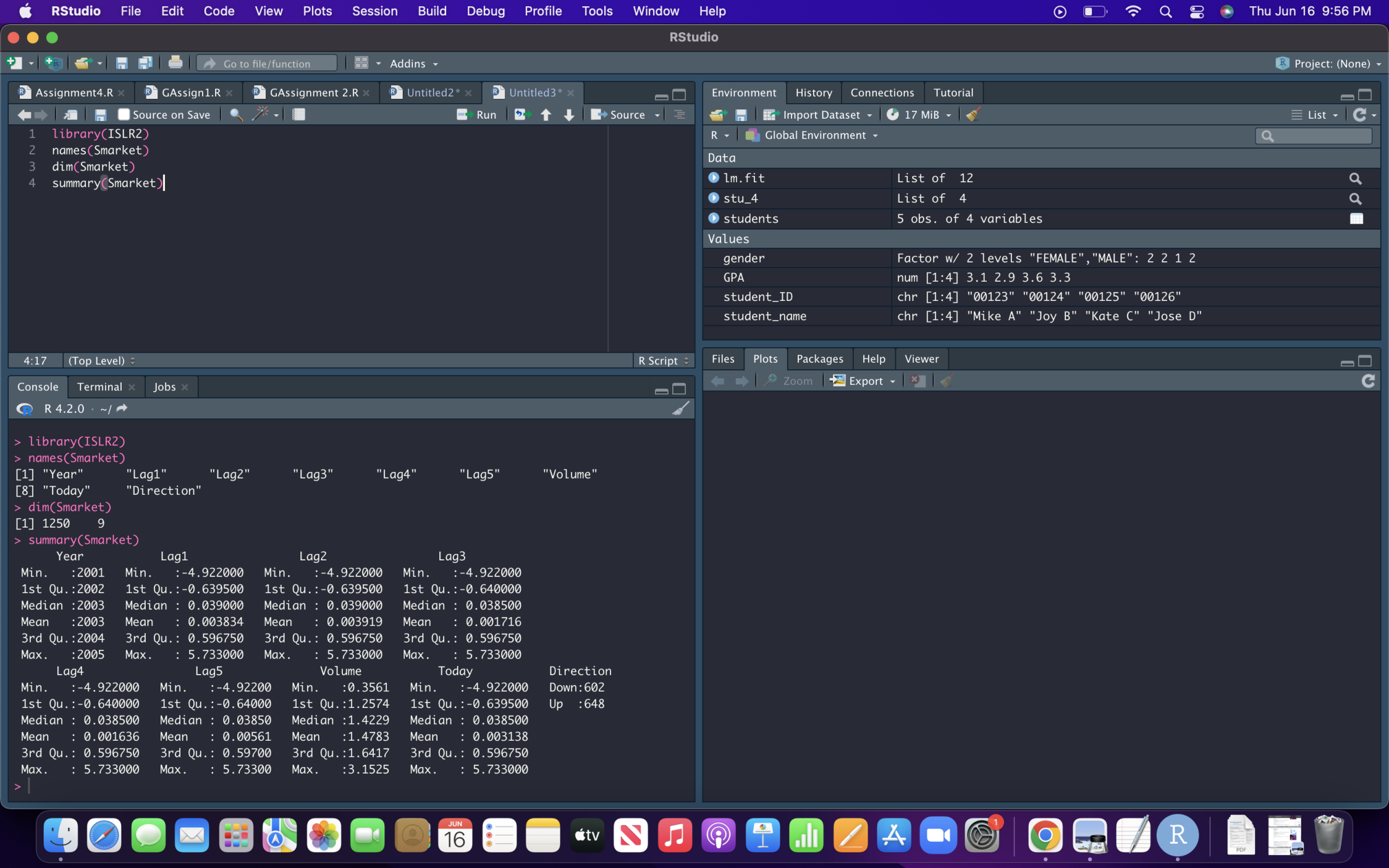
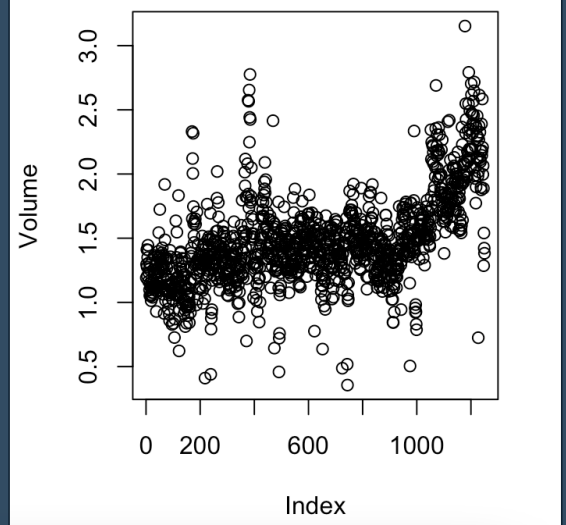
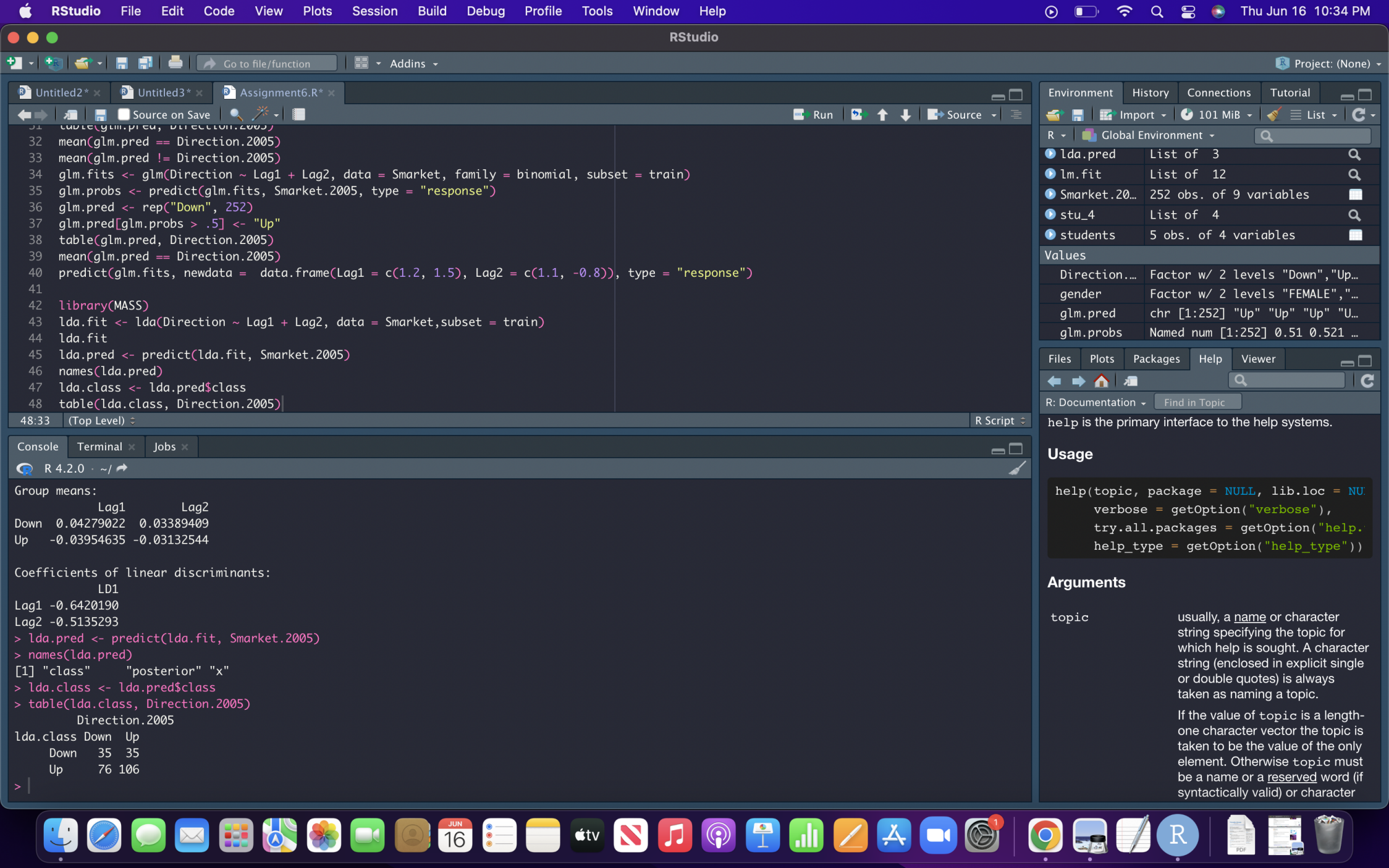
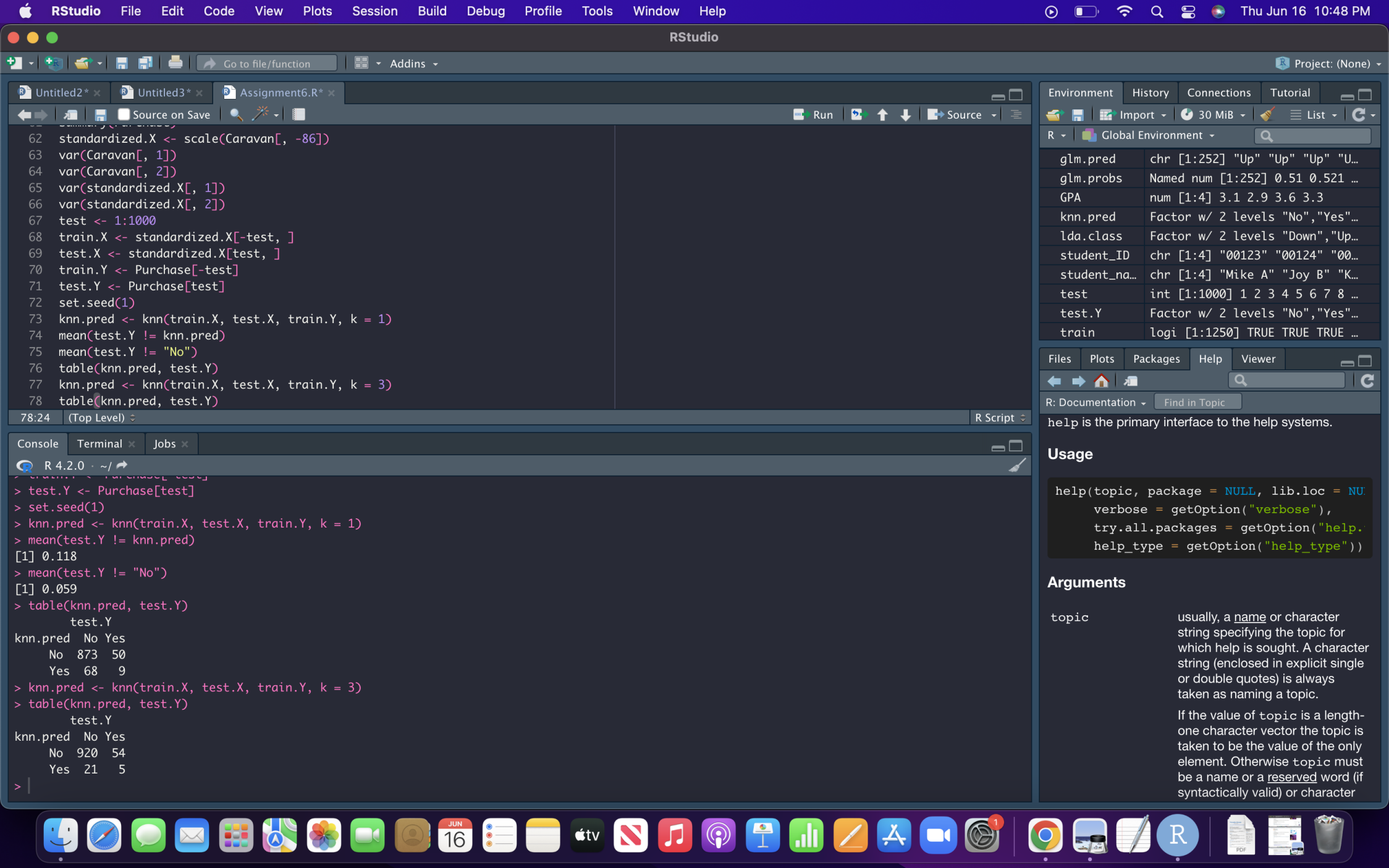

RStudio File Edit Code View Plots Session Build Debug Profile Tools Window Help
Thu Jun 16 9:56 PM
RStudio
Go to file/function
Addins
Project: (None)
Assignment4.R
GAssign1.R
GAssignment 2.R
Untitled2*
Untitled3*
Source on Save
Run
Source
1 library(ISLR2)
2 names(Smarket)
3 dim(Smarket)
4 summary(Smarket)
4:17
(Top Level)
R Script
Console
Terminal
Jobs
R 4.2.0 · ~/
> library(ISLR2)
> names(Smarket)
[1] "Year" "Lag1" "Lag2" "Lag3" "Lag4" "Lag5" "Volume"
[8] "Today" "Direction"
> dim(Smarket)
[1] 1250 9
> summary(Smarket)
Year Lag1 Lag2 Lag3
Min. :2001 Min. :-4.922000 Min. :-4.922000 Min. :-4.922000
1st Qu.:2002 1st Qu.:-0.639500 1st Qu.:-0.639500 1st Qu.:-0.640000
Median :2003 Median : 0.039000 Median : 0.039000 Median : 0.038500
Mean :2003 Mean : 0.003834 Mean : 0.003919 Mean : 0.001716
3rd Qu.:2004 3rd Qu.: 0.596750 3rd Qu.: 0.596750 3rd Qu.: 0.596750
Max. :2005 Max. : 5.733000 Max. : 5.733000 Max. : 5.733000
Lag4 Lag5 Volume Today Direction
Min. :-4.922000 Min. :-4.92200 Min. :0.3561 Min. :-4.922000 Down:602
1st Qu.:-0.640000 1st Qu.:-0.64000 1st Qu.:1.2574 1st Qu.:-0.639500 Up :648
