

*IBM SPSS Statistics - Essentials
for R: Инструкции по установке
для Linux*

IBM

Содержание

IBM SPSS Statistics - Essentials for R: Инструкции по установке для Linux . . . 1

IBM SPSS Statistics - Essentials for R: Инструкции по установке для Linux	1
Обзор	1
Установка программы IBM SPSS Statistics	1
Скачивание и установка R 3,2	1
Проверка R.	3
Скачивание и установка IBM SPSS Statistics - Essentials for R	4

Сконфигурируйте среду для IBM SPSS Statistics - Integration Plug-in for R	5
Перед началом использования IBM SPSS Statistics - Integration Plug-in for R	5
Доступ к примерам команд на языке R	5
Деинсталляция компонентов IBM SPSS Statistics - Essentials for R.	8

IBM SPSS Statistics - Essentials for R: Инструкции по установке для Linux

IBM SPSS Statistics - Essentials for R: Инструкции по установке для Linux

Следующие инструкции предназначены для установки IBM® SPSS Statistics - Essentials for R в операционных системах Linux.

Обзор

IBM SPSS Statistics - Essentials for R содержит инструменты, которые потребуются вам, чтобы начать разработку пользовательских программ R для использования с IBM SPSS Statistics. Она содержит:

1. IBM SPSS Statistics - Integration Plug-in for R для IBM SPSS Statistics 24
2. Набор работающих примеров программ R для IBM SPSS Statistics

Установка программы IBM SPSS Statistics

IBM SPSS Statistics - Essentials for R предназначен для следующих программ:

- IBM SPSS Statistics для Linux
- Сервер IBM SPSS Statistics для Linux

Компоненты, установленные с IBM SPSS Statistics - Essentials for R, работают с любой допустимой лицензией IBM SPSS Statistics.

Если вы этого еще не сделали, следуйте инструкциям, прилагаемым к программному обеспечению, чтобы установить одну из программ IBM SPSS Statistics на компьютере, на который будет установлен IBM SPSS Statistics - Essentials for R. Если вы устанавливаете IBM SPSS Statistics - Essentials for R на настольном компьютере, установите на настольном компьютере IBM SPSS Statistics 24. Если вы устанавливаете IBM SPSS Statistics - Essentials for R на компьютере сервера, установите сервер IBM SPSS Statistics 24 на компьютере сервера.

Скачивание и установка R 3,2

Для версии 24 IBM SPSS Statistics - Essentials for R требуется R версии 3,2 (рекомендуется версия 3.2.2). Установите R на компьютере, где вы собираетесь установить IBM SPSS Statistics - Essentials for R.

Необходимые условия работы

На компьютере назначения, где вы собираетесь установить Essentials for R, должна быть система X11. Если на компьютере назначения есть физический дисплей, скорее всего, на нем есть и X11. Далее описаны действия по установке X11, если это потребуется.

1. Установите клиент и сервер X11
 - Для дистрибутивов, которые используют yum, установите программное обеспечение клиента и сервера X11 при помощи команд:

```
yum groupinstall "X Window System" "Desktop" "Fonts" "General Purpose Desktop"
yum update xorg-x11-server-Xorg
yum install xorg-x11-server-Xvfb.x86_64
```
 - Для дистрибутивов, которые используют apt-get, установите программное обеспечение клиента и сервера X11 при помощи команд:

```
apt-get install xorg xterm
apt-get install xserver-xorg xserver-xorg-core xserver-xorg-dev
apt-get install xvfb
```

2. Установите OpenGL

- Для дистрибутивов, которые используют yum, установите OpenGL при помощи команды:
`yum install mesa-libGL-devel mesa-libGLU-devel libpng-devel`
- Для дистрибутивов, которые используют apt-get, установите OpenGL при помощи команды:
`apt-get install libgl1-mesa-glx libgl1-mesa-dev libglu1-mesa libglu1-mesa-dev`

3. Запустите Xvfb. Дополнительную информацию смотрите в разделе <http://www.x.org/archive/X11R7.6/doc/man/man1/Xvfb.1.xhtml>.

