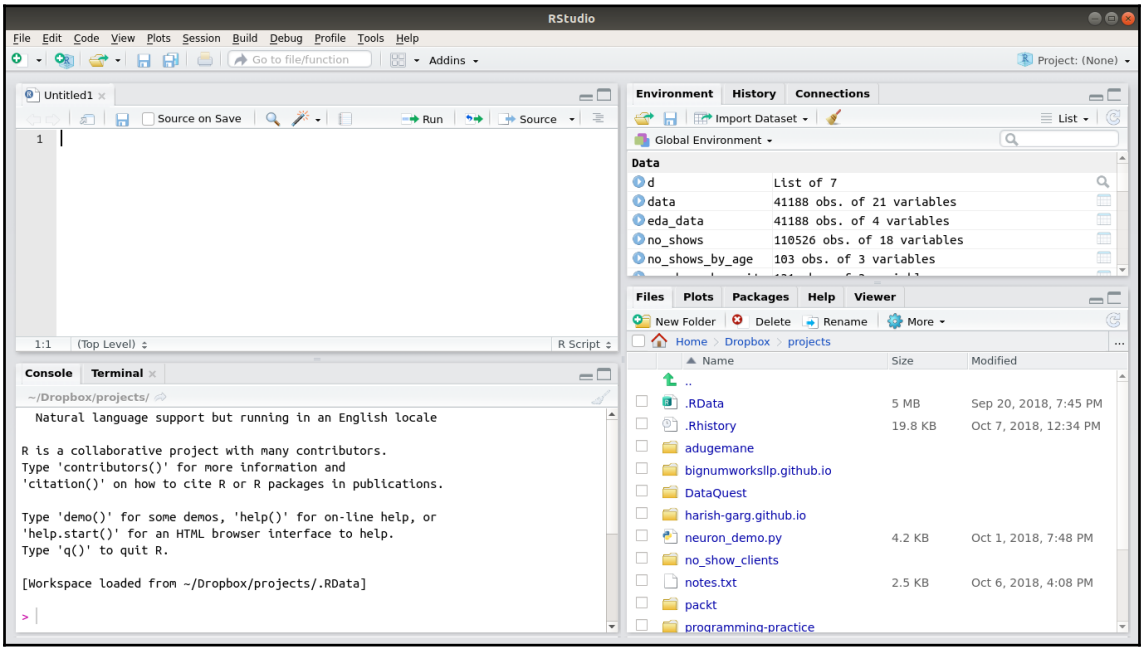
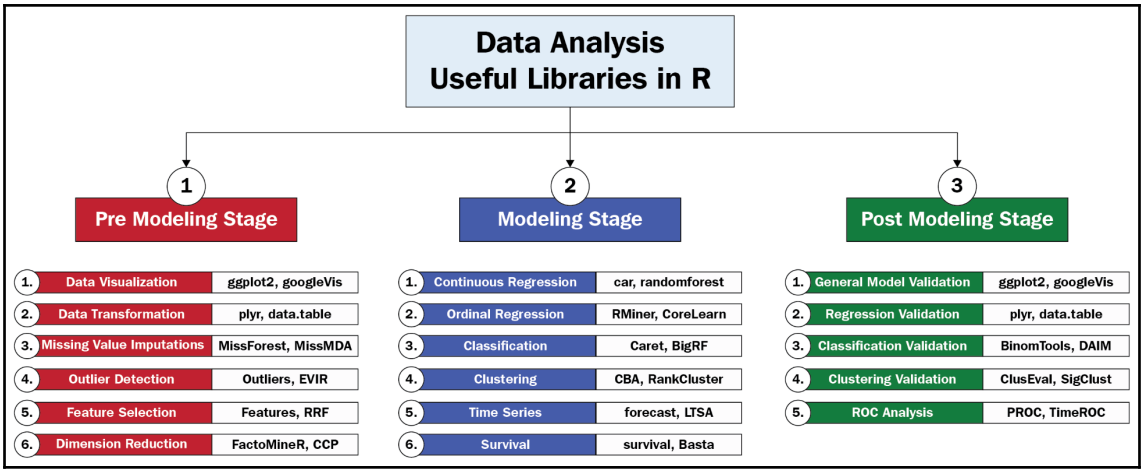
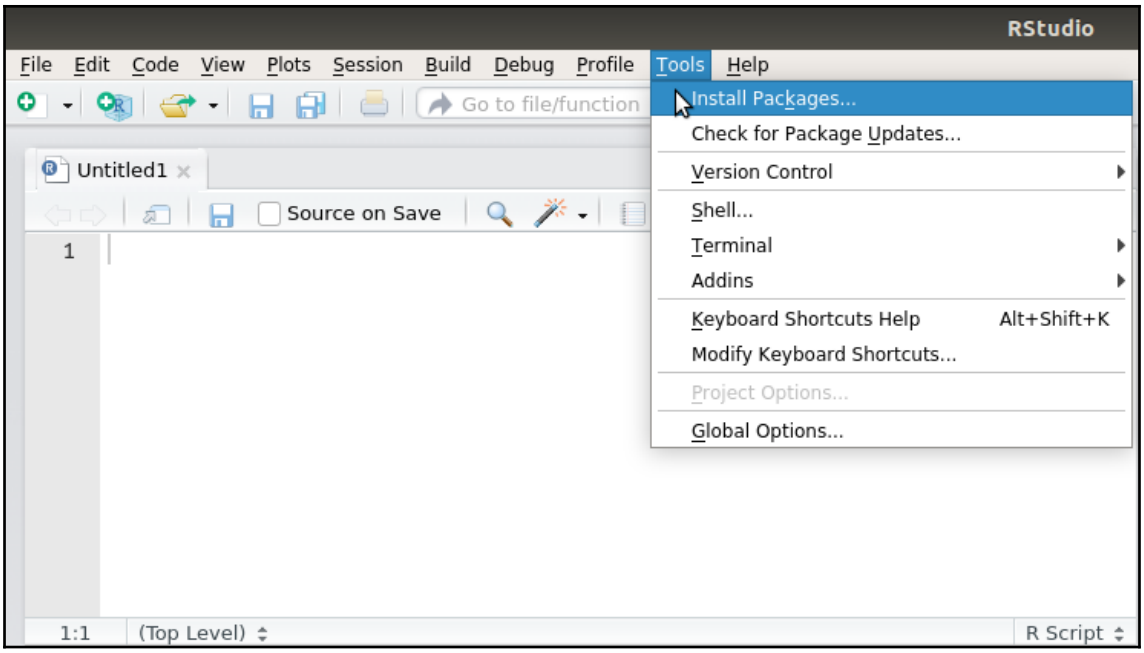
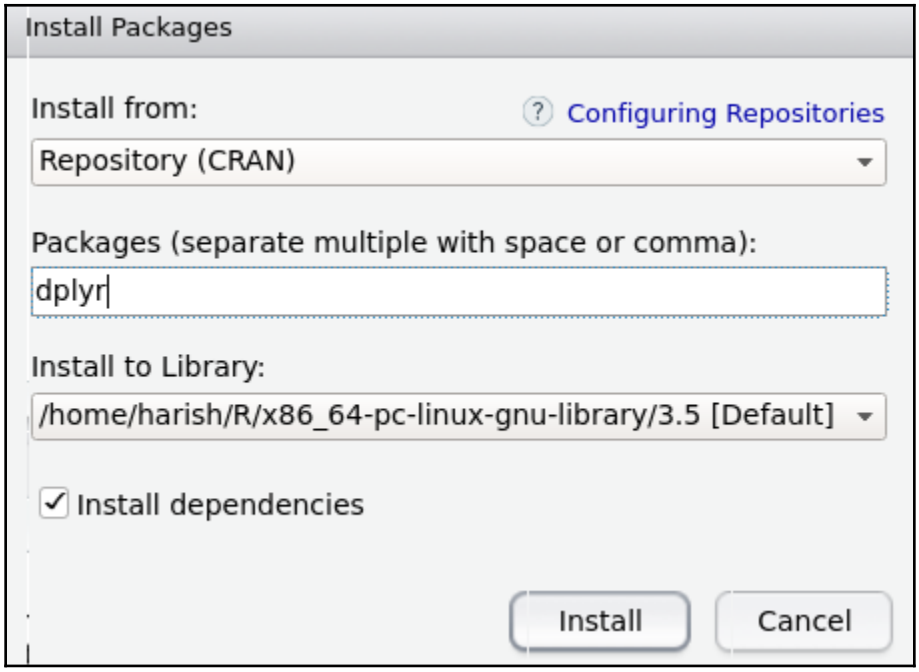
Table of Contents

Chapter 1: Setting Up Our Data Analysis Environment






RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function
Untitled1
Source on Save
1
Install Packages...
Check for Package Updates...
Version Control
Shell...
Terminal
Addins
Keyboard Shortcuts Help Alt+Shift+K
Modify Keyboard Shortcuts...
Project Options...
Global Options...
1:1
(Top Level)
R Script


Install Packages
Install from:
Configuring Repositories
Repository (CRAN)
Packages (separate multiple with space or comma):
dplyr
Install to Library:
/home/harish/R/x86_64-pc-linux-gnu-library/3.5 [Default]
Install dependencies
Install
Cancel

