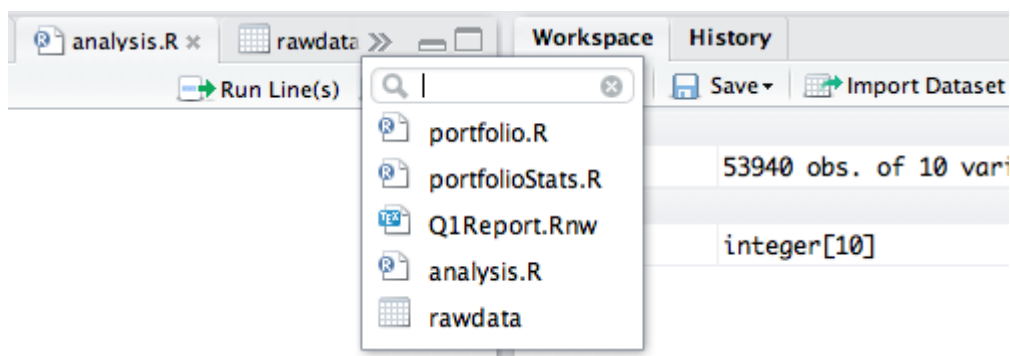
RStudio поддерживает подсветку синтаксиса и другие специализированные опции по работе с кодом следующих типов файлов:

- R-скрипты
- документы Sweave
- документы TeX

Для создания нового файла используйте меню File -> New:



Для открытия существующего файла необходимо воспользоваться меню File -> Open или Open Recent (для открытия файла, с которым работа осуществлялась недавно). Если открыто несколько файлов одновременно, быстрый переход от одного документа к другому выполняется при помощи соответствующих закладок в верхней части окна редактора кода. Для навигация между большим количеством открытых файлов служит иконка >> в правой верхней части окна Редактора; можно также воспользоваться меню View -> Switch to.

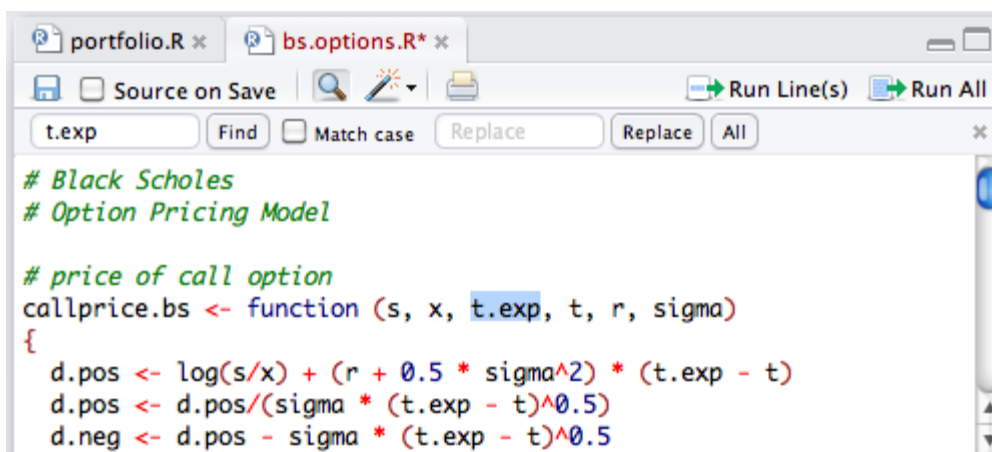


3.3. Завершение кода

Как уже отмечалось выше при описании работы с Консолью, RStudio поддерживает автоматическое завершение кода при помощи клавиши Tab. Например, если в рабочем пространстве имеется объект с именем `pollResults`, то можно набрать на клавиатуре `poll`, нажать Tab, и RStudio автоматически завершит название этого объекта.

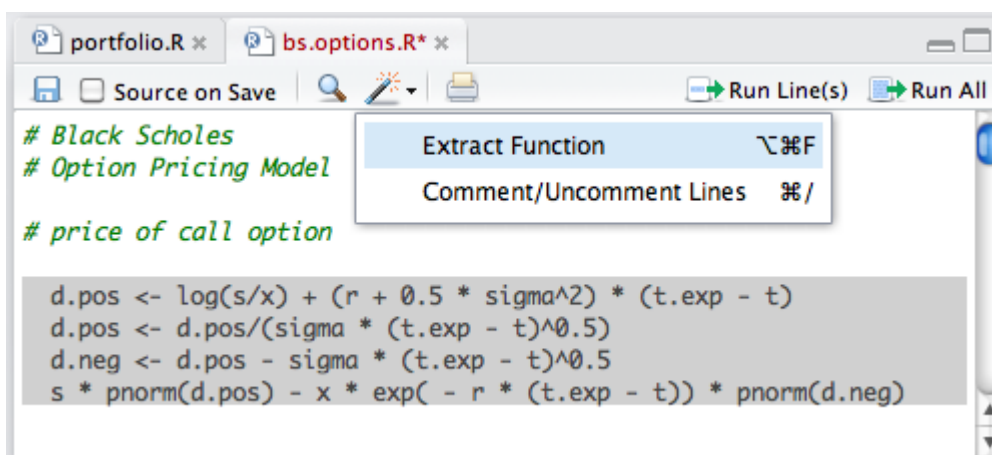
3.4. Поиск и замена частей кода

В Редакторе кода RStudio имеется возможность найти и заменить необходимые части текста. Для выполнения этой операции необходимо воспользоваться сочетанием клавиш Ctrl+F или разделом меню Edit -> Find and Replace. В верхней части окна редактора имеется кнопка в виде увеличительного стекла - ее нажатие также приводит к открытию диалогового окна, в котором задаются параметры поиска.