4. Задайте переменную среды *DISPLAY*. Общая форма для переменной *DISPLAY* выглядит так:

```
export DISPLAY=<имя_хоста>:<D>.<S>
```

В предыдущем операторе <имя_хоста> - это имя компьютера, где расположен дисплейный сервер X. Чтобы указать localhost, опустите значение <имя_хоста>. <D> - номер дисплея экземпляра Xvfb. <S> - номер экрана, обычно 0.

Примечание: Переменную среды *DISPLAY* надо задать до того, как вы запустите сервер IBM SPSS Statistics.

В дополнение к X11 рекомендуется до установки R установить tcl/tk.

Установка R из менеджера пакетов

Репозиторий вашего дистрибутива может содержать R 3,2. В таком случае можно установить R при помощи стандартного менеджера пакетов распространения (такого как RPM Package Manager или Synaptic Package Manager).

- Для дистрибутивов, которые используют yum, можно установить R при помощи команды `yum install R`.
- Для дистрибутивов, которые используют apt-get, можно установить R при помощи команды:
`apt-get install r-base=<Версия> r-base-core=<Версия> r-base-dev=<Версия>`
где <Версия> - имя версии. Заметим, что вам может понадобиться исправить файл `/etc/apt/source.list`, добавив в него новые источники.

Построение и установка R из исходного кода

Источник R версии 3,2 доступен на <http://www.r-project.org/>. Можно также скачать его непосредственно с <ftp://ftp.stat.math.ethz.ch/Software/CRAN/src/base/R-3/>.

1. Создайте временный каталог, куда вы сможете распаковать и разархивировать исходный код R. Например, введите в командной строке:
`mkdir ~/Rsource`
2. Скачайте исходный код для построения R, например, `R-3.2.2.tar.gz`, и сохраните его во временном каталоге.
3. Перейдите во временный каталог. Например, введите в командной строке:
`cd ~/Rsource`
4. Разархивируйте и распакуйте исходный код R во временный каталог. Например, введите в командной строке:
`tar xzf R-3.2.2.tar.gz`
5. Перейдите в каталог исходного кода. Например, введите в командной строке:
`cd R-3.2.2`

Примечание: Чтобы установить R в каталог по умолчанию, надо выполнить следующий шаг от имени пользователя root, либо войдя в систему как root, либо используя команду `sudo`. Перед тем, как вы перейдете к конфигурированию, построению и установке R, рекомендуем прочитать информацию в файле `doc/html/R-admin.html` (он расположен в подкаталоге каталога, куда вы распаковали исходный код R).

- Введите следующие команды, чтобы задать необходимые параметры компилятора (смотрите специальные параметры для PowerLinux):

```
export CC="gcc -m64"
export CXXFLAGS="-m64 -O2 -g"
export FFLAGS="-m64 -O2 -g"
export FCFLAGS="-m64 -O2 -g"
export LDFLAGS="-L/usr/local/lib64"
export LIBn=lib
```

Параметры PowerLinux:

```
export CC="<XLC_PATH>/bin/xlc_r -q64"
export CFLAGS="-g -O2 -qstrict -qfloat=nomaf:fenv"
export F77="<XLF_PATH>xlf_r -q64"
export FFLAGS="-g -O3 -qstrict -qfloat=nomaf:fenv -qextname"
export CXX="<XLC_PATH>xlc_r -q64"
export CPICFLAGS=-qpvc
export CXXPICFLAGS=-qpvc
export FPICFLAGS=-qpvc
export SHLIB_LDFLAGS=-qmksrshobj
export SHLIB_CXXLDFLAGS=-G
export FC="<XLF_PATH>xlf95_r -q64"
export FCFLAGS="-g -O3 -qstrict -qfloat=nomaf:fenv -qextname"
export FCPICFLAGS=-qpvc
export CXX1XSTD=-qlanglvl=extended0x
```

Здесь `<XLC_PATH>` и `<XLF_PATH>` - положения IBM XL C/C++ for Linux и IBM XL Fortran for Linux соответственно.