Chapter 2: Importing Diverse Datasets

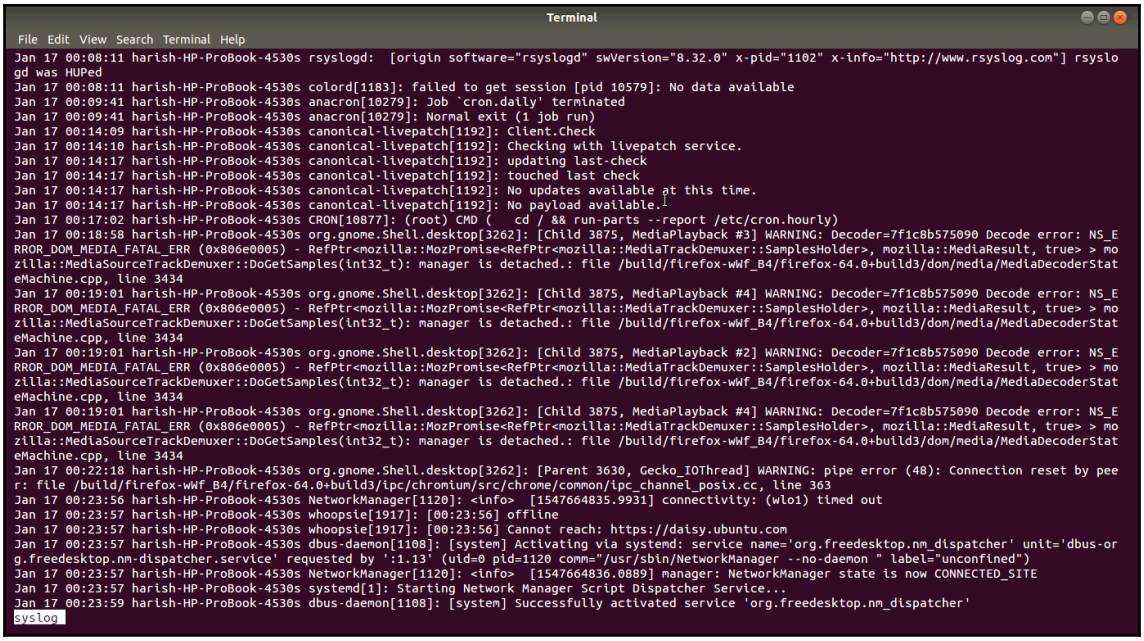
```
model,mpg,cyl,disp,hp,drat,wt,qsec,vs,am,gear,carb
Mazda RX4,21,6,160,110,3.9,2.62,16.46,0,1,4,4
Mazda RX4 Wag,21,6,160,110,3.9,2.875,17.02,0,1,4,4
Datsun 710,22.8,4,108,93,3.85,2.32,18.61,1,1,4,1
Hornet 4 Drive,21.4,6,258,110,3.08,3.215,19.44,1,0,3,1
Hornet Sportabout,18.7,8,360,175,3.15,3.44,17.02,0,0,3,2
Valiant,18.1,6,225,105,2.76,3.46,20.22,1,0,3,1
Duster 360,14.3,8,360,245,3.21,3.57,15.84,0,0,3,4
Merc 240D,24.4,4,146.7,62,3.69,3.19,20,1,0,4,2
Merc 230,22.8,4,140.8,95,3.92,3.15,22.9,1,0,4,2
Merc 280,19.2,6,167.6,123,3.92,3.44,18.3,1,0,4,4
Merc 280C,17.8,6,167.6,123,3.92,3.44,18.9,1,0,4,4
Merc 450SE,16.4,8,275.8,180,3.07,4.07,17.4,0,0,3,3
Merc 450SL,17.3,8,275.8,180,3.07,3.73,17.6,0,0,3,3
Merc 450SLC,15.2,8,275.8,180,3.07,3.78,18,0,0,3,3
Cadillac Fleetwood,10.4,8,472,205,2.93,5.25,17.98,0,0,3,4
Lincoln Continental,10.4,8,460,215,3,5.424,17.82,0,0,3,4
Chrysler Imperial,14.7,8,440,230,3.23,5.345,17.42,0,0,3,4
Fiat 128,32.4,4,78.7,66,4.08,2.2,19.47,1,1,4,1
Honda Civic,30.4,4,75.7,52,4.93,1.615,18.52,1,1,4,2
Toyota Corolla,33.9,4,71.1,65,4.22,1.835,19.9,1,1,4,1
Toyota Corona,21.5,4,120.1,97,3.7,2.465,20.01,1,0,3,1
Dodge Challenger,15.5,8,318,150,2.76,3.52,16.87,0,0,3,2
AMC Javelin,15.2,8,304,150,3.15,3.435,17.3,0,0,3,2
Camaro Z28,13.3,8,350,245,3.73,3.84,15.41,0,0,3,4
Pontiac Firebird,19.2,8,400,175,3.08,3.845,17.05,0,0,3,2
Fiat X1-9,27.3,4,79,66,4.08,1.935,18.9,1,1,4,1
Porsche 914-2,26,4,120.3,91,4.43,2.14,16.7,0,1,5,2
Lotus Europa,30.4,4,95.1,113,3.77,1.513,16.9,1,1,5,2
Ford Pantera L,15.8,8,351,264,4.22,3.17,14.5,0,1,5,4
Ferrari Dino,19.7,6,145,175,3.62,2.77,15.5,0,1,5,6
Maserati Bora,15,8,301,335,3.54,3.57,14.6,0,1,5,8
Volvo 142E,21.4,4,121,109,4.11,2.78,18.6,1,1,4,2
```

