

RStudio

File

Edit

Code

View

Plots

Session

Build

Debug

Profile

Tools

Help

Go to file/function

Addins

Project: (None)

EnvironmentHistoryConnectionsTutorial

Global Environment

cars

50 obs. of 2 variables

cars_qm

List of 12

dframe

6 obs. of 4 variables

foo

List of 3

gg

Large gg (9 elements, 3.5 MB)

mydata

5 obs. of 3 variables

mydata2

5 obs. of 3 variables

mymat

num [1:4, 1:3] 1 5 100 11 3 4 20 12 4...

Values

dist

num [1:50] 2 10 4 22 16 10 18 26 34 17 ...

predictedcoun

Named num [1:251] 2.47 2.56 2.66 2.75 2

FilesPlotsPackagesHelpViewer

ZoomExport

Publish

Untitled1* xUntitled2* x

Source on Save

Run

Source

```
1 # Q.1
2
3 data("cars")
4 head(cars)
5 |
```

5:1 (Top Level)

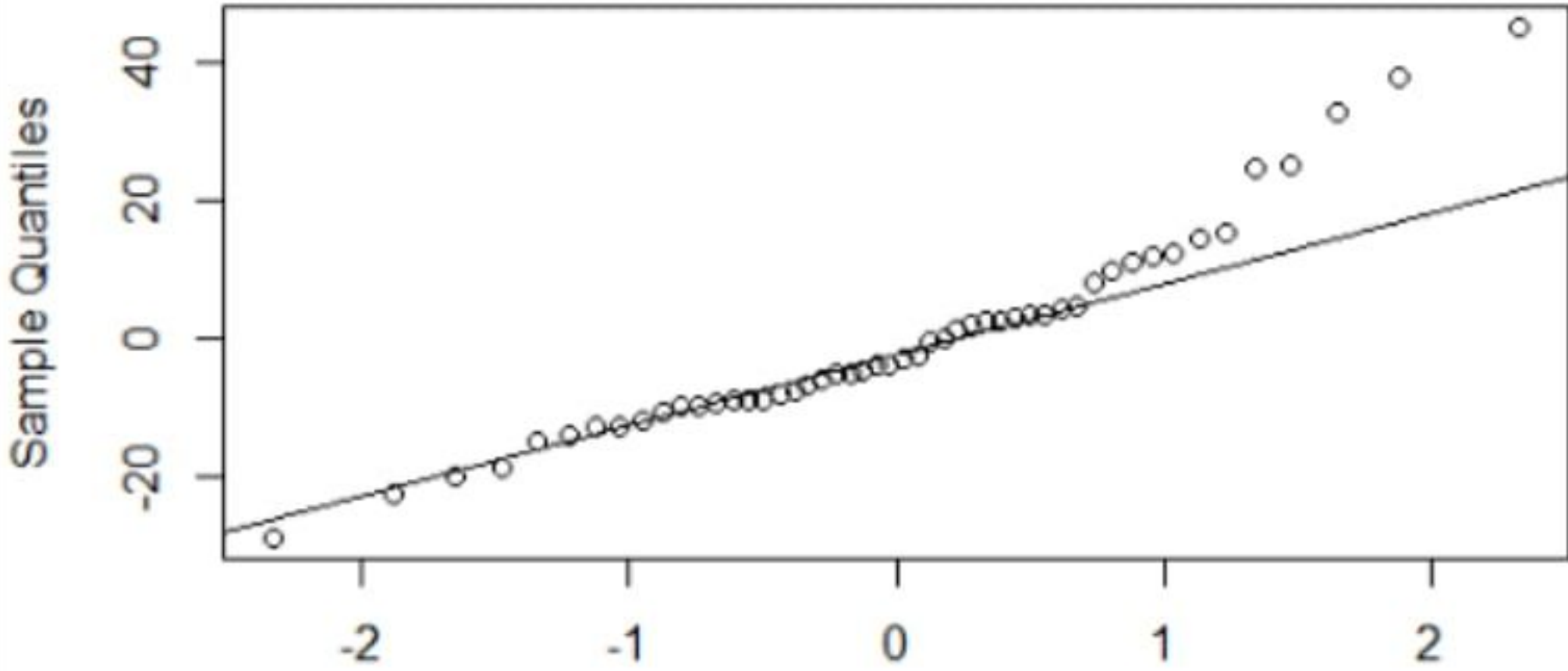
R Script

ConsoleTerminalJobs

~/

```
> # Q.1
>
> data("cars")
> head(cars)
  speed dist
1     4    2
2     4   10
3     7    4
4     7   22
5     8   16
6     9   10
> |
```