Median : 0.038500 Median : 0.03850 Median :1.4229 Median : 0.038500
Mean : 0.001636 Mean : 0.00561 Mean :1.4783 Mean : 0.003138
3rd Qu.: 0.596750 3rd Qu.: 0.59700 3rd Qu.:1.6417 3rd Qu.: 0.596750
Max. : 5.733000 Max. : 5.73300 Max. :3.1525 Max. : 5.733000
>
Environment
History
Connections
Tutorial
Import Dataset
17 MiB
List
R
Global Environment
Data
lm.fit List of 12
stu_4 List of 4
students 5 obs. of 4 variables
Values
gender Factor w/ 2 levels "FEMALE","MALE": 2 2 1 2
GPA num [1:4] 3.1 2.9 3.6 3.3
student_ID chr [1:4] "00123" "00124" "00125" "00126"
student_name chr [1:4] "Mike A" "Joy B" "Kate C" "Jose D"
Files
Plots
Packages
Help
Viewer
Zoom
Export

```
dim(Smarket.2005)
Direction.2005 <- Direction[!train]
glm.fits <- glm(Direction ~ Lag1 + Lag2 + Lag3 + Lag4 + Lag5 + Volume, data = Smarket, family = binomial, subset = train)
glm.probs <- predict(glm.fits, Smarket.2005,type = "response")
glm.pred <- rep("Down", 252)
glm.pred[glm.probs > .5] <- "Up"
table(glm.pred, Direction.2005)
mean(glm.pred == Direction.2005)
mean(glm.pred != Direction.2005)
glm.fits <- glm(Direction ~ Lag1 + Lag2, data = Smarket, family = binomial, subset = train)
glm.probs <- predict(glm.fits, Smarket.2005, type = "response")
glm.pred <- rep("Down", 252)