mtcars

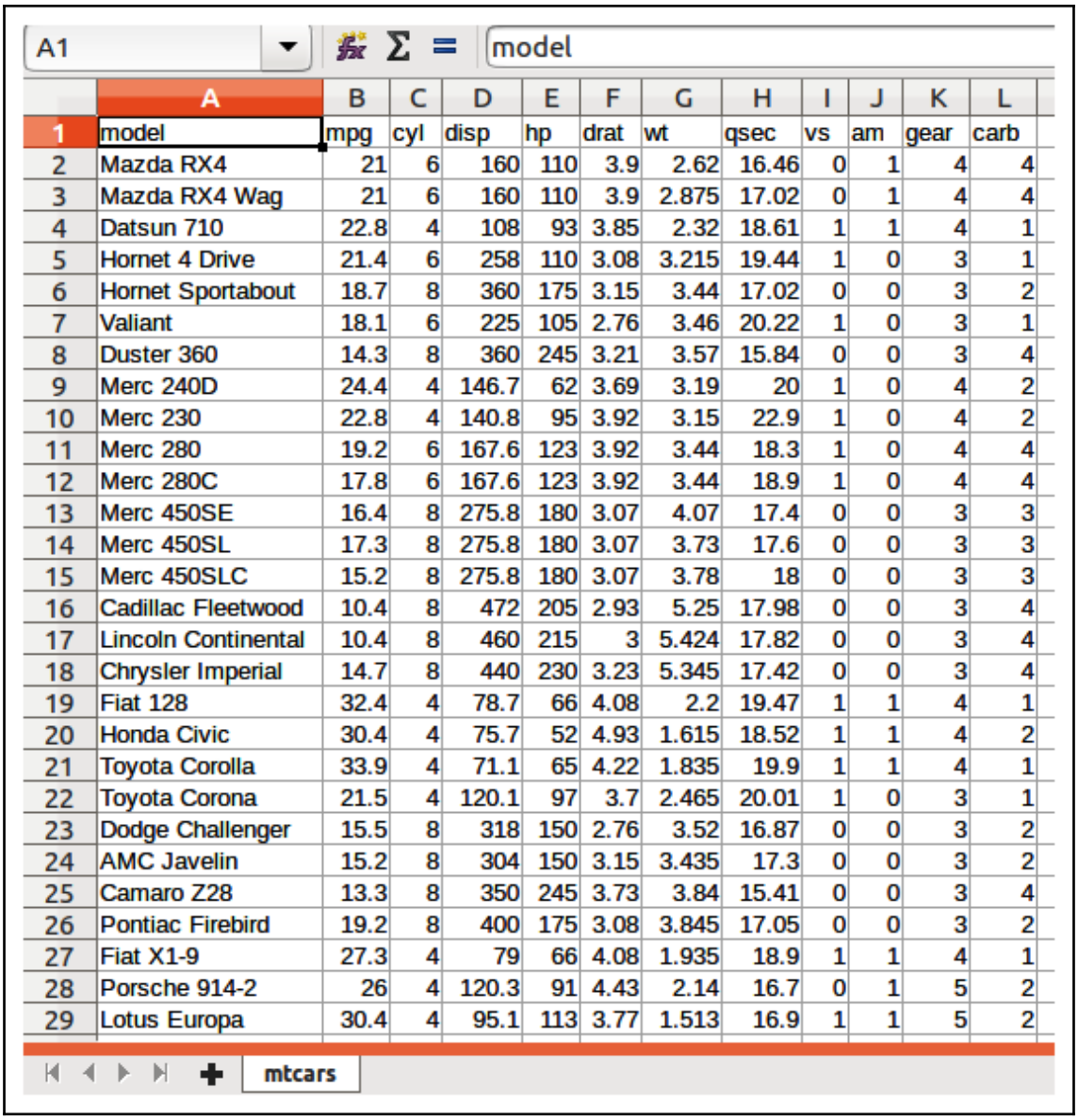
Filter

	mpg	cyl	disp	hp	drat	wt	qsec	vs	am	gear	carb
Mazda RX4	21.0	6	160.0	110	3.90	2.620	16.46	0	1	4	4
Mazda RX4 Wag	21.0	6	160.0	110	3.90	2.875	17.02	0	1	4	4
Datsun 710	22.8	4	108.0	93	3.85	2.320	18.61	1	1	4	1
Hornet 4 Drive	21.4	6	258.0	110	3.08	3.215	19.44	1	0	3	1
Hornet Sportabout	18.7	8	360.0	175	3.15	3.440	17.02	0	0	3	2
Valiant	18.1	6	225.0	105	2.76	3.460	20.22	1	0	3	1
Duster 360	14.3	8	360.0	245	3.21	3.570	15.84	0	0	3	4
Merc 240D	24.4	4	146.7	62	3.69	3.190	20.00	1	0	4	2
Merc 230	22.8	4	140.8	95	3.92	3.150	22.90	1	0	4	2
Merc 280	19.2	6	167.6	123	3.92	3.440	18.30	1	0	4	4

```
model|mpg|cyl|disp|hp|drat|wt|qsec|vs|am|gear|carb
Mazda RX4|21|6|160|110|3.9|2.62|16.46|0|1|4|4
Mazda RX4 Wag|21|6|160|110|3.9|2.875|17.02|0|1|4|4
Datsun 710|22.8|4|108|93|3.85|2.32|18.61|1|1|4|1
Hornet 4 Drive|21.4|6|258|110|3.08|3.215|19.44|1|0|3|1
Hornet Sportabout|18.7|8|360|175|3.15|3.44|17.02|0|0|3|2
Valiant|18.1|6|225|105|2.76|3.46|20.22|1|0|3|1
Duster 360|14.3|8|360|245|3.21|3.57|15.84|0|0|3|4
Merc 240D|24.4|4|146.7|62|3.69|3.19|20|1|0|4|2
Merc 230|22.8|4|140.8|95|3.92|3.15|22.9|1|0|4|2
Merc 280|19.2|6|167.6|123|3.92|3.44|18.3|1|0|4|4
Merc 280C|17.8|6|167.6|123|3.92|3.44|18.9|1|0|4|4
Merc 450SE|16.4|8|275.8|180|3.07|4.07|17.4|0|0|3|3
Merc 450SL|17.3|8|275.8|180|3.07|3.73|17.6|0|0|3|3
Merc 450SLC|15.2|8|275.8|180|3.07|3.78|18|0|0|3|3
Cadillac Fleetwood|10.4|8|472|205|2.93|5.25|17.98|0|0|3|4
Lincoln Continental|10.4|8|460|215|3|5.424|17.82|0|0|3|4
Chrysler Imperial|14.7|8|440|230|3.23|5.345|17.42|0|0|3|4
Fiat 128|32.4|4|78.7|66|4.08|2.2|19.47|1|1|4|1
Honda Civic|30.4|4|75.7|52|4.93|1.615|18.52|1|1|4|2
Toyota Corolla|33.9|4|71.1|65|4.22|1.835|19.9|1|1|4|1
Toyota Corona|21.5|4|120.1|97|3.7|2.465|20.01|1|0|3|1
Dodge Challenger|15.5|8|318|150|2.76|3.52|16.87|0|0|3|2
AMC Javelin|15.2|8|304|150|3.15|3.435|17.3|0|0|3|2
Camaro Z28|13.3|8|350|245|3.73|3.84|15.41|0|0|3|4
Pontiac Firebird|19.2|8|400|175|3.08|3.845|17.05|0|0|3|2
Fiat X1-9|27.3|4|79|66|4.08|1.935|18.9|1|1|4|1
Porsche 914-2|26|4|120.3|91|4.43|2.14|16.7|0|1|5|2
Lotus Europa|30.4|4|95.1|113|3.77|1.513|16.9|1|1|5|2
Ford Pantera L|15.8|8|351|264|4.22|3.17|14.5|0|1|5|4
Ferrari Dino|19.7|6|145|175|3.62|2.77|15.5|0|1|5|6
Maserati Bora|15|8|301|335|3.54|3.57|14.6|0|1|5|8
Volvo 142E|21.4|4|121|109|4.11|2.78|18.6|1|1|4|2
```



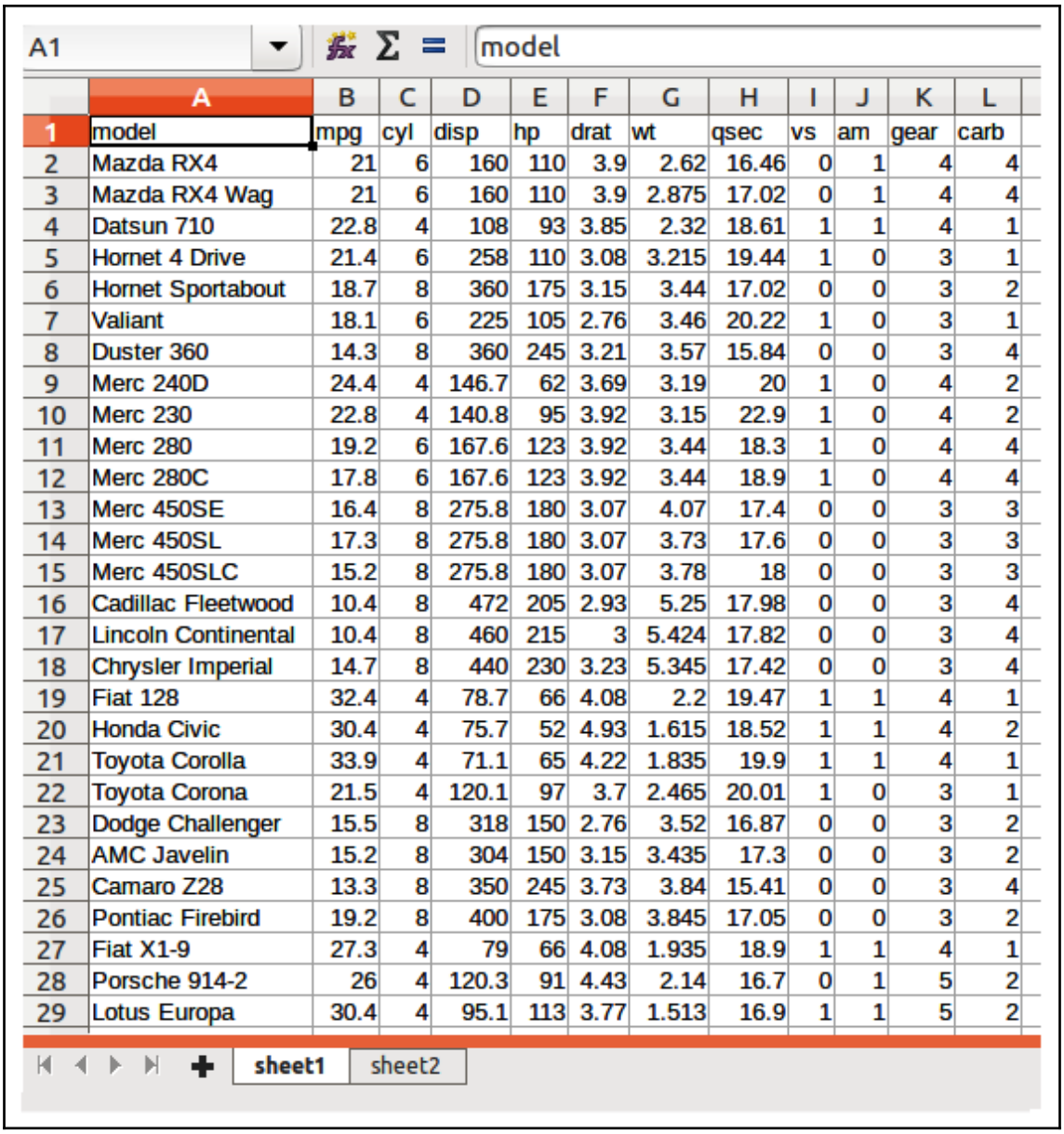

```
Parsed with column specification:
cols(
  X1 = col_character(),
  X2 = col_character(),
  X3 = col_character(),
  X4 = col_character(),
  X5 = col_character(),
  X6 = col_integer(),
  X7 = col_integer()
)
# A tibble: 2 x 7
  X1          X2    X3              X4                  X5
      X6    X7
  <chr>       <chr> <chr>           <chr>               <chr>
    <int> <int>
1 172.21.13… NA    "Microsoft\\Jo… 08/Apr/2001:17:3… GET /scripts/iisadmin/ism.dll
?…    200  3401
2 127.0.0.1  NA    frank           10/Oct/2000:13:5… GET /apache_pb.gif HTTP/1.0
      200  2326
> |
```