- Сконфигурируйте, постройте и установите R. Не забудьте при конфигурировании R указать аргументы `--enable-R-shlib` и `--with-x`. Например, в командной строке введите (смотрите специальные параметры для PowerLinux):

```
./configure --enable-R-shlib --with-x && make && make install
```

Параметры PowerLinux:

```
./configure --enable-R-shlib --with-x --with-readline=no && gmake && gmake install
```

Проверка R

Для проверки R из командной строки вам понадобится добавить библиотеку R к переменной среды `LD_LIBRARY_PATH`. Далее `<R_HOME>` обозначает положение установки R 3,2 - например, `/usr/local/R-3.2.2`.

Например, в командной строке UNIX введите:

```
export LD_LIBRARY_PATH=<R_HOME>/lib/R/lib:.$LD_LIBRARY_PATH
```

Проверьте R, чтобы убедиться в его работоспособности. Например, в командной строке UNIX введите:

```
cd /usr/local/R-3.2.2/bin
```

```
./R
```

Вы должны увидеть сообщение от R и приглашение R (например, `>`). Можно завершить сеанс R, введя `q()`. Теперь можно удалить исходный каталог R, чтобы освободить место на диске.

Скачивание и установка IBM SPSS Statistics - Essentials for R

Убедитесь, что вы используете версию IBM SPSS Statistics - Essentials for R, совместимую с версией IBM SPSS Statistics на вашем компьютере. С основной версией IBM SPSS Statistics, такой как 24, надо использовать версию IBM SPSS Statistics - Essentials for R с той же первой цифрой версии. Нет необходимости устанавливать ранние версии IBM SPSS Statistics - Essentials for R перед установкой более новой версии.

Если пользователи работают в распределенном режиме (с сервером IBM SPSS Statistics Server), надо установить IBM SPSS Statistics - Essentials for R и на клиентском, и на серверном компьютере.

Скачайте версию 24 IBM SPSS Statistics - Essentials for R, доступную в сообществе IBM SPSS Predictive Analytics по адресу <https://developer.ibm.com/predictiveanalytics/predictive-extensions/>.

Важное замечание:

- Если вы устанавливаете Essentials for R на компьютере, у которого нет выхода в интернет, и вы планируете использовать примеры работы с R, входящие в комплект Essentials for R, вам потребуются все пакеты R, требуемые для этих примеров, необходимо будет вручную установить их в R. Чтобы определить, какие пакеты R требуются для конкретного примера R, откройте хаб расширений (Расширения > Хаб расширений), перейдите на вкладку Установлено и нажмите кнопку **Дополнительная информация** для нужного вам примера (каждый пример устанавливается как отдельное расширение). Требуемые пакеты R перечислены в разделе Зависимости диалогового окна Подробности расширения. Пакеты R можно скачать с любого из сайтов-зеркал R CRAN, их адреса приводятся здесь: <http://www.r-project.org/>. Убедитесь, что вы получаете версии пакетов, соответствующие вашей версии R. Пакеты для конкретных версий можно скачать по ссылкам со страницы "Contributed Packages" на сайте-зеркале CRAN.

1. Запустите прикладную программу терминала.
2. Перейдите в каталог, в который вы скачали IBM SPSS Statistics - Essentials for R. В командной строке введите:

```
./<<имя_файла>>
```

где <<имя_файла>> - это имя полученного вами файла.

Примечание: Предыдущую команду надо выполнить как пользователь root, либо войдя в систему как root, либо получив полномочия по команде sudo. Кроме того, если вы не строили R 3,2 из исходного кода, вам может потребоваться установить компиляторы gcc и gfortran до установки IBM SPSS Statistics - Essentials for R.

3. Следуйте инструкциям на экране. Когда у вас запросят положение R, введите положение, где находится каталог *lib* для R - например, */usr/lib/R*.