Normal Q-Q Plot



The plot displays Sample Quantiles on the y-axis (ranging from -20 to 40) against Theoretical Quantiles on the x-axis (ranging from -2 to 2). A solid diagonal line represents the expected normal distribution. Data points are plotted as open circles. Most points follow the line closely, but there is a noticeable upward curve at the higher end of the theoretical quantiles, suggesting the data is right-skewed.

Search for anything

7

16:02

29-11-2020

ENG

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function

Addins

Project: (None)

Untitled1* x

Untitled2* x

Source on Save

Run

Source

```
1 # Q. 2
2
3 attach(cars)
4 plot(speed, dist, pch=16, las=1, cex=1.2, cex.lab=1.2, xlab='speed (mph)',
5       ylab='stopping distance (ft)')
6
7 #There is a co relation between stopping distance and speed.
8 #Stopping distances increases as speed increases.
```

8:50 (Top Level)

R Script

Console

Terminal x

Jobs x

```
~/
> attach(cars)
The following objects are masked _by_ .GlobalEnv:
    dist, speed

The following objects are masked from cars (pos = 3):
    dist, speed

The following objects are masked from cars (pos = 4):
    dist, speed

> plot(speed, dist, pch=16, las=1, cex=1.2, cex.lab=1.2, xlab='speed (mph)',
+       ylab='stopping distance (ft)')
>
> #There is a co relation between stopping distance and speed.
> #Stopping distances increases as speed increases.
>
```

Environment

History

Connections

Tutorial

Import Dataset

Global Environment

cars	50 obs. of 2 variables
cars_qm	List of 12
dframe	6 obs. of 4 variables
foo	List of 3
gg	Large gg (9 elements, 3.5 MB)
mydata	5 obs. of 3 variables
mydata2	5 obs. of 3 variables
mymat	num [1:4, 1:3] 1 5 100 11 3 4 20 12 4 3 ...

Values

dist num [1:50] 2 10 4 22 16 10 18 26 34 17 ...

predictedcounts Named num [1:251] 2.47 2.56 2.66 2.75 2.85 ...

Files

Plots

Packages

Help

Viewer

Zoom

Export

Publish

Search for anything

7

ENG

16:07

29-11-2020

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function Addins Project: (None)

```
1 # Q. 3
2
3 lm_model <- lm(dist~speed)
4 plot(speed,dist,pch=16,las=1,cex=1.2,cex.lab=1.2,
5       xlab='speed (mph)',ylab='stopping distance (ft)')
6 abline(lm_model,lwd=3,col='red')
```

6:33 (Top Level) R Script

Console Terminal Jobs

```
> # Q. 3
>
> lm_model <- lm(dist~speed)
> plot(speed,dist,pch=16,las=1,cex=1.2,cex.lab=1.2,
+       xlab='speed (mph)',ylab='stopping distance (ft)')
> abline(lm_model,lwd=3,col='red')
>
```

Environment History Connections Tutorial

Global Environment

cars_qm	List of 12
dframe	6 obs. of 4 variables
foo	List of 3
gg	Large gg (9 elements, 3.5 MB)
lm_model	List of 12
mydata	5 obs. of 3 variables
mydata2	5 obs. of 3 variables
mymat	num [1:4, 1:3] 1 5 100 11 3 4 20 12 4 3 ...

Values

dist	num [1:50] 2 10 4 22 16 10 18 26 34 17 ...
predictedcounts	Named num [1:251] 2.47 2.56 2.66 2.75 2.85 ...