A1 | model

	A	B	C	D	E	F	G	H	I	J	K	L
1	model	mpg	cyl	disp	hp	drat	wt	qsec	vs	am	gear	carb
2	Mazda RX4	21	6	160	110	3.9	2.62	16.46	0	1	4	4
3	Mazda RX4 Wag	21	6	160	110	3.9	2.875	17.02	0	1	4	4
4	Datsun 710	22.8	4	108	93	3.85	2.32	18.61	1	1	4	1
5	Hornet 4 Drive	21.4	6	258	110	3.08	3.215	19.44	1	0	3	1
6	Hornet Sportabout	18.7	8	360	175	3.15	3.44	17.02	0	0	3	2
7	Valiant	18.1	6	225	105	2.76	3.46	20.22	1	0	3	1
8	Duster 360	14.3	8	360	245	3.21	3.57	15.84	0	0	3	4
9	Merc 240D	24.4	4	146.7	62	3.69	3.19	20	1	0	4	2
10	Merc 230	22.8	4	140.8	95	3.92	3.15	22.9	1	0	4	2
11	Merc 280	19.2	6	167.6	123	3.92	3.44	18.3	1	0	4	4
12	Merc 280C	17.8	6	167.6	123	3.92	3.44	18.9	1	0	4	4
13	Merc 450SE	16.4	8	275.8	180	3.07	4.07	17.4	0	0	3	3
14	Merc 450SL	17.3	8	275.8	180	3.07	3.73	17.6	0	0	3	3
15	Merc 450SLC	15.2	8	275.8	180	3.07	3.78	18	0	0	3	3
16	Cadillac Fleetwood	10.4	8	472	205	2.93	5.25	17.98	0	0	3	4
17	Lincoln Continental	10.4	8	460	215	3	5.424	17.82	0	0	3	4
18	Chrysler Imperial	14.7	8	440	230	3.23	5.345	17.42	0	0	3	4
19	Fiat 128	32.4	4	78.7	66	4.08	2.2	19.47	1	1	4	1
20	Honda Civic	30.4	4	75.7	52	4.93	1.615	18.52	1	1	4	2
21	Toyota Corolla	33.9	4	71.1	65	4.22	1.835	19.9	1	1	4	1
22	Toyota Corona	21.5	4	120.1	97	3.7	2.465	20.01	1	0	3	1
23	Dodge Challenger	15.5	8	318	150	2.76	3.52	16.87	0	0	3	2
24	AMC Javelin	15.2	8	304	150	3.15	3.435	17.3	0	0	3	2
25	Camaro Z28	13.3	8	350	245	3.73	3.84	15.41	0	0	3	4
26	Pontiac Firebird	19.2	8	400	175	3.08	3.845	17.05	0	0	3	2
27	Fiat X1-9	27.3	4	79	66	4.08	1.935	18.9	1	1	4	1
28	Porsche 914-2	26	4	120.3	91	4.43	2.14	16.7	0	1	5	2
29	Lotus Europa	30.4	4	95.1	113	3.77	1.513	16.9	1	1	5	2

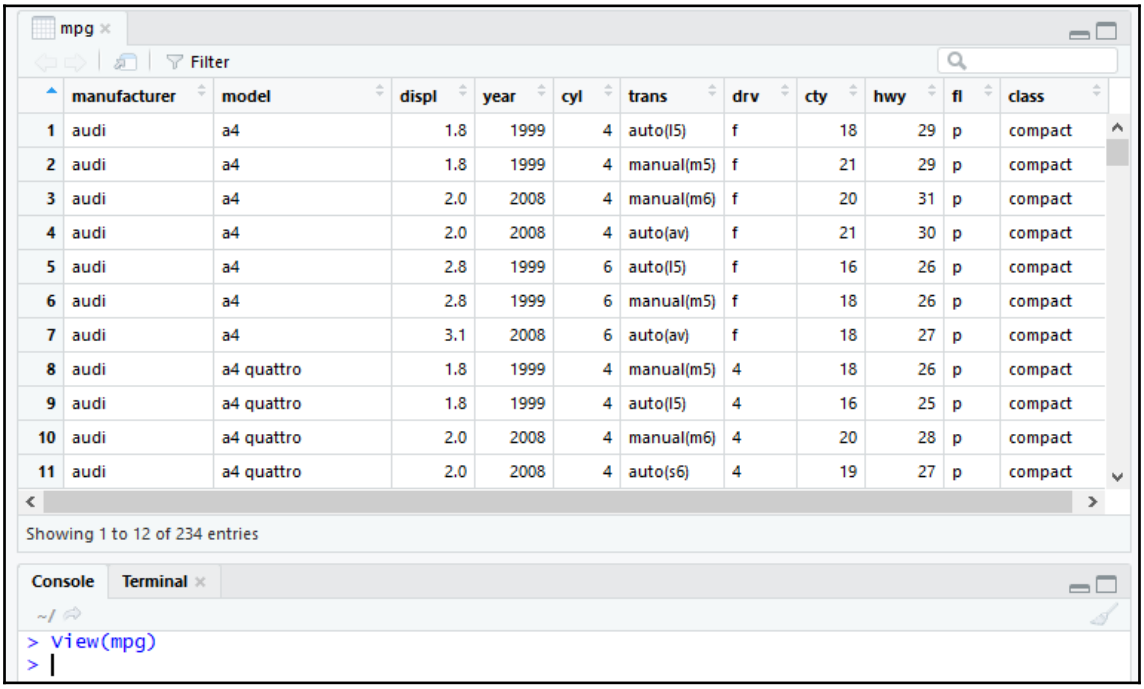
mtcars


A1 model

	A	B	C	D	E	F	G	H	I	J	K	L
1	model	mpg	cyl	disp	hp	drat	wt	qsec	vs	am	gear	carb
2	Mazda RX4	21	6	160	110	3.9	2.62	16.46	0	1	4	4
3	Mazda RX4 Wag	21	6	160	110	3.9	2.875	17.02	0	1	4	4
4	Datsun 710	22.8	4	108	93	3.85	2.32	18.61	1	1	4	1
5	Hornet 4 Drive	21.4	6	258	110	3.08	3.215	19.44	1	0	3	1
6	Hornet Sportabout	18.7	8	360	175	3.15	3.44	17.02	0	0	3	2
7	Valiant	18.1	6	225	105	2.76	3.46	20.22	1	0	3	1
8	Duster 360	14.3	8	360	245	3.21	3.57	15.84	0	0	3	4
9	Merc 240D	24.4	4	146.7	62	3.69	3.19	20	1	0	4	2
10	Merc 230	22.8	4	140.8	95	3.92	3.15	22.9	1	0	4	2
11	Merc 280	19.2	6	167.6	123	3.92	3.44	18.3	1	0	4	4
12	Merc 280C	17.8	6	167.6	123	3.92	3.44	18.9	1	0	4	4
13	Merc 450SE	16.4	8	275.8	180	3.07	4.07	17.4	0	0	3	3
14	Merc 450SL	17.3	8	275.8	180	3.07	3.73	17.6	0	0	3	3
15	Merc 450SLC	15.2	8	275.8	180	3.07	3.78	18	0	0	3	3
16	Cadillac Fleetwood	10.4	8	472	205	2.93	5.25	17.98	0	0	3	4
17	Lincoln Continental	10.4	8	460	215	3	5.424	17.82	0	0	3	4
18	Chrysler Imperial	14.7	8	440	230	3.23	5.345	17.42	0	0	3	4
19	Fiat 128	32.4	4	78.7	66	4.08	2.2	19.47	1	1	4	1
20	Honda Civic	30.4	4	75.7	52	4.93	1.615	18.52	1	1	4	2
21	Toyota Corolla	33.9	4	71.1	65	4.22	1.835	19.9	1	1	4	1
22	Toyota Corona	21.5	4	120.1	97	3.7	2.465	20.01	1	0	3	1
23	Dodge Challenger	15.5	8	318	150	2.76	3.52	16.87	0	0	3	2
24	AMC Javelin	15.2	8	304	150	3.15	3.435	17.3	0	0	3	2
25	Camaro Z28	13.3	8	350	245	3.73	3.84	15.41	0	0	3	4
26	Pontiac Firebird	19.2	8	400	175	3.08	3.845	17.05	0	0	3	2
27	Fiat X1-9	27.3	4	79	66	4.08	1.935	18.9	1	1	4	1
28	Porsche 914-2	26	4	120.3	91	4.43	2.14	16.7	0	1	5	2
29	Lotus Europa	30.4	4	95.1	113	3.77	1.513	16.9	1	1	5	2

sheet1 sheet2

Chapter 3: Examining, Cleaning, and Filtering



mpg

Filter

	manufacturer	model	displ	year	cyl	trans	drv	cty	hwy	fl	class
1	audi	a4	1.8	1999	4	auto(l5)	f	18	29	p	compact
2	audi	a4	1.8	1999	4	manual(m5)	f	21	29	p	compact
3	audi	a4	2.0	2008	4	manual(m6)	f	20	31	p	compact
4	audi	a4	2.0	2008	4	auto(av)	f	21	30	p	compact
5	audi	a4	2.8	1999	6	auto(l5)	f	16	26	p	compact
6	audi	a4	2.8	1999	6	manual(m5)	f	18	26	p	compact
7	audi	a4	3.1	2008	6	auto(av)	f	18	27	p	compact
8	audi	a4 quattro	1.8	1999	4	manual(m5)	4	18	26	p	compact
9	audi	a4 quattro	1.8	1999	4	auto(l5)	4	16	25	p	compact
10	audi	a4 quattro	2.0	2008	4	manual(m6)	4	20	28	p	compact
11	audi	a4 quattro	2.0	2008	4	auto(s6)	4	19	27	p	compact
12	audi	a4 quattro	2.8	1999	6	auto(l5)	4	15	25	p	compact