Установка без вывода сообщений

Кроме варианта установки вручную, описанного выше, можно выполнить установку без вывода сообщений. Это особенно полезно для администраторов сетей, которые выполняют установку для многих конечных пользователей. Чтобы запустить установку без вывода сообщений, выполните следующее:

1. Запустите прикладную программу терминала.
2. Перейдите в каталог, в который вы скачали IBM SPSS Statistics - Essentials for R.
3. При помощи текстового редактора создайте файл ответов под именем *installer.properties*.
4. Добавьте в этот файл ответов следующие свойства и связанные с ними значения:

```
INSTALLER_UI=silent  
USER_SPSS_HOME=<положение IBM SPSS Statistics>  
USER_R_HOME=<Домашний каталог R 3,2>
```

где <положение IBM SPSS Statistics> - это положение установки IBM SPSS Statistics, а <домашний каталог R 3,2> - положение установки R 3,2. Например:

```
USER_SPSS_HOME=/opt/IBM/SPSS/Statistics/24  
USER_R_HOME=/usr/lib/R
```

5. Сохраните *installer.properties* в каталоге, содержащем файл *.bin* для IBM SPSS Statistics - Essentials for R, и перейдите в этот каталог.
6. Запустите программу установки, введя следующую команду:

```
./<имя_программы_установки>
```

где *<имя_программы_установки>* - это имя файла *.bin* для IBM SPSS Statistics - Essentials for R.

Примечание: Предыдущую команду надо выполнить как пользователь *root*, либо войдя в систему как *root*, либо получив полномочия по команде *sudo*.

В ходе установки все пакеты R, требуемые для примеров R, будут автоматически скачаны из Интернета, если возможно. Это может занять несколько минут. Если вы сталкиваетесь с проблемами в связи с необходимыми пакетами R, может понадобиться скачать эти пакеты и установить их вручную.

Примечание: Чтобы использовать другой файл ответов (вместо *installer.properties*), запустите программу установки следующей командой:

```
./<имя_программы_установки> -f <имя_файла_ответов>
```

Сконфигурируйте среду для IBM SPSS Statistics - Integration Plug-in for R

Для IBM SPSS Statistics - Integration Plug-in for R требуется добавить путь к переменной среды *LD_LIBRARY_PATH*. Добавьте эти параметры в файл *statsenv.sh*, находящийся в положении установки IBM SPSS Statistics.

- Добавьте путь к библиотекам *libpng*, *libjpeg* и *zlib* на вашем компьютере к переменной среды *LD_LIBRARY_PATH*.
- Если вы строите R из источника (а не устанавливаете R из менеджера пакетов), надо также изменить переменную среды *LD_LIBRARY_PATH* и для Linux, и для zLinux следующим образом:

```
export LD_LIBRARY_PATH=<R_HOME>/lib/R/lib:<SPSS_HOME>/lib:/lib64:.$LD_LIBRARY_PATH
```

Для PowerLinux используйте:

```
export LD_LIBRARY_PATH=<R_HOME>/lib/R/lib:<SPSS_HOME>/lib:/<COMPILER_HOME>/lib:.$LD_LIBRARY_PATH
```

В предыдущих операторах *<SPSS_HOME>* - это положение, где установлена версия 24 программы IBM SPSS Statistics, а *<R_HOME>* - положение, где установлен R 3,2 - например, */usr/local/R-3.2.2*. Для PowerLinux *<COMPILER_HOME>* - положение, где установлен компилятор - например, */opt/ibm*.

Перед началом использования IBM SPSS Statistics - Integration Plug-in for R

После установки IBM SPSS Statistics - Essentials for R вы сможете начать разработку прикладных программ R с использованием IBM SPSS Statistics - Integration Plug-in for R. Полная документация для этого подключаемого модуля доступна в теме Integration Plug-in for R в системе справки IBM SPSS Statistics. Вы можете также захотеть просмотреть учебные материалы, доступные из темы Работа с R в справочной системе.

Доступ к примерам команд на языке R

В IBM SPSS Statistics - Essentials for R входит набор рабочих примеров команд расширения для IBM SPSS Statistics, реализованных на языке R, что предоставляет дополнительные возможности, сверх доступных встроенным процедурам SPSS Statistics. Все расширения команд R включают в себя настраиваемое диалоговое окно и команду расширения. Команды расширения можно запускать с использованием синтаксиса команд SPSS Statistics, как и любую встроенную команду, такую как *FREQUENCIES*. Синтаксис команды можно сгенерировать для каждой команды расширения в связанном настраиваемом диалоговом окне.

Таблица 1. Список команд расширения R.