Files Plots Packages Help Viewer

Zoom Export Publish

The plot shows a positive linear relationship between speed (mph) on the x-axis and stopping distance (ft) on the y-axis. The x-axis ranges from 5 to 25, and the y-axis ranges from 0 to 120. Black dots represent individual data points, and a thick red line represents the linear regression model. The data points are scattered around the line, showing some variability.

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function Addins Project: (None)

```
1 # Q. 4
2
3 summary(lm_model)
4
5 #There estimated slope coefficient is significantly different than zero.
6 #The R-squared values indicate that about 65% of the variance
7 #   in stopping distance can be explained by speed.
```

7:55 (Top Level) R Script

Console Terminal Jobs

~/

Call:
lm(formula = dist ~ speed)

Residuals:

	Min	1Q	Median	3Q	Max
	-29.069	-9.525	-2.272	9.215	43.201

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-17.5791	6.7584	-2.601	0.0123 *
speed	3.9324	0.4155	9.464	1.49e-12 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 15.38 on 48 degrees of freedom
Multiple R-squared: 0.6511, Adjusted R-squared: 0.6438
F-statistic: 89.57 on 1 and 48 DF, p-value: 1.49e-12

> |

Environment History Connections Tutorial

Global Environment

cars_qm	List of 12
dframe	6 obs. of 4 variables
foo	List of 3
gg	Large gg (9 elements, 3.5 MB)
lm_model	List of 12
mydata	5 obs. of 3 variables
mydata2	5 obs. of 3 variables
mymat	num [1:4, 1:3] 1 5 100 11 3 4 20 12 4 3 ...

Values

dist	num [1:50] 2 10 4 22 16 10 18 26 34 17 ...
predictedcounts	Named num [1:251] 2.47 2.56 2.66 2.75 2.85 ...

Files Plots Packages Help Viewer

Zoom Export Publish

The plot shows a positive linear relationship between speed (mph) on the x-axis and stopping distance (ft) on the y-axis. The x-axis ranges from 5 to 25, and the y-axis ranges from 0 to 120. Black dots represent individual data points, and a solid red line represents the linear regression model. The data points are scattered around the line, showing some variance.

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function

Addins

Project: (None)

Untitled1* x Untitled2* x

Source on Save

Run

Source

```
1 # Q. 5
2
3 par(mfrow = c(2,2))
4 plot(lm_model)
```

4:15 (Top Level) R Script

Console Terminal Jobs

```
> # Q. 5
>
> par(mfrow = c(2,2))
> plot(lm_model)
>
```

Environment History Connections Tutorial

Global Environment

cars_qm

List of 12

dframe

6 obs. of 4 variables

foo

List of 3

gg

Large gg (9 elements, 3.5 MB)

lm_model

List of 12

mydata

5 obs. of 3 variables

mydata2

5 obs. of 3 variables

mymat

num [1:4, 1:3] 1 5 100 11 3 4 20 12 4 3 ...

Values

dist

num [1:50] 2 10 4 22 16 10 18 26 34 17 ...

predictedcounts

Named num [1:251] 2.47 2.56 2.66 2.75 2.85 ...

Files Plots Packages Help Viewer

Zoom Export

Publish

Residuals vs Fitted

Normal Q-Q

Scale-Location

Residuals vs Leverage

Windows Search for anything

7

R

16:13

29-11-2020

ENG

RStudio

FileEditCodeViewPlotsSessionBuildDebugProfileToolsHelp

Go to file/function

Addins

Project: (None)

EnvironmentHistoryConnectionsTutorial

Import Dataset

List

Global Environment

cars_qm

List of 12

dframe

6 obs. of 4 variables

foo

List of 3

gg

Large gg (9 elements, 3.5 MB)

lm_model

List of 12

mydata

5 obs. of 3 variables

mydata2

5 obs. of 3 variables

mymat

num [1:4, 1:3] 1 5 100 11 3 4 20 12 4 3 ...

Values

FilesPlotsPackagesHelpViewer

ZoomExport

Publish

Residuals vs Fitted

Normal Q-Q

Scale-Location

Residuals vs Leverage

Untitled1* xUntitled2* x

Source on Save

RunSource

1# Q. 6

2

3# Yes , the data does appear to be homoscedastic.