Showing 1 to 12 of 234 entries

```
> install.packages("tidyr")
Installing package into 'C:/Users/Radhika/Documents/R/win-library/3.5'
(as 'lib' is unspecified)
trying URL 'https://cran.rstudio.com/bin/windows/contrib/3.5/tidyr_0.8.3.zip'
Content type 'application/zip' length 953866 bytes (931 KB)
downloaded 931 KB

package 'tidyr' successfully unpacked and MD5 sums checked

The downloaded binary packages are in
        C:\Users\Radhika\AppData\Local\Temp\Rtmpgt3r8E\downloaded_packages
```

Console Terminal ×

~/

```
> library(tidyr)

Attaching package: 'tidyr'

The following object is masked _by_ '.GlobalEnv':

    who

The following object is masked from 'package:Matrix':

    expand

Warning message:
package 'tidyr' was built under R version 3.5.3
>
```

mpg × mpg2 ×

Filter

	manufacturer	model	displ	cyl	trans	drv	cty	hwy	fl	class	mpg	Year of Establishment
4	audi	a4	2.0	4	auto(av)	f	21	30	p	compact	year	2008
5	audi	a4	2.8	6	auto(l5)	f	16	26	p	compact	year	1999
6	audi	a4	2.8	6	manual(m5)	f	18	26	p	compact	year	1999
7	audi	a4	3.1	6	auto(av)	f	18	27	p	compact	year	2008
8	audi	a4 quattro	1.8	4	manual(m5)	4	18	26	p	compact	year	1999
9	audi	a4 quattro	1.8	4	auto(l5)	4	16	25	p	compact	year	1999
10	audi	a4 quattro	2.0	4	manual(m6)	4	20	28	p	compact	year	2008
11	audi	a4 quattro	2.0	4	auto(s6)	4	19	27	p	compact	year	2008
12	audi	a4 quattro	2.8	6	auto(l5)	4	15	25	p	compact	year	1999
13	audi	a4 quattro	2.8	6	manual(m5)	4	17	25	p	compact	year	1999
14	audi	a4 quattro	3.1	6	auto(s6)	4	17	25	p	compact	year	2008

mpg

Filter

	manufacturer	model	displ	year	cyl	trans	drv	cty	hwy	fl	class
1	audi	a4	1.8	1999	4	auto(l5)	f	18	29	p	compact
2	audi	a4	1.8	1999	4	manual(m5)	f	21	29	p	compact
3	audi	a4	2.0	2008	4	manual(m6)	f	20	31	p	compact
4	audi	a4	2.0	2008	4	auto(av)	f	21	30	p	compact
5	audi	a4	2.8	1999	6	auto(l5)	f	16	26	p	compact
6	audi	a4	2.8	1999	6	manual(m5)	f	18	26	p	compact
7	audi	a4	3.1	2008	6	auto(av)	f	18	27	p	compact
8	audi	a4 quattro	1.8	1999	4	manual(m5)	4	18	26	p	compact
9	audi	a4 quattro	1.8	1999	4	auto(l5)	4	16	25	p	compact
10	audi	a4 quattro	2.0	2008	4	manual(m6)	4	20	28	p	compact
11	audi	a4 quattro	2.0	2008	4	auto(s6)	4	19	27	p	compact

Showing 1 to 12 of 234 entries

Console Terminal

```
> view(mpg)
>
```

mpg4

Filter

	manufacturer	model	displ	year	cyl	trans	FuelEfficiency	cty	hwy	class
1	audi	a4	1.8	1999	4	auto(l5)	f_p	18	29	compact
2	audi	a4	1.8	1999	4	manual(m5)	f_p	21	29	compact
3	audi	a4	2.0	2008	4	manual(m6)	f_p	20	31	compact
4	audi	a4	2.0	2008	4	auto(av)	f_p	21	30	compact
5	audi	a4	2.8	1999	6	auto(l5)	f_p	16	26	compact
6	audi	a4	2.8	1999	6	manual(m5)	f_p	18	26	compact
7	audi	a4	3.1	2008	6	auto(av)	f_p	18	27	compact
8	audi	a4 quattro	1.8	1999	4	manual(m5)	4_p	18	26	compact
9	audi	a4 quattro	1.8	1999	4	auto(l5)	4_p	16	25	compact
10	audi	a4 quattro	2.0	2008	4	manual(m6)	4_p	20	28	compact

mpg4 mpg3

Filter

	manufacturer	model	displ	year	cyl	trans	drv	f1	cty	hwy	class
1	audi	a4	1.8	1999	4	auto(l5)	f	p	18	29	compact
2	audi	a4	1.8	1999	4	manual(m5)	f	p	21	29	compact
3	audi	a4	2.0	2008	4	manual(m6)	f	p	20	31	compact
4	audi	a4	2.0	2008	4	auto(av)	f	p	21	30	compact
5	audi	a4	2.8	1999	6	auto(l5)	f	p	16	26	compact
6	audi	a4	2.8	1999	6	manual(m5)	f	p	18	26	compact
7	audi	a4	3.1	2008	6	auto(av)	f	p	18	27	compact
8	audi	a4 quattro	1.8	1999	4	manual(m5)	4	p	18	26	compact
9	audi	a4 quattro	1.8	1999	4	auto(l5)	4	p	16	25	compact
10	audi	a4 quattro	2.0	2008	4	manual(m6)	4	p	20	28	compact

```
> install.packages("dplyr")
Installing package into 'C:/Users/Radhika/Documents/R/win-library/3.5'
(as 'lib' is unspecified)
Warning in install.packages :
  package 'dplyr' is in use and will not be installed
>
```

Console Terminal

~/

```
> library(dplyr)
>
```

mpgMutate

Filter

	manufacturer	model	displ	year	cyl	trans	drv	cty	hwy	fl	class	nv
1	audi	a4	1.8	1999	4	auto(l5)	f	18	29	p	compact	5.8
2	audi	a4	1.8	1999	4	manual(m5)	f	21	29	p	compact	5.8
3	audi	a4	2.0	2008	4	manual(m6)	f	20	31	p	compact	6.0
4	audi	a4	2.0	2008	4	auto(av)	f	21	30	p	compact	6.0
5	audi	a4	2.8	1999	6	auto(l5)	f	16	26	p	compact	8.8
6	audi	a4	2.8	1999	6	manual(m5)	f	18	26	p	compact	8.8
7	audi	a4	3.1	2008	6	auto(av)	f	18	27	p	compact	9.1
8	audi	a4 quattro	1.8	1999	4	manual(m5)	4	18	26	p	compact	5.8
9	audi	a4 quattro	1.8	1999	4	auto(l5)	4	16	25	p	compact	5.8
10	audi	a4 quattro	2.0	2008	4	manual(m6)	4	20	28	p	compact	6.0

Showing 1 to 11 of 234 entries

mpgGroupBy

Filter

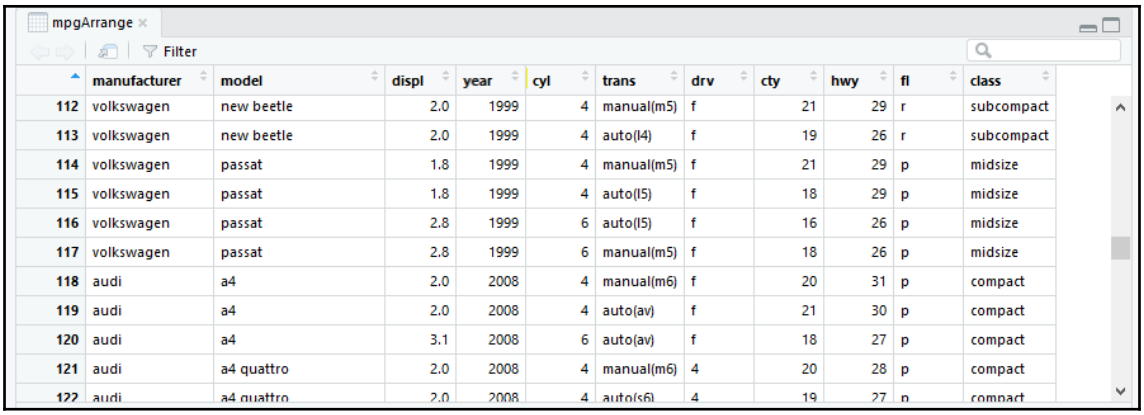
	manufacturer	model	displ	year	cyl	trans	drv	cty	hwy	fl	class
1	audi	a4	1.8	1999	4	auto(l5)	f	18	29	p	compact
2	audi	a4	1.8	1999	4	manual(m5)	f	21	29	p	compact
3	audi	a4	2.0	2008	4	manual(m6)	f	20	31	p	compact
4	audi	a4	2.0	2008	4	auto(av)	f	21	30	p	compact
5	audi	a4	2.8	1999	6	auto(l5)	f	16	26	p	compact
6	audi	a4	2.8	1999	6	manual(m5)	f	18	26	p	compact
7	audi	a4	3.1	2008	6	auto(av)	f	18	27	p	compact
8	audi	a4 quattro	1.8	1999	4	manual(m5)	4	18	26	p	compact
9	audi	a4 quattro	1.8	1999	4	auto(l5)	4	16	25	p	compact
10	audi	a4 quattro	2.0	2008	4	manual(m6)	4	20	28	p	compact
11	audi	a4 quattro	2.0	2008	4	auto(s6)	4	19	27	p	compact
12	audi	a4 quattro	2.8	1999	6	auto(l5)	4	15	25	p	compact