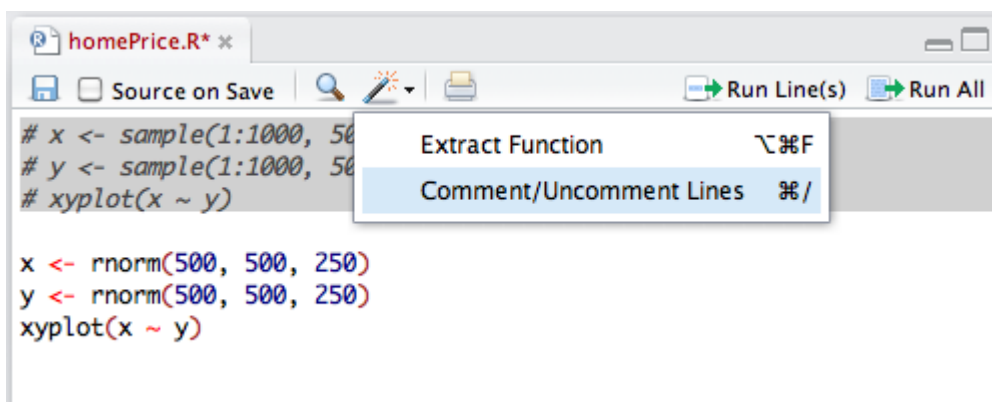
3.5. Автоматическое создание функций

Редактор кода RStudio может проанализировать выделенную часть текста и автоматически конвертировать ее в функцию для последующего многократного использования. Все "свободные" переменные в выделенном коде (т.е. объекты, на которые имеется ссылка, но которые не созданы внутри выделения) будут при этом преобразованы в аргументы функции:



3.6. Комментирование кода

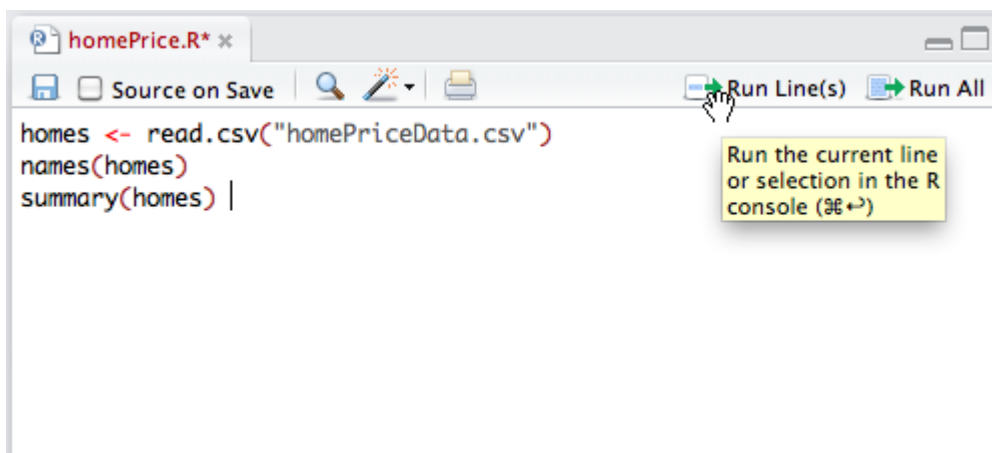
Для комментирования выделенных частей кода (или для снятия комментирования) используйте пункт меню Edit -> Comment/Uncomment Lines (см. также рисунок ниже):



3.7. Выполнение кода

RStudio поддерживает выполнение кода непосредственно из окна Редактора (выполняемые команды посылаются в Консоль, где появляется также результат их выполнения).

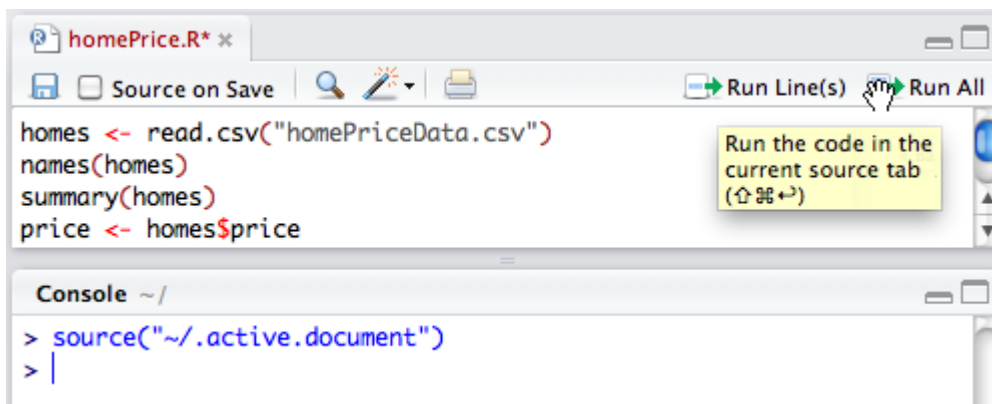
Для выполнения текущей строки кода можно воспользоваться сочетанием клавиш Ctrl+Enter или кнопкой Run Line(s), расположенной в верхней части окна Редактора:



После выполнения кода курсор автоматически перейдет на следующую строку. Такой механизм обеспечивает возможность пошагового выполнения команд.