glm.pred[glm.probs > .5] <- "Up"
table(glm.pred, Direction.2005)
```

```
> glm.probs <- predict(glm.fits, Smarket.2005,type = "response")
> glm.pred <- rep("Down", 252)
> glm.pred[glm.probs > .5] <- "Up"
> table(glm.pred, Direction.2005)
        Direction.2005
glm.pred Down Up
    Down   77 97
    Up     34 44
> mean(glm.pred == Direction.2005)
[1] 0.4801587
>
> mean(glm.pred != Direction.2005)
[1] 0.5198413
> glm.fits <- glm(Direction ~ Lag1 + Lag2, data = Smarket, family = binomial, subset = train)
> glm.probs <- predict(glm.fits, Smarket.2005, type = "response")
> glm.pred <- rep("Down", 252)
> glm.pred[glm.probs > .5] <- "Up"
> table(glm.pred, Direction.2005)
        Direction.2005
glm.pred Down  Up
    Down   35  35
    Up     76 106
>
```




RStudio File Edit Code View Plots Session Build Debug Profile Tools Window Help
Thu Jun 16 10:34 PM
RStudio
Go to file/function
Addins
Project: (None)
Untitled2*
Untitled3*
Assignment6.R*
Source on Save
Run
Source
mean(glm.pred == Direction.2005)
mean(glm.pred != Direction.2005)
glm.fits <- glm(Direction ~ Lag1 + Lag2, data = Smarket, family = binomial, subset = train)
glm.probs <- predict(glm.fits, Smarket.2005, type = "response")
glm.pred <- rep("Down", 252)
glm.pred[glm.probs > .5] <- "Up"
table(glm.pred, Direction.2005)