Положение меню	Имя команды	Описание
Анализ>Отчеты>Apriori	SPSSINC APRIORI	Найти часто встречающиеся наборы элементов и правила связывания по алгоритму Apriori.
Анализ>Корреляции>Гетерогенные корреляции	SPSSINC HETCOR	Вычисляет корреляции между номинальными, порядковыми и количественными переменными.
Анализ>Описательная статистика>Q-Q диаграмма, для двух переменных или групповая	SPSSINC QQPLOT2	Q-Q диаграмма, для двух переменных или групповая.
Анализ>Регрессия>Регрессия квантилей	SPSSINC QUANTREG	Оценит одну или несколько условных квантилей для линейной модели.
Анализ>Оценка RanFor	SPSSINC RANFOR	Оценка случайного леса.
Анализ>Предсказание RanFor	SPSSINC RANPRED	Вычислить предсказанные значения для новых данных с использованием лесов из SPSSINC RANFOR.
Анализ>Регрессия>Устойчивая регрессия	SPSSINC ROBUST REGR	Оценивает модель линейной регрессии по робастной регрессии, используя M-оценитель.
Анализ>Регрессия>Регрессия Tobit	SPSSINC TOBIT REGR	Оценить модель регрессии, в которой зависимая переменная имеет фиксированную нижнюю границу, верхнюю границу или обе.
Анализ>Выживание>Расширение регрессии Кокса	STATS COXREGR	Регрессия Кокса (пропорциональных рисков).
Анализ>Классификация>Предсказ. с использованием кластера плотности	STATS DBPRED	Предсказание на основе кластеризации по плотности.
Анализ>Классификации>Кластериз. на основе плотности	STATS DBSCAN	Кластеризация на основе плотности.
Анализ>Регрессия>Системы уравнений	STATS EQNSYSTEM	Оценить систему линейных уравнений.
Анализ>Масштаб>Расширенная модель Рэша	STATS EXTRASCH	Вычислить стандартную и расширенную модели Рэша.
Анализ>Регрессия>Логистическая регрессия Фирта	STATS FIRTHLOG	Логистическая регрессия Фирта.
Анализ>Прогнозирование>Модели GARCH	STATS GARCH	Модели GARCH.
Анализ>Обобщенные линейные модели>Обобщенная регрессия с бустингом	STATS GBM	Оценить модели обобщенной регрессии с бустингом (generalized boosted regression, GBM).
Анализ>Обобщенные линейные модели>Предсказания обобщенной регрессии с бустингом	STATS GBMPRED	Вычислить предсказания для моделей обобщенной регрессии с бустингом.
Файл>Получить рабочее пространство R	STATS GET R	Получить информацию о содержимом рабочего пространства R и создать наборы данных SPSS.
Анализ>Масштаб>Ступенчатая модель откликов	STATS GRM	Подобрать ступенчатые модели откликов для порядковых данных.
Анализ>Масштаб>Модель откликов элементов	STATS IRM	Подобрать трехпараметрические модели откликов элементов.

Таблица 1. Список команд расширения R (продолжение).

Положение меню	Имя команды	Описание
Анализ>Логлинейный анализ>Анализ скрытых классов	STATS LATENT CLASS	Анализ скрытых классов.
Анализ>Описательная статистика>Вычисление скорректированных Р-значений	STATS PADJUST	Вычислить скорректированные Р-значения для нескольких проверок.
Анализ>Обобщенные линейные модели>Пропорциональная регрессия	STATS PROPOR REGR	Линейные модели для зависимых переменных (пропорции).
Анализ>Обобщенные линейные модели>Предсказание пропорциональной регрессии	STATS PROPOR REGRPRED	Вычислить предсказанные значения для моделей пропорциональной регрессии.
Анализ>Регрессия>Разрывность регрессии	STATS RDD	Анализ разрывности регрессии.
Анализ>Регрессия>Относительная важность регрессии	STATS RELIMP	Меры относительной важности для регрессии.
Анализ>Выживание>Параметрич. регрессия	STATS SURVREG	Параметрическая регрессия выживания.
Анализ>Классификация>Поддержка векторных компьютеров	STATS SVM	Поддержка векторного компьютера.
Анализ>Обобщенные линейные модели>Модели подсчета с поправкой на избыток нулевых значений	STATS ZEROINFL	Оценить и предсказать модель подсчета с поправкой на избыток нулевых значений.