Имеются две возможности для одновременного выполнения нескольких строк кода:

- Выделить необходимые строки и нажать клавиши Ctrl+Enter (или воспользоваться кнопкой Run Line(s), расположенной в верхней части окна Редактора);
- Нажать Ctrl+Shift+Enter для выполнения всех строк, имеющихя в скрипте (или воспользоваться кнопкой Run All, расположенной в верхней части окна Редактора).



Разница между выполнением кода из выделенной части скрипта и нажатием клавиши Run All состоит в том, что в первом случае строки кода посылаются непосредственно в Консоль, тогда как во втором случае создается временный файл, который затем используется как источник кода для Консоли (так предотвращается "захламление" окна Консоли текстом кода).

3.8. Горячие клавиши

Ctrl+Shift+N - создать новый документ

Ctrl+O - открыть документ

Ctrl+S - сохранить активный документ

Ctrl+1 - переместить курсор в окно Редактора кода

Ctrl+2 - переместить курсор в командную строку Консоли

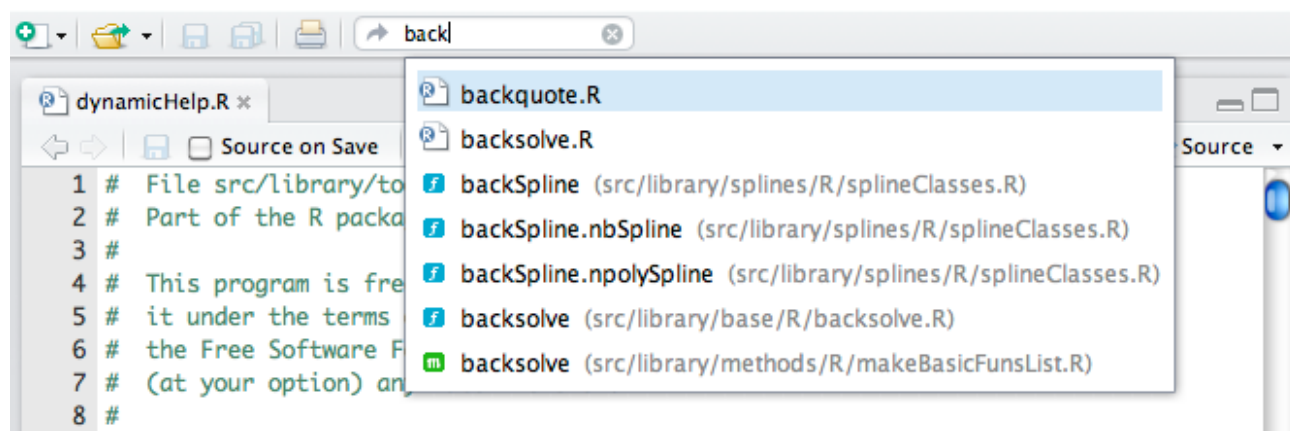
4. Навигация по коду

4.1. Общая информация

RStudio включает ряд опций, обеспечивающих быструю навигацию по R-коду. Изучение этих опций может привести к существенному повышению эффективности работы, а также к лучшему пониманию кода, написанного другими.

4.2. Переход к файлу или функции

Если Вы знаете имя файла со скриптом или название функции, Вы можете быстро найти их, используя поле поиска Go to File/Function (Перейти к Файлу/Функции) в главной панели инструментов RStudio:

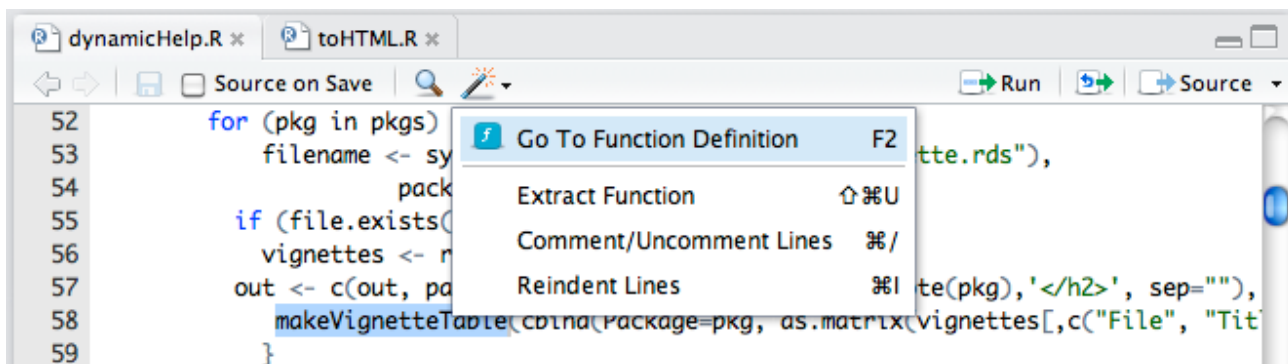


Опция "Go to File/Function" работает благодаря постоянному индексированию создаваемого Вами кода. Конкретные файлы, подлежащие индексированию, определяются следующим образом:

- Если какой-либо проект RStudio активен: индексируются все файлы, находящиеся в папке этого проекта;
- Если проект не активен: индексируются все открытые в данный момент файлы с R-скриптами.