mean(glm.pred == Direction.2005)
predict(glm.fits, newdata = data.frame(Lag1 = c(1.2, 1.5), Lag2 = c(1.1, -0.8)), type = "response")
library(MASS)
lda.fit <- lda(Direction ~ Lag1 + Lag2, data = Smarket,subset = train)
lda.fit
lda.pred <- predict(lda.fit, Smarket.2005)
names(lda.pred)
lda.class <- lda.pred$class
table(lda.class, Direction.2005)
48:33
(Top Level)
R Script
Console
Terminal
Jobs
R 4.2.0 · ~/
Group means:
Lag1 Lag2
Down 0.04279022 0.03389409
Up -0.03954635 -0.03132544
Coefficients of linear discriminants:
LD1
Lag1 -0.6420190
Lag2 -0.5135293
> lda.pred <- predict(lda.fit, Smarket.2005)
> names(lda.pred)
[1] "class" "posterior" "x"
> lda.class <- lda.pred$class
> table(lda.class, Direction.2005)
Direction.2005
lda.class Down Up
Down 35 35
Up 76 106
Environment
History
Connections
Tutorial
Import
101 MiB
List
R
Global Environment
lda.pred List of 3
lm.fit List of 12
Smarket.20… 252 obs. of 9 variables
stu_4 List of 4
students 5 obs. of 4 variables
Values
Direction.… Factor w/ 2 levels "Down","Up…
gender Factor w/ 2 levels "FEMALE","…
glm.pred chr [1:252] "Up" "Up" "Up" "U…
glm.probs Named num [1:252] 0.51 0.521 …
Files
Plots
Packages
Help
Viewer
R: Documentation
Find in Topic
help is the primary interface to the help systems.