4#The mean for the errors is the same for all values of the explanatory variable.

5#This can be seen in the top-left or bottom-left plots of the 4-panel fig in Q.5

6#A perfectly horizontal line at 0 would indicate the mean of the residuals in 0

7# and the same for all values of the explanatory variable.

8#We see there is some deviance from a horizontal line at the trails of the

9# explanatory variable.

9:26 (Top Level)

R Script

ConsoleTerminalJobs

~/

> |

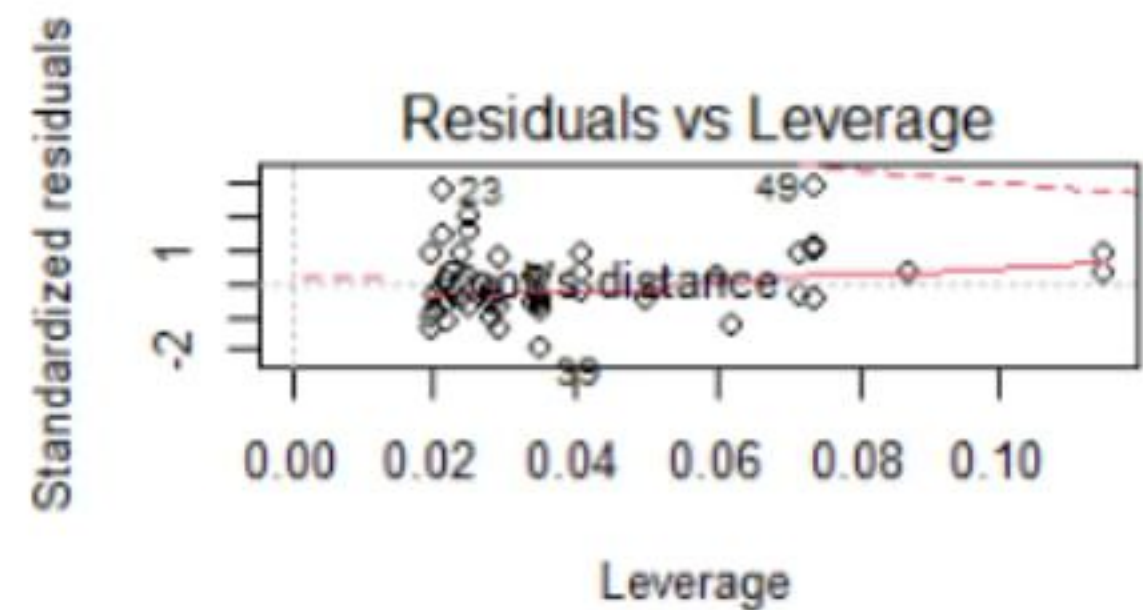
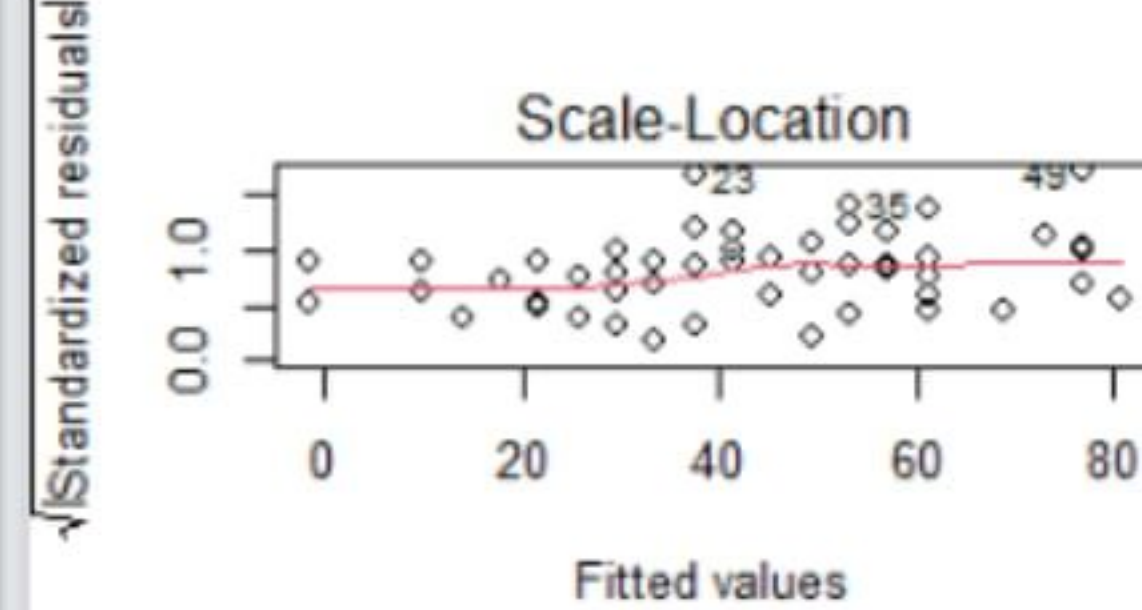
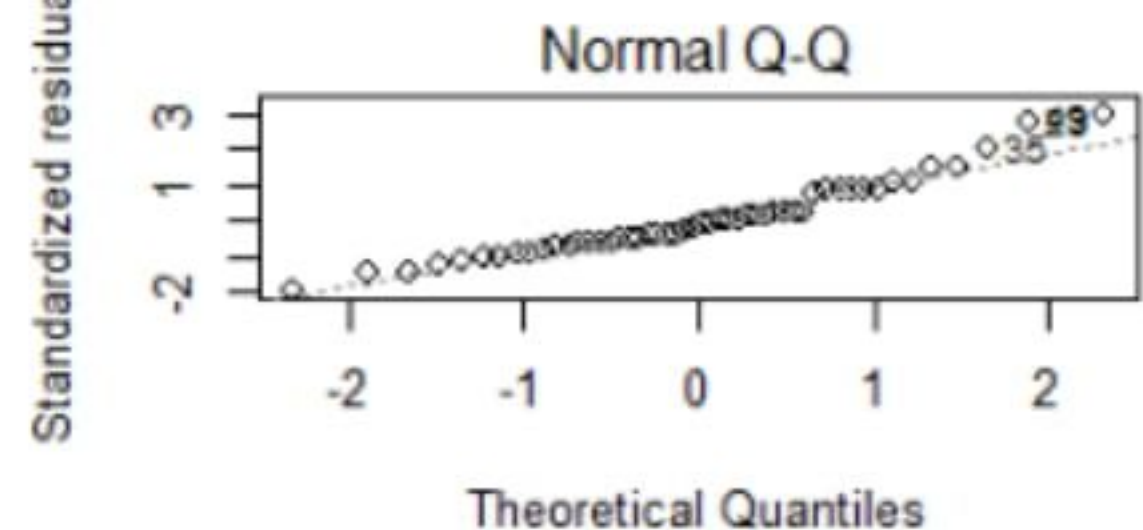
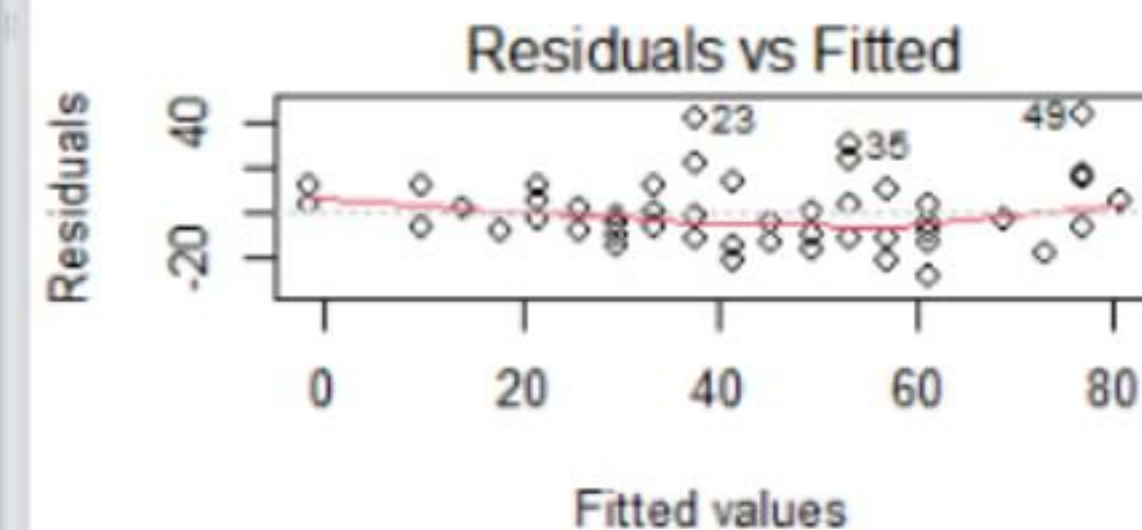
Windows Taskbar