4.3. Переход к определению функции

Благодаря постоянно выполняемому индексированию R-кода (см. выше), RStudio может также помочь Вам найти определение любой функции. Для перехода к коду функции наведите курсор на ее имя (не обязательно выделять его полностью) и выберите команду Go to Function Definition (Перейти к определению функции):



Перейти к определению функции можно также следующими способами:

- Используя горячую клавишу F2.
- Используя Ctrl + щелчок по имени функции.
- Используя Редактор кода или Консоль.

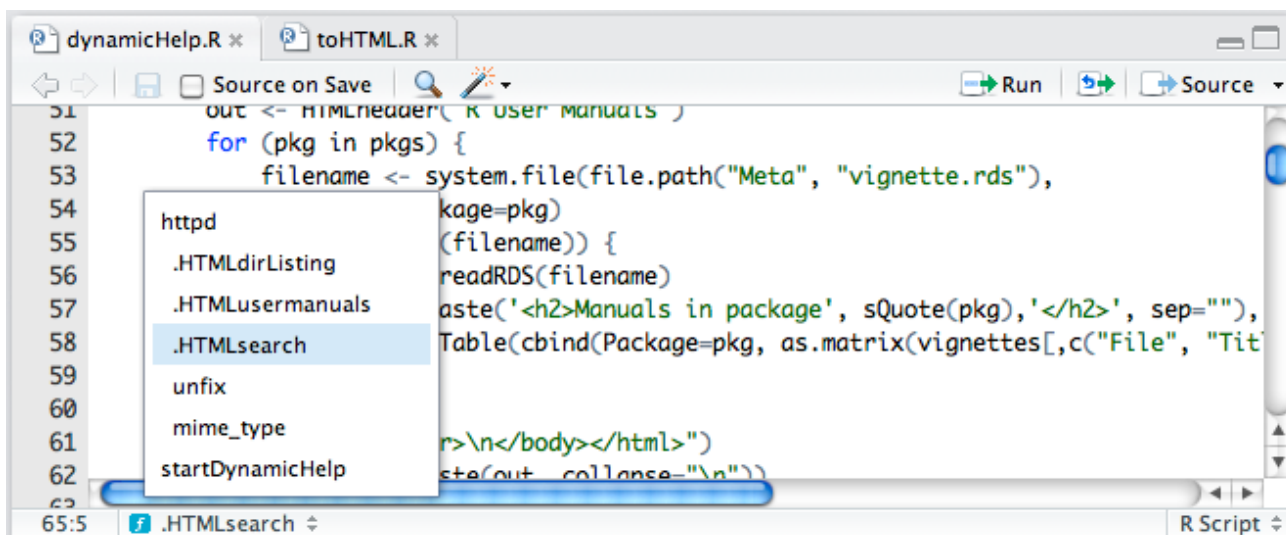
Имеется возможность переходить к определению не только Ваших собственных функций, но также функций из любого пакета R. Код функций из дополнительных пакетов R отображается в специальном Окне просмотра исходного кода (Source Viewer pane) и предназначен только для чтения.

4.4. Навигация по скрипту

Существует несколько способов навигации по коду, хранящемуся в виде скрипта в отдельном файле:

- Рассмотренная выше опция "Go to File/Function" позволяет выполнять поиск в пределах открытого файла со скриптом;
- Команда Jump to Line (Перейти к строке) предназначена для перехода к определенной строке скрипта.

Можно также использовать опции из соответствующего меню (Code Tools):



4.5. Двусторонняя навигация по коду

При навигации по коду (особенно при последовательных переходах от одной функции к другой) часто возникает необходимость быстро вернуться к предыдущему месту, где выполнялось редактирование кода. RStudio хранит список таких редактирований и позволяет переходить от одного к другому в любом направлении, используя команды Back (Назад) и Forward (Вперед) (доступны из меню Edit (Редактирование), а также в верхнем левом углу окна Редактора кода). Команды Back и Forward позволяют повторять следующие действия:

- Открытие документа (или переключение между закладками)
- Переход к определению функции
- Переход к определенной строке кода
- Переход к функции с использованием меню Function

Команды Back и Forward можно вызвать, используя сочетания клавиш Ctrl+F9 / Ctrl+F10 (Cmd+F9 / Cmd+F10 на Mac компьютерах).

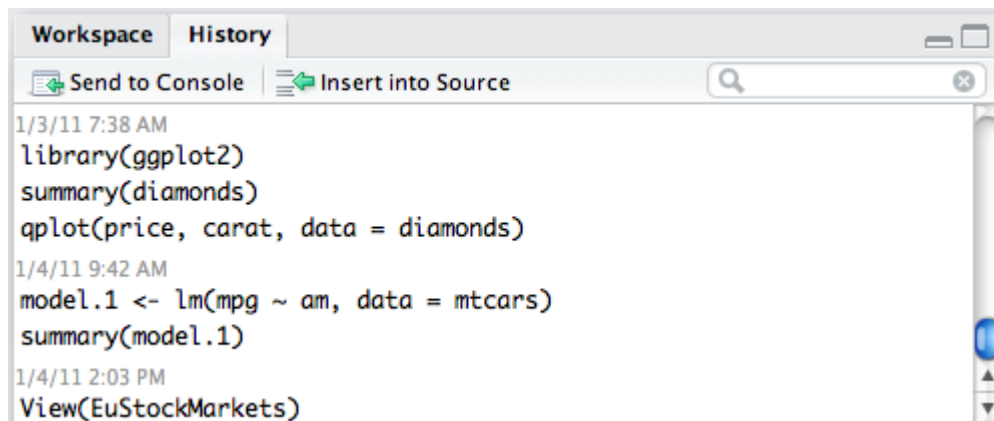
5. Использование истории команд

5.1. Общая информация

Во время работы RStudio создает базу данных всех команд, которые пользователь вводит в Консоль. Имеется возможность просмотра этой базы данных при помощи панели History (История).