Showing 1 to 12 of 234 entries

mpgSummarize

Filter

	displ	avg_displ
1	1.6	1.6
2	1.8	1.8
3	1.9	1.9
4	2.0	2.0
5	2.2	2.2
6	2.4	2.4
7	2.5	2.5
8	2.7	2.7
9	2.8	2.8
10	3.0	3.0
11	3.1	3.1
12	3.3	3.3



mpgArrange

Filter

	manufacturer	model	displ	year	cyl	trans	drv	cty	hwy	fl	class
112	volkswagen	new beetle	2.0	1999	4	manual(m5)	f	21	29	r	subcompact
113	volkswagen	new beetle	2.0	1999	4	auto(l4)	f	19	26	r	subcompact
114	volkswagen	passat	1.8	1999	4	manual(m5)	f	21	29	p	midsize
115	volkswagen	passat	1.8	1999	4	auto(l5)	f	18	29	p	midsize
116	volkswagen	passat	2.8	1999	6	auto(l5)	f	16	26	p	midsize
117	volkswagen	passat	2.8	1999	6	manual(m5)	f	18	26	p	midsize
118	audi	a4	2.0	2008	4	manual(m6)	f	20	31	p	compact
119	audi	a4	2.0	2008	4	auto(av)	f	21	30	p	compact
120	audi	a4	3.1	2008	6	auto(av)	f	18	27	p	compact
121	audi	a4 quattro	2.0	2008	4	manual(m6)	4	20	28	p	compact
122	audi	a4 quattro	2.0	2008	4	auto(s6)	4	19	27	p	compact

```
> glimpse(mpg)
Observations: 234
Variables: 11
$ manufacturer <chr> "audi", "audi", "audi", "audi", "audi", "audi", "audi", "audi", "audi", "audi", "audi...
$ model        <chr> "a4", "a4", "a4", "a4", "a4", "a4", "a4", "a4 quattro", "a4 quattro", "a4 quattro", "...
$ displ        <dbl> 1.8, 1.8, 2.0, 2.0, 2.8, 2.8, 3.1, 1.8, 1.8, 2.0, 2.0, 2.8, 2.8, 3.1, 3.1, 2.8, 3.1, ...
$ year         <int> 1999, 1999, 2008, 2008, 1999, 1999, 2008, 1999, 1999, 2008, 2008, 1999, 1999, 2008, 2...
$ cyl          <int> 4, 4, 4, 4, 6, 6, 6, 4, 4, 4, 6, 6, 6, 6, 6, 6, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8...
$ trans        <chr> "auto(l5)", "manual(m5)", "manual(m6)", "auto(av)", "auto(l5)", "manual(m5)", "auto(a...
$ drv          <chr> "f", "f", "f", "f", "f", "f", "f", "4", "4", "4", "4", "4", "4", "4", "4", "4", "4", ...
$ cty          <int> 18, 21, 20, 21, 16, 18, 18, 18, 16, 20, 19, 15, 17, 17, 15, 15, 17, 16, 14, 11, 14, 1...
$ hwy          <int> 29, 29, 31, 30, 26, 26, 27, 26, 25, 28, 27, 25, 25, 25, 25, 24, 25, 23, 20, 15, 20, 1...
$ fl           <chr> "p", "p", "p", "p", "p", "p", "p", "p", "p", "p", "p", "p", "p", "p", "p", "p", "p", ...
$ class        <chr> "compact", "compact", "compact", "compact", "compact", "compact", "compact", "compact...
>
```

mpgSubset

Filter

	manufacturer	model
1	audi	a4
2	audi	a4
3	audi	a4
4	audi	a4
5	audi	a4
6	audi	a4
7	audi	a4
8	audi	a4 quattro
9	audi	a4 quattro
10	audi	a4 quattro
11	audi	a4 quattro
12	audi	a4 quattro

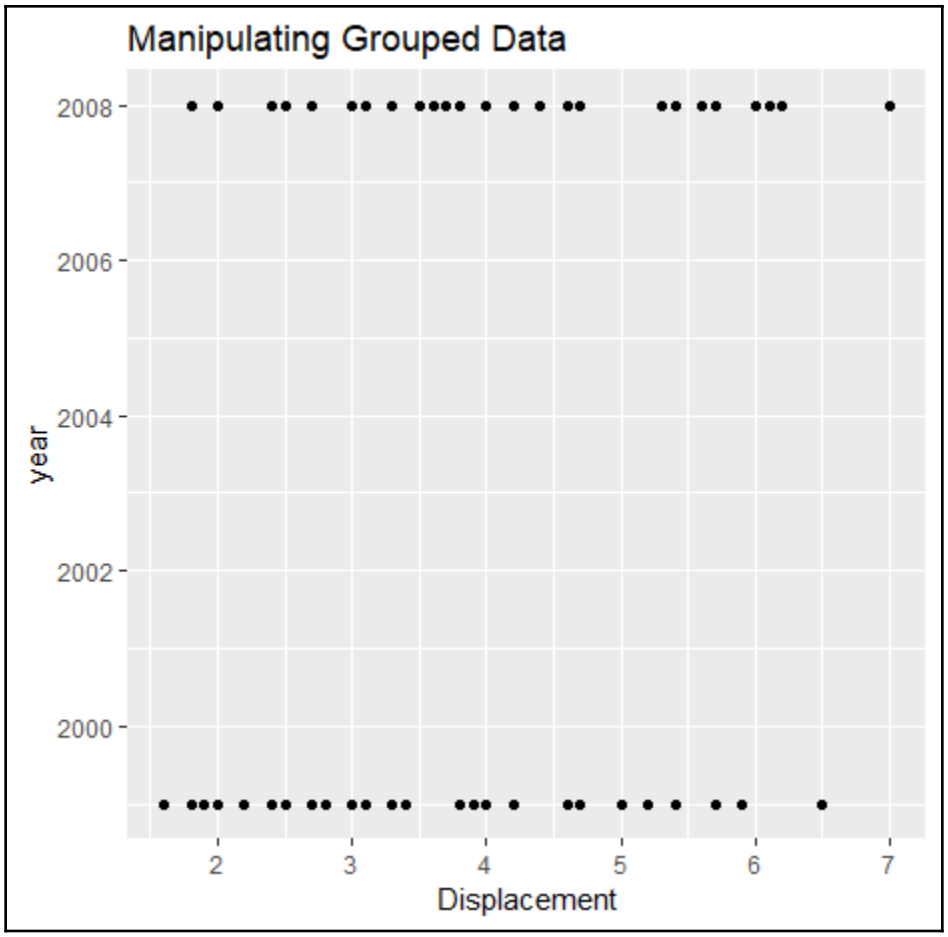
Showing 1 to 12 of 234 entries

mpgFilter

Filter

	manufacturer	model	displ	year	cyl	trans	drv	cty	hwy	fl	class
1	audi	a4	2.0	2008	4	manual(m6)	f	20	31	p	compact
2	audi	a4	2.0	2008	4	auto(av)	f	21	30	p	compact
3	audi	a4	3.1	2008	6	auto(av)	f	18	27	p	compact
4	audi	a4 quattro	2.0	2008	4	manual(m6)	4	20	28	p	compact
5	audi	a4 quattro	2.0	2008	4	auto(s6)	4	19	27	p	compact
6	audi	a4 quattro	3.1	2008	6	auto(s6)	4	17	25	p	compact
7	audi	a4 quattro	3.1	2008	6	manual(m6)	4	15	25	p	compact
8	audi	a6 quattro	3.1	2008	6	auto(s6)	4	17	25	p	midsize
9	audi	a6 quattro	4.2	2008	8	auto(s6)	4	16	23	p	midsize
10	chevrolet	c1500 suburban 2wd	5.3	2008	8	auto(l4)	r	14	20	r	suv
11	chevrolet	c1500 suburban 2wd	5.3	2008	8	auto(l4)	r	11	15	e	suv
12	chevrolet	c1500 suburban 2wd	5.3	2008	8	auto(l4)	r	14	20	r	suv