Важное замечание:

Для расширения неоднородных корреляций требуются и IBM SPSS Statistics - Integration Plug-in for R, и IBM SPSS Statistics - Integration Plug-in for Python. IBM SPSS Statistics - Integration Plug-in for Python включен в IBM SPSS Statistics - Essentials for Python, который устанавливается по умолчанию с вашим продуктом IBM SPSS Statistics.

Примечания

- Справка для каждой из команд расширения R доступна при нажатии кнопки **Справка** в соответствующем диалоговом окне. Однако эта справка не интегрирована со справочной системой SPSS Statistics.
- Получение полной справки по синтаксису каждой из команд расширения возможно при наведении указателя мыши на команду (в окне синтаксиса) и нажатии клавиши F1. Справку можно получить, запустив команду с подкомандой /HELP. Например:

```
SPSSINC HETCOR /HELP.
```

Однако эта справка по синтаксису команд не интегрирована со справочной системой SPSS Statistics и не включена в *Справочник по синтаксису команд*.

Примечание: Механизм F1 для показа справки не поддерживается в распределенном режиме.

- Если положение меню, заданное для команды расширения, отсутствует в вашем продукте IBM SPSS Statistics, посмотрите в меню **Расширения** связанное диалоговое окно.
- Эти диалоговые окна были созданы при помощи Построителя настраиваемых диалоговых окон в IBM SPSS Statistics. Можно просмотреть структуру любых диалоговых окон и настроить ее с помощью Построителя настраиваемых диалоговых окон. Это можно сделать в окне **Расширения>Утилиты>Построитель пользовательских диалогов (Режим совместимости)...** Чтобы посмотреть структуру диалогового окна, выберите в построителе пользовательских диалоговых окон **Файл>Открыть установленный**.

- Код реализации (файл исходного кода R) и файлы спецификации XML для каждой из команд расширения R можно найти в том каталоге на вашем компьютере, где установлены команды расширения. Для просмотра этого каталога запустите команду синтаксиса SHOW EXTRATHS. На выходе этой команды будет список положений с заголовком "Положения для команд расширения". Файлы устанавливаются в первом доступном для записи положении из списка.
- Даже при работе в режиме Unicode может потребоваться задать вашу локаль SPSS Statistics совпадающей с языком вывода SPSS Statistics (OLANG), чтобы правильно выводить символы расширения. Например, если язык вывода японский, может потребоваться задать и японскую локаль для SPSS Statistics, то есть SET LOCALE= 'japanese'.
- Прочие команды расширения, не включенные в IBM SPSS Statistics - Essentials for R, доступны для скачивания с хаба расширений, доступного в меню **Расширения>Хаб расширений**. На хабе расширений можно также посмотреть доступность обновления для команд расширения, включенных в IBM SPSS Statistics - Essentials for R, а также обновлений любых других установленных вами расширений.

Примечание: Расширения всегда устанавливаются или скачиваются на ваш локальный компьютер с хаба расширений. Если вы работаете в режиме распределенного анализа, вам надо отдельно установить расширения на сервер. Информацию смотрите в разделе системы справки **Основная система > Расширения > Установка комплектов локальных расширений**.

- Если вы устанавливаете расширения на сервере SPSS Statistics, можно использовать сценарий, чтобы установить сразу несколько расширений. Информацию смотрите в разделе системы справки **Основная система > Расширения > Установка комплектов локальных расширений > Пакетная установка комплектов расширений**.

Деинсталляция компонентов IBM SPSS Statistics - Essentials for R.

1. Запустите программу терминала.
2. Перейдите в подкаталог `Uninstall_IBM_SPSS_Statistics_Essentials_for_R_24` в каталоге установки IBM SPSS Statistics.
3. В командной строке введите:

```
./Uninstall_IBM_SPSS_Statistics_Essentials_for_R_24
```

Важно: Необходимо располагать правами на удаление каталога установки, иначе процесс деинсталляции завершится неудачно.



Напечатано в Дании