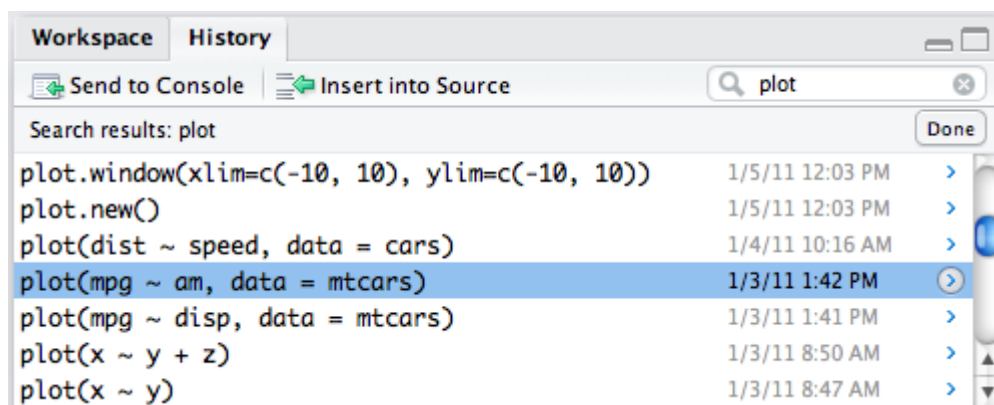
5.2. Обзор истории команд

Закладка History в одной из служебных панелей RStudio служит для просмотра истории выполненных ранее команд. Команды отображаются в порядке их выполнения (недавние команды - внизу списка) и сгруппированы по времени:

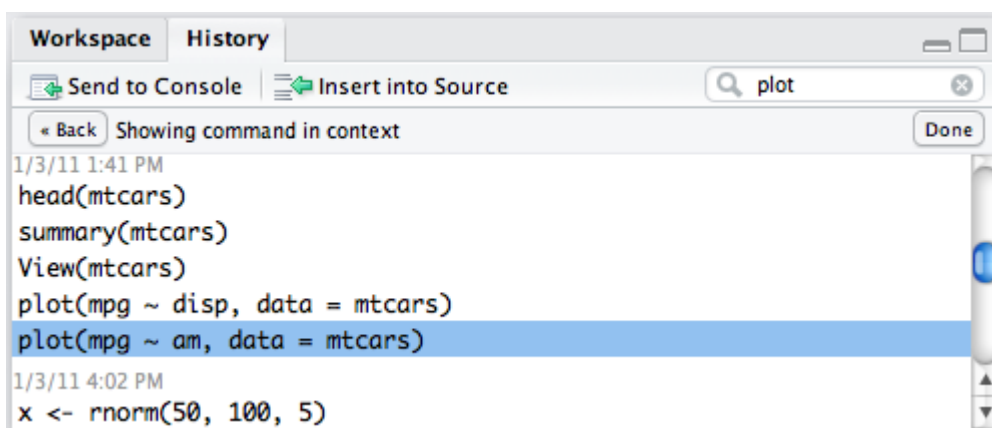


5.3. Поиск в истории команд

В правом верхнем углу списка ранее выполненных команд имеется поле для ввода поисковых запросов (например, `plot`). Условия поиска можно детализировать, введя несколько слов, разделенных пробелами (например, имя какой-либо таблицы данных):



После обнаружения необходимой команды можно также просмотреть и другие команды, которые выполнялись примерно в то же время. Для этого необходимо кликнуть по стрелке, расположенной справа от времени выполнения соответствующей команды:

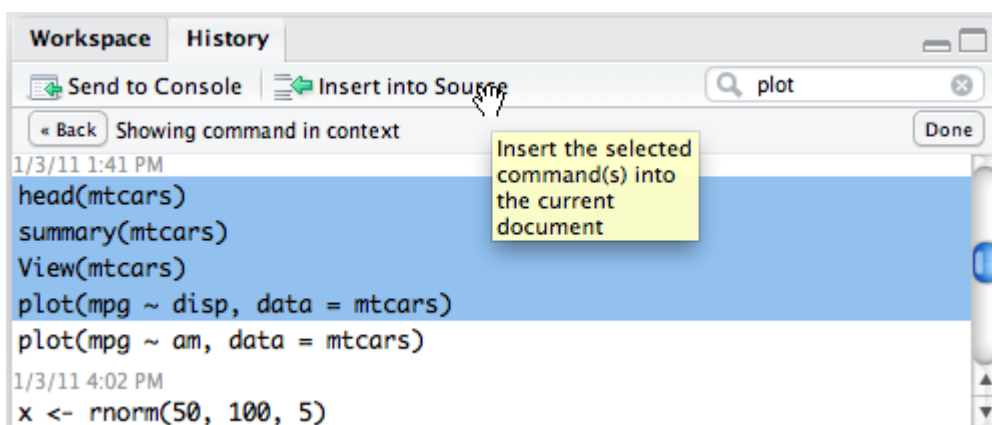


5.4. Выполнение команд

Команды в панели History можно выделить и далее использовать двумя способами (в соответствии с названиями двух кнопок в левой верхней части панели):

- Послать в Консоль (Send to Console) - отправляет выделенные команды в Консоль. Заметьте, однако, что отправленные в Консоль команды не будут выполнены пока не будет нажата клавиша Enter.
- Вставить в Источник (Send to Source) - отправляет выбранные команды в активный документ Редактора кода. Если в данный момент в Редакторе нет активных открытых документов, будет создан новый документ без имени.