Showing 1 to 12 of 117 entries

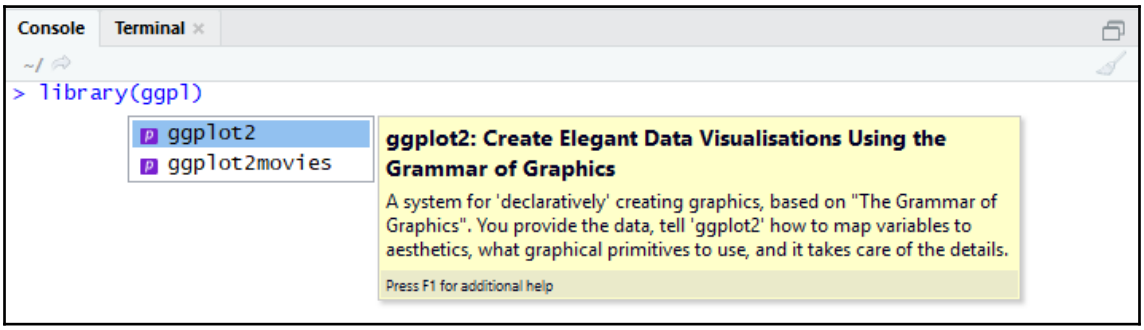

Manipulating Grouped Data
2008
2006
2004
2002
2000
year
2
3
4
5
6
7
Displacement

Chapter 4: Visualizing Data Graphically with ggplot2

```
> install.packages("ggplot2")
Installing package into 'C:/Users/Radhika/Documents/R/win-library/3.5'
(as 'lib' is unspecified)
trying URL 'https://cran.rstudio.com/bin/windows/contrib/3.5/ggplot2_3.1.1.zip'
Content type 'application/zip' length 3622219 bytes (3.5 MB)
downloaded 3.5 MB

package 'ggplot2' successfully unpacked and MD5 sums checked

The downloaded binary packages are in
        C:\Users\Radhika\AppData\Local\Temp\RtmpmeKTcs\downloaded_packages
```



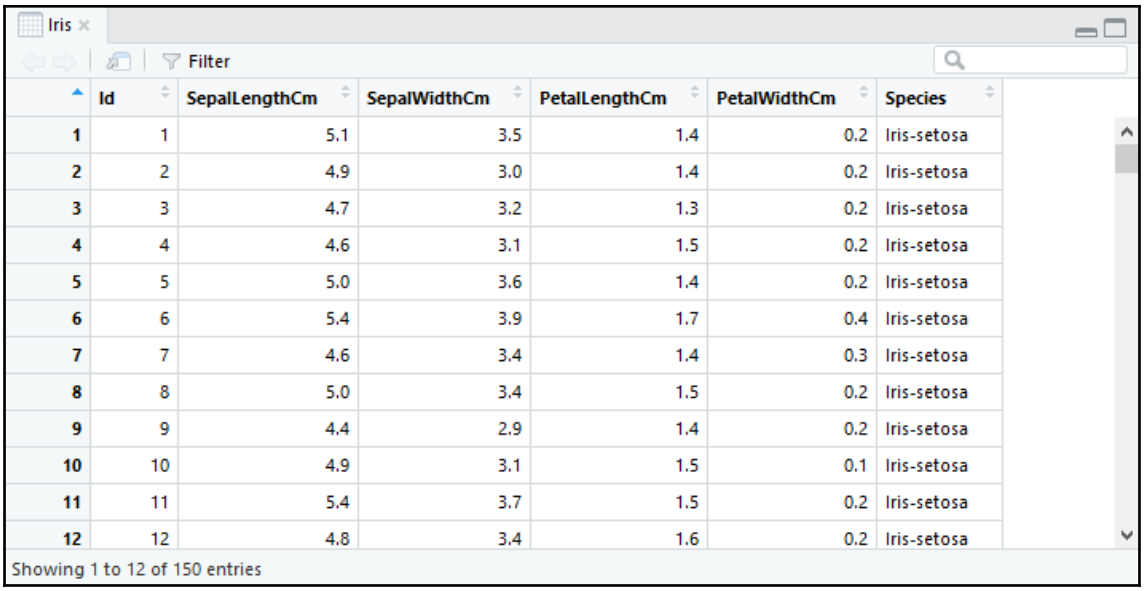

```
> library(ggplot2)

Attaching package: 'ggplot2'

The following object is masked _by_ '.GlobalEnv':

    mpg

Warning message:
package 'ggplot2' was built under R version 3.5.3
>
```