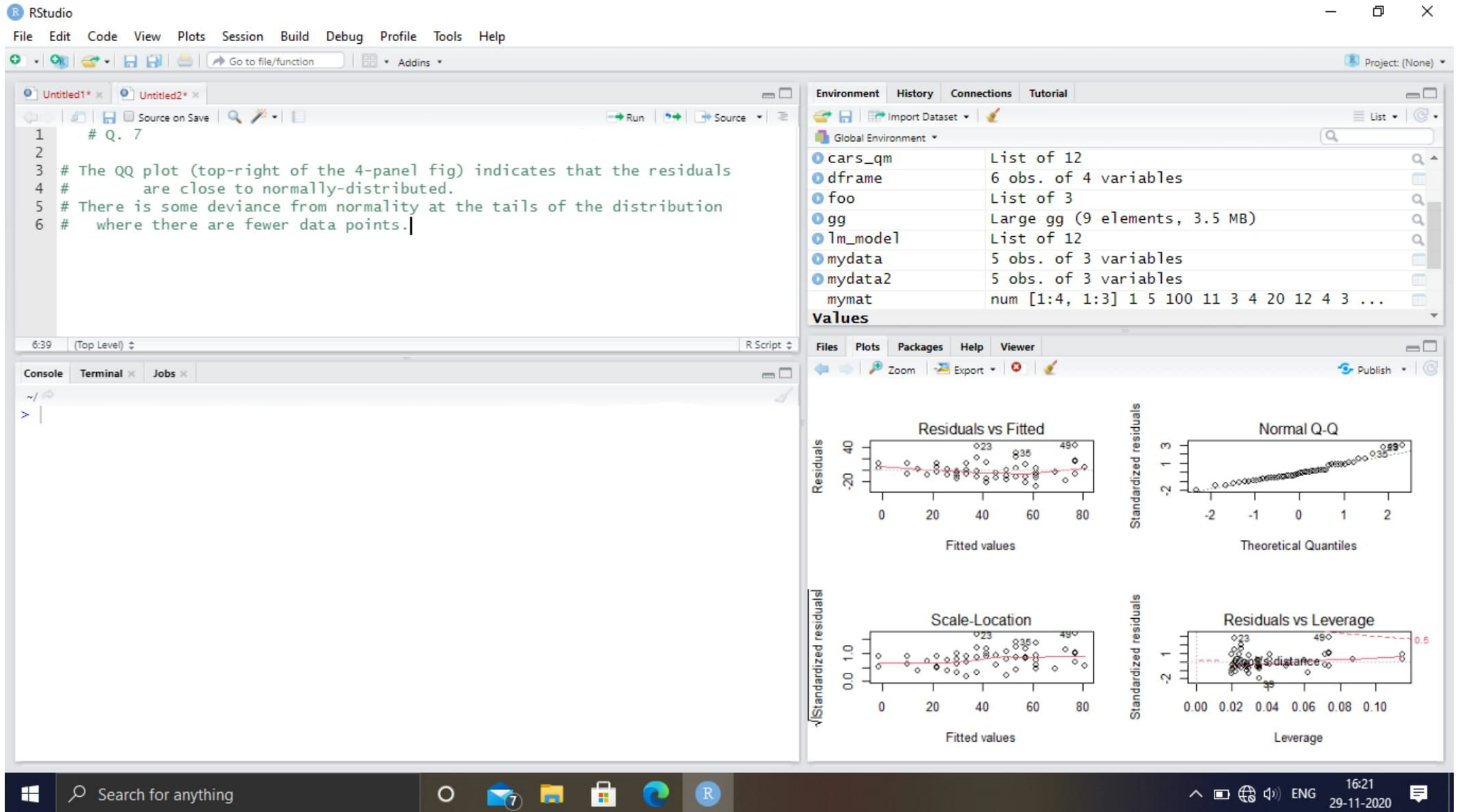
Search for anything

7

ENG

16:19 29-11-2020





RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function Addins Project: (None)

```
1 # Q. 8
2
3 par(mfrow=c(1,1))
4 plot(speed,lm_model$residuals,pch=16,ylab='residuals')
5
6
7 # The residuals are not correlated with the explanatory variable.
```

7:66 (Top Level) R Script

Console Terminal Jobs

```
> # Q. 8
>
> par(mfrow=c(1,1))
> plot(speed,lm_model$residuals,pch=16,ylab='residuals')
>
```