В списке ранее выполненных команд имеется возможность выбрать одну или несколько строк:



6. Рабочие директории и рабочие пространства

6.1. Общая информация

Учитывая то, как R по умолчанию обращается с файлами `.RData` и своими рабочими пространствами (workspaces), имеет смысл организовывать работу таким образом, чтобы файлы из разных проектов хранились в разных папках. Представленный ниже материал описывает имеющиеся в RStudio возможности для организации такой модели работы.

6.2. Рабочая директория по умолчанию

Аналогично стандартной версии R, RStudio руководствуется принципом "исходной рабочей директории". Как правило, это домашняя папка пользователя. При запуске RStudio происходит следующее:

- Выполняется файл `.Rprofile` (если таковой имеется) из исходной рабочей папки.
- Загружается файл `.RData` (если таковой имеется) из этой папки в рабочее пространство R.
- Выполняются другие действия, описанные в служебном файле R [Startup](#).

Если в рабочее пространство были внесены изменения, то при завершении работы RStudio появится диалоговая панель с вопросом о необходимости сохранить это рабочее пространство в виде файла `.RData` в текущей рабочей папке. Такое поведение можно изменить в настройках RStudio (меню Tools -> Options -> R General) следующим образом:

- Изменить исходную рабочую директорию (Initial working directory)
- Активировать/деактивировать автоматическую загрузку файла `.RData` из исходной рабочей папки R (Restore `.RData` into workspace at startup)
- Выбрать, нужно ли файл `.RData` сохранять всегда (Always), никогда (Never), или спрашивать о сохранении при завершении работы (Ask).

6.3. Изменение рабочей директории

Текущая рабочая директория отображается в верхней части окна Консоли. Имеется несколько способов изменить рабочую директорию:

- Использовать функцию [setwd\(\)](#)
- Воспользоваться разделом меню Tools -> Change Working Dir...
- На закладке Files выбрать More -> Set As Working Directory

Необходимо помнить о возможных последствиях изменения рабочей папки:

- Имеющиеся в коде относительные ссылки на файлы станут недействительными.
- При завершении работы файл `.RData` будет сохранен в новую папку.

Поскольку эти побочные эффекты могут привести к ошибкам при выполнении кода, рекомендуется начинать работу из папки, соответствующей текущему проекту, и не менять ее до завершения сессии.

Если все файлы, имеющие отношение к определенному проекту, хранятся в одной папке, имеет смысл сделать ее исходной для R. Способы для этого будут варьировать в зависимости от операционной системы. Тем не менее, на всех платформах при установке RStudio происходит ассоциация этой программы с `.RData`, `.R` и другими имеющими отношение к R файлами. При открытии таких файлов RStudio автоматически будет делать рабочей папкой ту, в которой хранится открываемый файл. Однако следует помнить, что если RStudio уже запущена, открытие новых ассоциированных с ней файлов не приведет к изменению рабочей папки.

6.4. Загрузка и сохранение рабочих пространств

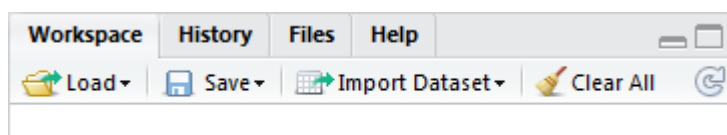
Для сохранения рабочего пространства в рабочую папку R или загрузки пространства из рабочей папки во время активной сессии RStudio можно использовать следующие команды:

```
save.image()  
load(".RData")
```

Следует помнить, что функция `load()` добавит загружаемое рабочее пространство к текущему. Следовательно, если в активном и загружаемом пространствах имеются объекты с одинаковыми именами, то объекты из загружаемого пространства заменят одноименные объекты в активном пространстве. Это может привести к нежелательным последствиям. Поэтому перед загрузкой нового рабочего пространства может иметь смысл очистить активное пространство. Это можно сделать при помощи следующей команды:

```
rm(list=ls())
```

Закладка **Workspace** также имеет опции для выполнения описанных выше операций:



7. Проекты

7.1. Общая информация

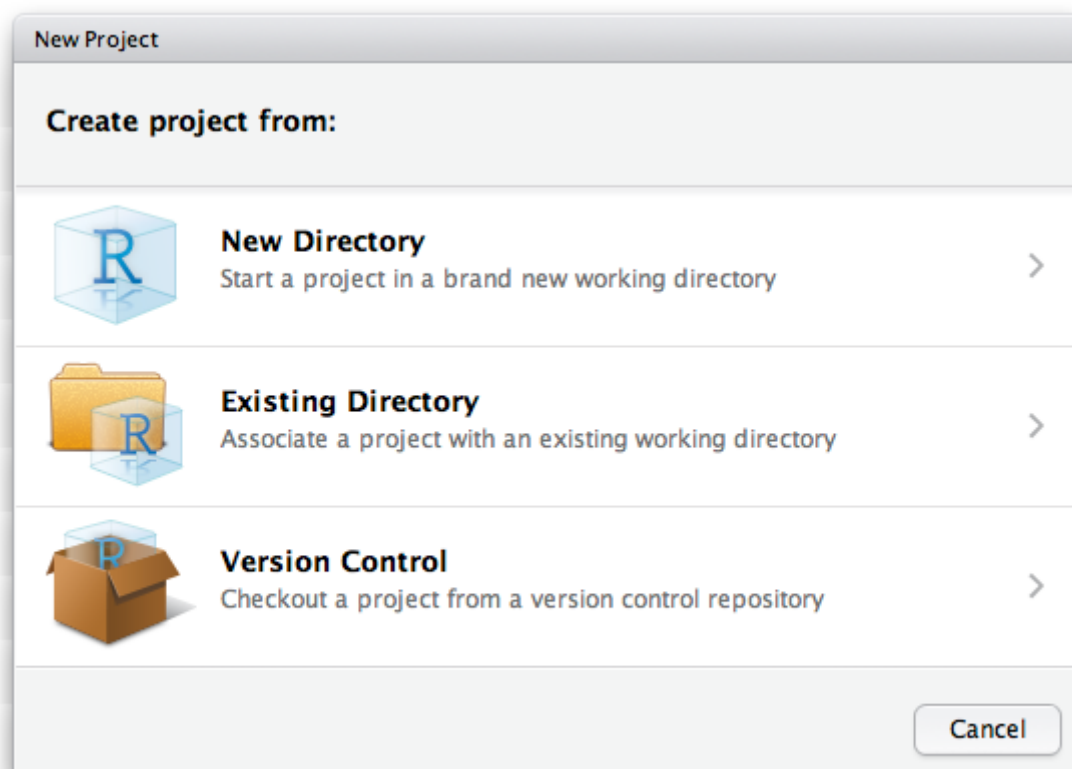
RStudio позволяет легко организовать Вашу работу в соответствующие контексту проекты так, что каждый проект будет иметь свою собственную рабочую директорию, рабочее пространство, историю и скрипты.

7.2. Создание проектов

Проекты RStudio ассоциированы с рабочими директориями R. Следовательно, проект можно создать

- в совершенно новой директории;
- в существующей директории, где уже хранятся скрипты с R-кодом и данные;
- путем копирования файлов, хранящихся в одной из онлайн-систем контроля версий - [Git](#) или [Subversion](#).

Для создания нового проекта служит команда New Project (Новый Проект), доступная из закладки Projects главного меню и из панели инструментов (в дальнем правом углу рабочего окна программы):



При создании нового проекта RStudio выполняет следующие действия:

- Создает файл проекта (с расширением `.Rproj`) в соответствующей папке. Этот файл содержит различные опции по управлению проектом (см. ниже), а также может использоваться как ярлык для запуска проекта.
- Создает невидимую директорию (с именем `.Project.user`) для хранения временных служебных файлов, имеющих отношение к проекту (например, автоматически сохраненные копии скриптов, файлы состояния оконной системы, и т.п.).
- Загружает проект и отображает его название в панели инструментов `Projects` (находится в дальнем правом углу рабочего окна программы).

7.3. Работа с проектами

Открытие проекта

Имеется несколько способов открытия проектов:

- Использование команды `Open Project` (Открыть Проект), доступной из закладки `Projects` главного меню и панели инструментов, для навигации к файлу ранее созданного проекта (например, `MyProject.Rproj`).
- Выбор проекта из списка недавно отредактированных проектов (эта опция также доступна из закладки `Projects` главного меню и из панели инструментов).
- Двойной щелчок по файлу-ярлыку проекта (см. выше).

При открытии проекта в RStudio выполняются следующие действия:

- Запускается новая сессия R;
- Выполняются команды из служебного файла `.Rprofile` (если таковой имеется в основной папке проекта);
- Файл `.Rhistory` загружается из папки проекта в панель `History` (История) и используется для навигации по ранее выполненным командам;
- Директория проекта становится рабочей директорией R;
- В Редакторе кода открываются ранее отредактированные скрипты с R-кодом.
- Происходит возврат других настроек RStudio (например, активные закладки, состояние разделителей окон, и т.п.) к состоянию, имевшему место перед предыдущим закрытием проекта.

Выход из проекта

Когда Вы выбираете команду `Quit R` (Выход) из раздела `File` главного меню, команду `Close Project` (Закрыть Проект) из панели инструментов `Project`, или просто открываете новый проект, происходит следующее:

- В папку проекта записываются файлы `.RData` и/или `.Rhistory` (если текущие настройки программы предполагают это);

- Сохраняется список входящих в проект скриптов для возможности открыть их в следующий раз;
- Сохраняются другие настройки RStudio (см. выше);
- Текущая сессия R завершается.