Usage
help(topic, package = NULL, lib.loc = NU
verbose = getOption("verbose"),
try.all.packages = getOption("help.
help_type = getOption("help_type"))
Arguments
topic
usually, a name or character string specifying the topic for which help is sought. A character string (enclosed in explicit single or double quotes) is always taken as naming a topic.
If the value of topic is a length-one character vector the topic is taken to be the value of the only element. Otherwise topic must be a name or a reserved word (if syntactically valid) or character


RStudio File Edit Code View Plots Session Build Debug Profile Tools Window Help
Thu Jun 16 10:48 PM
RStudio
Go to file/function
Addins
Project: (None)
Untitled2*
Untitled3*
Assignment6.R*
Source on Save
Run
Source
62 standardized.X <- scale(Caravan[, -86])
63 var(Caravan[, 1])
64 var(Caravan[, 2])
65 var(standardized.X[, 1])
66 var(standardized.X[, 2])
67 test <- 1:1000
68 train.X <- standardized.X[-test,]
69 test.X <- standardized.X[test,]
70 train.Y <- Purchase[-test]
71 test.Y <- Purchase[test]
72 set.seed(1)
73 knn.pred <- knn(train.X, test.X, train.Y, k = 1)
74 mean(test.Y != knn.pred)
75 mean(test.Y != "No")
76 table(knn.pred, test.Y)
77 knn.pred <- knn(train.X, test.X, train.Y, k = 3)
78 table(knn.pred, test.Y)
78:24
(Top Level)
R Script
Console
Terminal
Jobs
R 4.2.0 · ~/
> test.Y <- Purchase[test]
> set.seed(1)
> knn.pred <- knn(train.X, test.X, train.Y, k = 1)
> mean(test.Y != knn.pred)
[1] 0.118
> mean(test.Y != "No")
[1] 0.059
> table(knn.pred, test.Y)
test.Y
knn.pred No Yes
No 873 50
Yes 68 9
> knn.pred <- knn(train.X, test.X, train.Y, k = 3)
> table(knn.pred, test.Y)
test.Y
knn.pred No Yes
No 920 54
Yes 21 5
>
Environment
History
Connections
Tutorial
Import
30 MiB
List
R
Global Environment
glm.pred chr [1:252] "Up" "Up" "Up" "U…
glm.probs Named num [1:252] 0.51 0.521 …
GPA num [1:4] 3.1 2.9 3.6 3.3
knn.pred Factor w/ 2 levels "No","Yes"…
lda.class Factor w/ 2 levels "Down","Up…
student_ID chr [1:4] "00123" "00124" "00…
student_na… chr [1:4] "Mike A" "Joy B" "K…
test int [1:1000] 1 2 3 4 5 6 7 8 …
test.Y Factor w/ 2 levels "No","Yes"…
train logi [1:1250] TRUE TRUE TRUE …
Files
Plots
Packages
Help
Viewer
R: Documentation
Find in Topic
help is the primary interface to the help systems.
Usage
help(topic, package = NULL, lib.loc = NU
verbose = getOption("verbose"),
try.all.packages = getOption("help.
help_type = getOption("help_type"))
Arguments
topic
usually, a name or character string specifying the topic for which help is sought. A character string (enclosed in explicit single or double quotes) is always taken as naming a topic.
If the value of topic is a length-one character vector the topic is taken to be the value of the only element. Otherwise topic must be a name or a reserved word (if syntactically valid) or character