Environment History Connections Tutorial

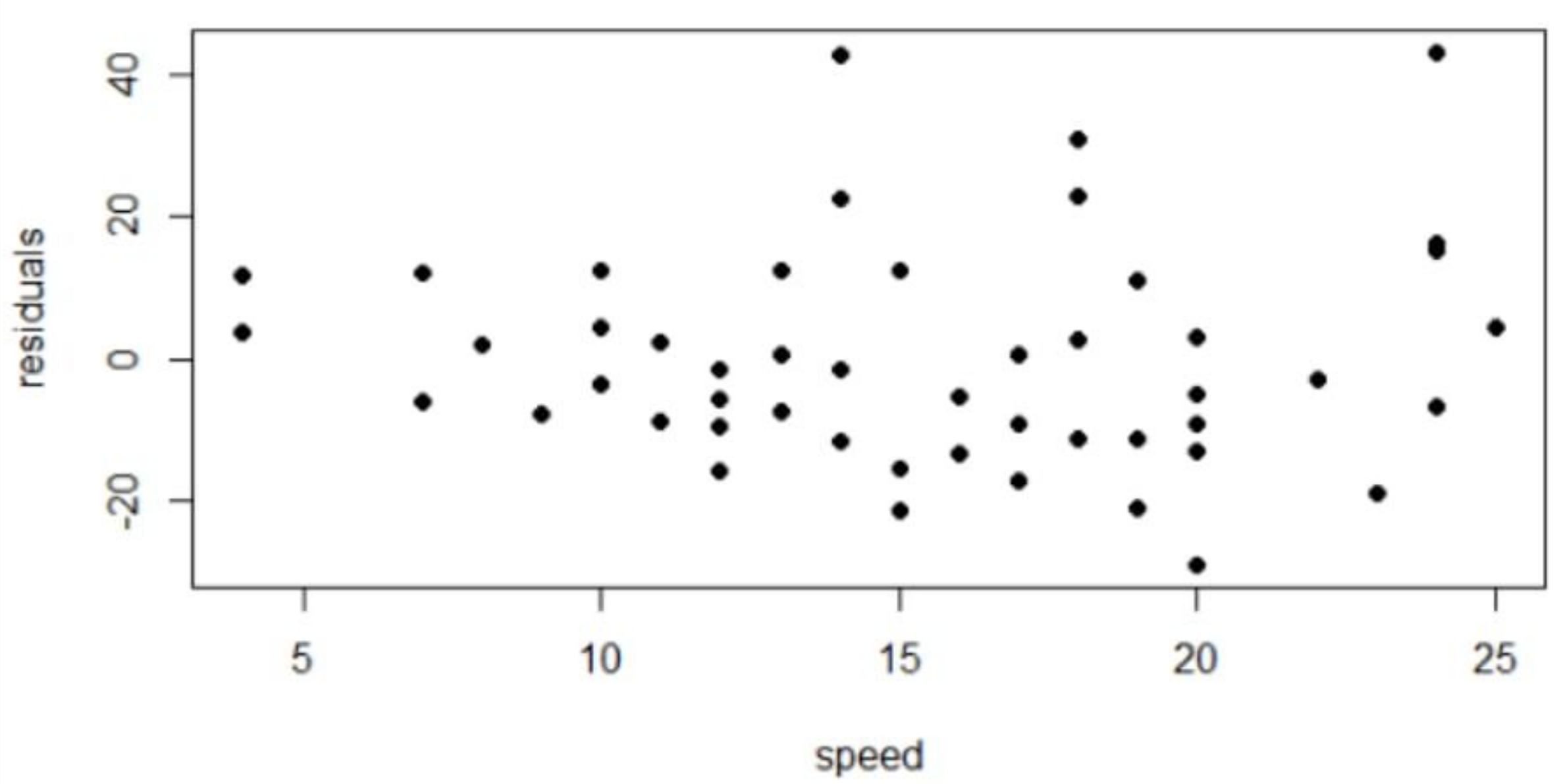
Global Environment

cars_qm	List of 12
dframe	6 obs. of 4 variables
foo	List of 3
gg	Large gg (9 elements, 3.5 MB)
lm_model	List of 12
mydata	5 obs. of 3 variables
mydata2	5 obs. of 3 variables
mymat	num [1:4, 1:3] 1 5 100 11 3 4 20 12 4 3 ...

Values

Files Plots Packages Help Viewer

Zoom Export Publish



The plot shows residuals on the y-axis (ranging from -20 to 40) against speed on the x-axis (ranging from 5 to 25). The residuals are plotted as black dots (pch=16). There is no apparent linear trend, suggesting the residuals are not correlated with the explanatory variable (speed).

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function

Addins

Project: (None)

Untitled1* x

Untitled2* x

Source on Save

Run

Source

1# Q. 9

2

3# There are several points (on the bottom-right plot of the 4-panel fig)

4# with a large Cook's distance that may have an influence on the regression,

5# including points labeled 49,23 and 39.

5:44 (Top Level) R Script

Console

Terminal

Jobs

~/

> |

Environment

History

Connections

Tutorial

Import Dataset

Global Environment

cars_qm

List of 12

dframe

6 obs. of 4 variables

foo

List of 3

gg

Large gg (9 elements, 3.5 MB)

lm_model

List of 12

mydata

5 obs. of 3 variables

mydata2

5 obs. of 3 variables

mymat

num [1:4, 1:3] 1 5 100 11 3 4 20 12 4 3 ...

Values

Files

Plots