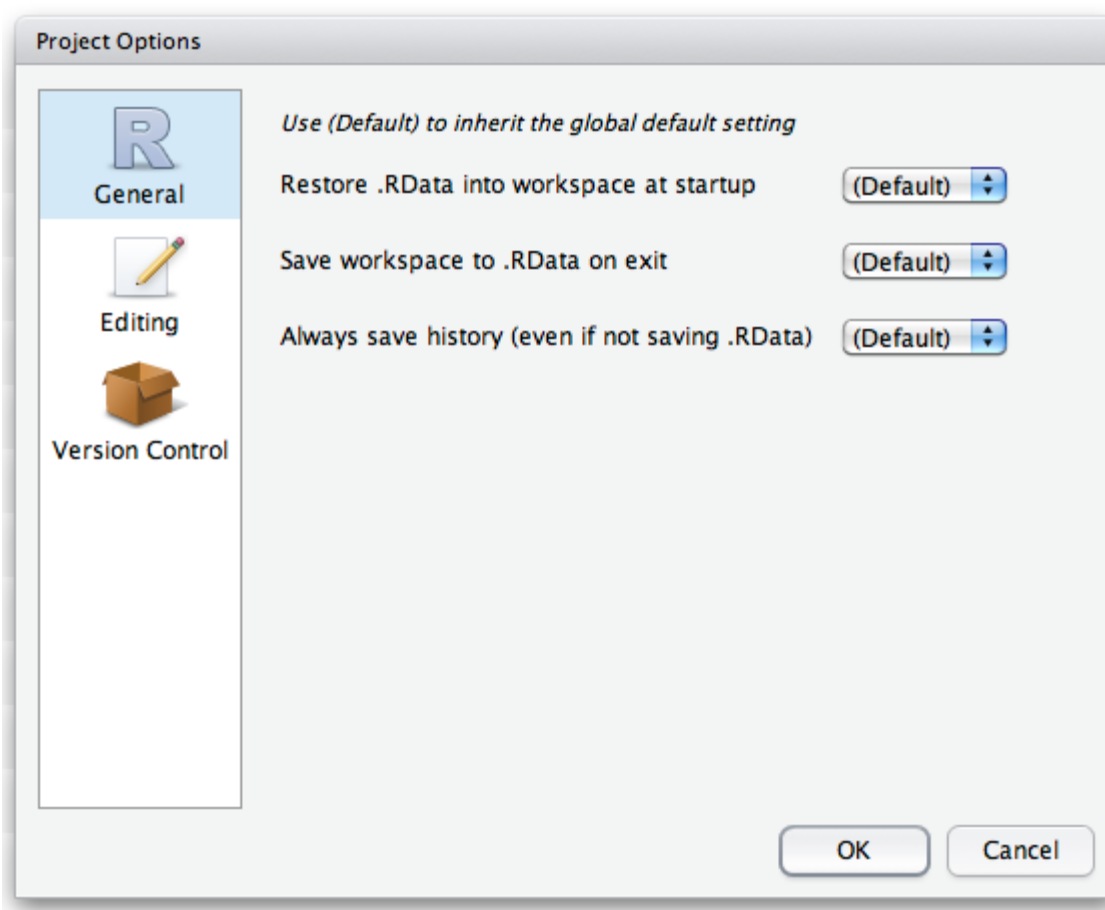
Работа с несколькими проектами одновременно

Вы можете работать с несколькими проектами одновременно, просто запуская соответствующее количество копий RStudio. Это можно сделать двумя способами:

- Использовать команду Open Project in New Window (Открыть проект в новом окне), находящуюся в разделе Project главного меню.
- Открыть несколько проектов из браузера файлов, щелкая по соответствующим ярлыкам.

Настройка проектов

Имеется несколько опций для настройки поведения каждого конкретного проекта в RStudio. Эти опции доступны по команде Project Options из раздела Project главного меню программы:



Закладка General (Общие): При помощи этой закладки можно изменить глобальные настройки поведения проектов. Если изменять их нет необходимости, оставьте значение Default (По умолчанию). Доступны следующие опции:

- Restore .RData into workspace at startup - загрузка файла .RData (если таковой имеется) из исходной рабочей директории R. Если файл .RData имеет большой размер, для данной опции лучше выбрать значение "No" (Нет). Это значительно сократит время запуска программы.
- Save workspace to .RData on exit - сохранение рабочего пространства в виде файла .RData при выходе из программы. Возможные варианты: "Default" (используются глобальные настройки), "Yes" (Да), "No" (Нет) и "Ask" (Спросить). Обратите внимание: если рабочее пространство не было изменено, запрос на сохранение изменений не появится, даже если задано значение "Ask".
- Always save history (even when not saving .RData) - позволяет сделать так, чтобы файл .Rhistory с историей команд всегда сохранялся, даже если Вы выбрали не сохранять файл .RData при выходе из программы.

Закладка Editing (Редактирование):

- Index R source files - позволяет сделать так, чтобы скрипты с R-кодом из конкретного проекта индексировались для возможности навигации по коду (см. выше). Обычно эта опция должна быть активирована, но если в папке проекта находятся "тысячи" файлов, эту опцию можно отключить во избежание замедления работы программы.
- Insert spaces for tab - задает действие клавиши Tab, т.е. сколько пробелов должно появляться при ее нажатии.
- Text encoding - задает кодировку текста в файлах-скриптах. Файлы, кодировка которых не соответствует заданной, все равно могут быть открыты и правильно распознаны при использовании команды Reopen with Encoding из меню File.

Закладка Version Control (Контроль версий):

- Version control system - позволяет указать систему, которая будет использоваться для контроля версий. RStudio автоматически определяет наличие контроля версий путем сканирования папки .git или .svn. Соответственно, эту настройку лучше не изменять. Причинами для внесения изменений могут быть следующие:
 1. В проекте имеется как папка .git, так и папка .svn, и Вы хотите указать однозначно, какая из систем контроля должна использоваться.
 2. Проект не привязан к какой-либо системе контроля версий и Вы хотите добавить логическое хранилище git (эквивалент выполнения команды git init из корневой папки проекта).
- Origin - отображение источника системы контроля версий.