Packages

Help

Viewer

Zoom

Export

Publish

residuals

40

20

0

-20

5

10

15

20

25

speed

Search for anything

7

ENG

16:24

29-11-2020

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function Addins

Project: (None)

Environment History Connections Tutorial

Global Environment

cars_qm	List of 12
dframe	6 obs. of 4 variables
foo	List of 3
gg	Large gg (9 elements, 3.5 MB)
lm_model1	List of 12
lm_model2	List of 12
mydata	5 obs. of 3 variables
mydata2	5 obs. of 3 variables
mymat	num [1:4, 1:3] 1 5 100 11 3 4 20 12 4 3 ...

Files Plots Packages Help Viewer

Zoom Export Publish

```
1 # Q. 10
2
3 cars_no_49 = cars[-49,]
4 lm_model2 <- lm(cars_no_49$dist~cars_no_49$speed)
5 summary(lm_model2)
6
7 #We see that removing the 49th record changed some values in the model regression
8 # but nothing to qualitatively alter the results.
9 # For ex., the estimated slope coefficient is now slightly smaller.
```

9:68 (Top Level) R Script

Console Terminal Jobs

~/

Call:
lm(formula = cars_no_49\$dist ~ cars_no_49\$speed)

Residuals:

	Min	1Q	Median	3Q	Max
	-26.789	-9.149	-1.672	8.013	43.048

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-14.0021	6.2951	-2.224	0.031 *
cars_no_49\$speed	3.6396	0.3918	9.290	3.26e-12 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 14.1 on 47 degrees of freedom
Multiple R-squared: 0.6474, Adjusted R-squared: 0.6399
F-statistic: 86.31 on 1 and 47 DF, p-value: 3.262e-12

> |