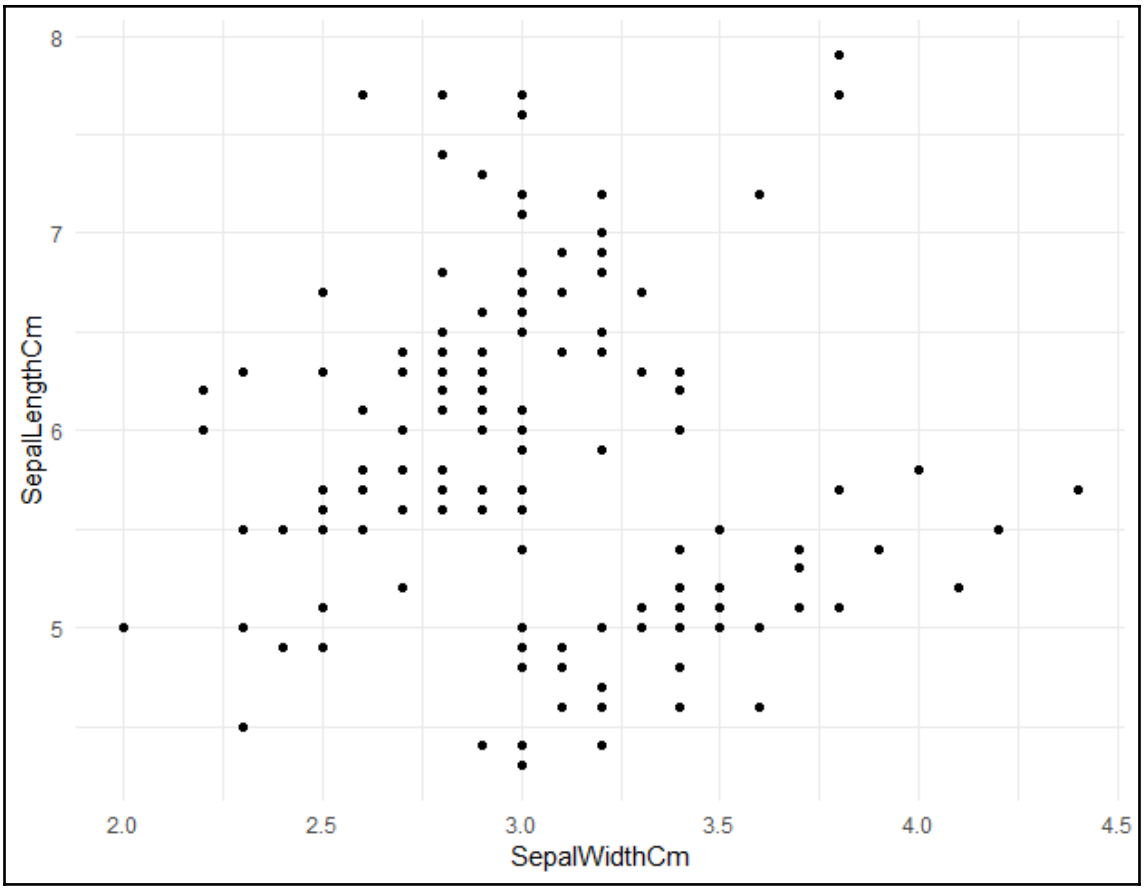


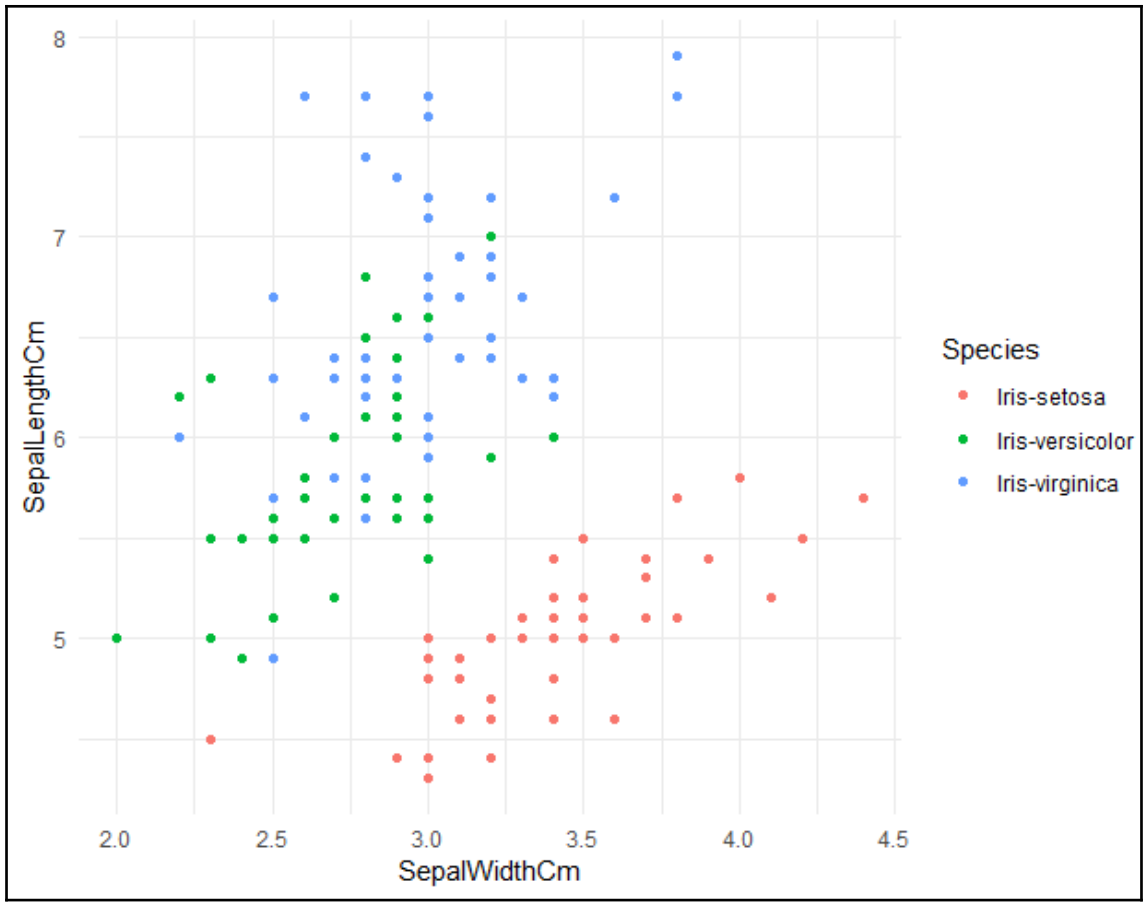
Iris

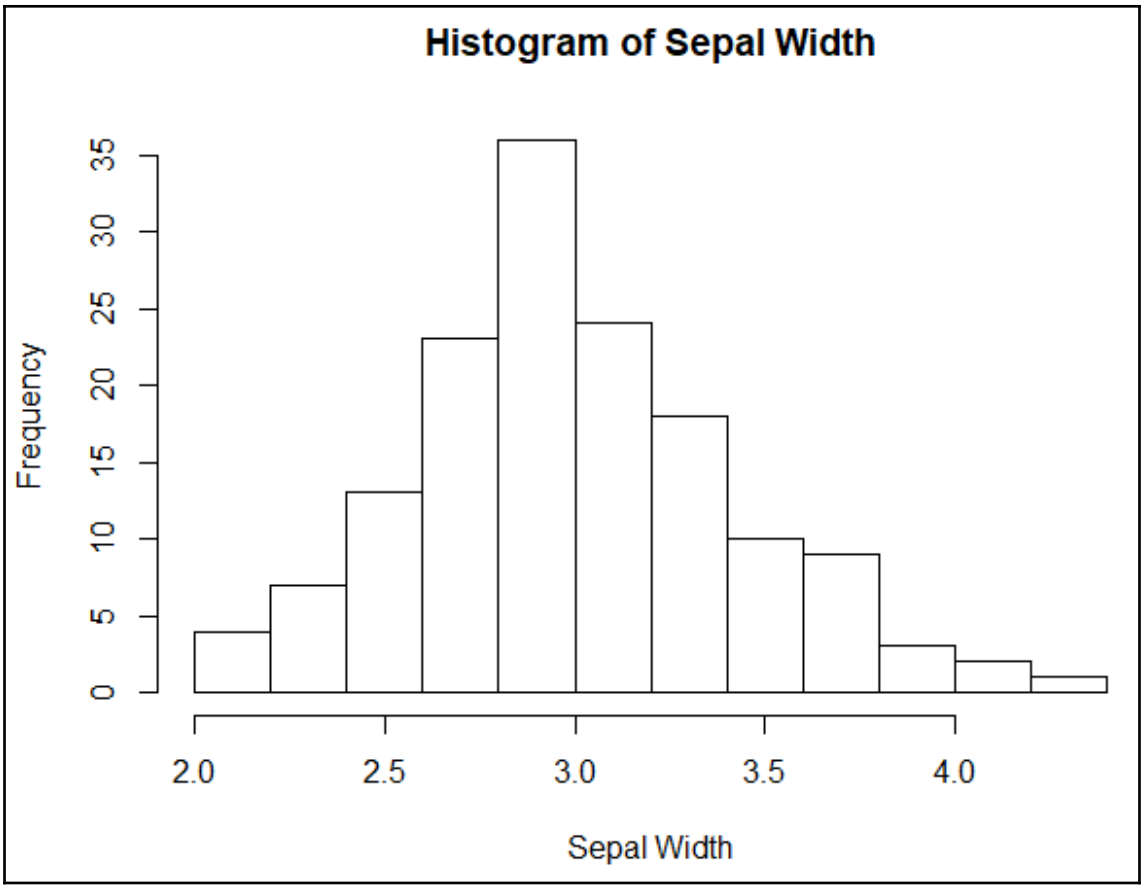
Filter

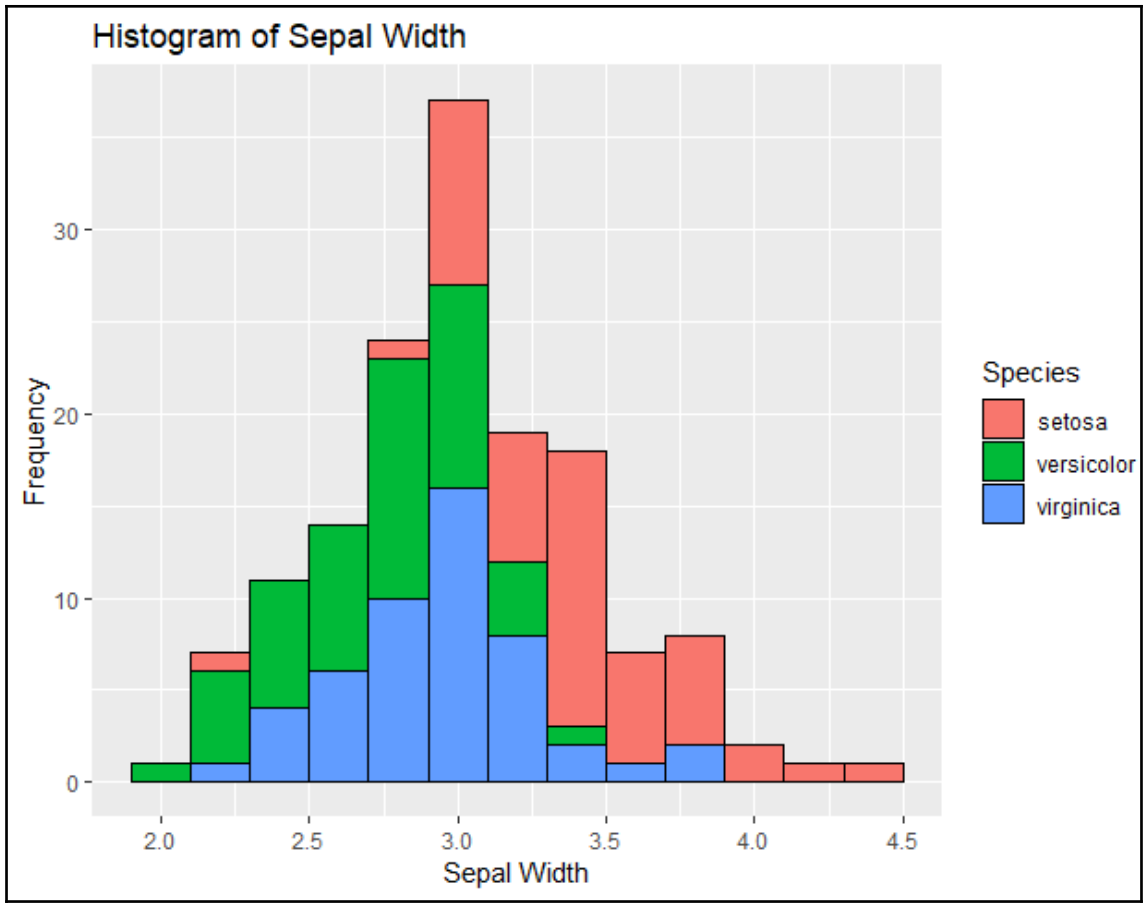
	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species
1	1	5.1	3.5	1.4	0.2	Iris-setosa
2	2	4.9	3.0	1.4	0.2	Iris-setosa
3	3	4.7	3.2	1.3	0.2	Iris-setosa
4	4	4.6	3.1	1.5	0.2	Iris-setosa
5	5	5.0	3.6	1.4	0.2	Iris-setosa
6	6	5.4	3.9	1.7	0.4	Iris-setosa
7	7	4.6	3.4	1.4	0.3	Iris-setosa
8	8	5.0	3.4	1.5	0.2	Iris-setosa
9	9	4.4	2.9	1.4	0.2	Iris-setosa
10	10	4.9	3.1	1.5	0.1	Iris-setosa
11	11	5.4	3.7	1.5	0.2	Iris-setosa
12	12	4.8	3.4	1.6	0.2	Iris-setosa

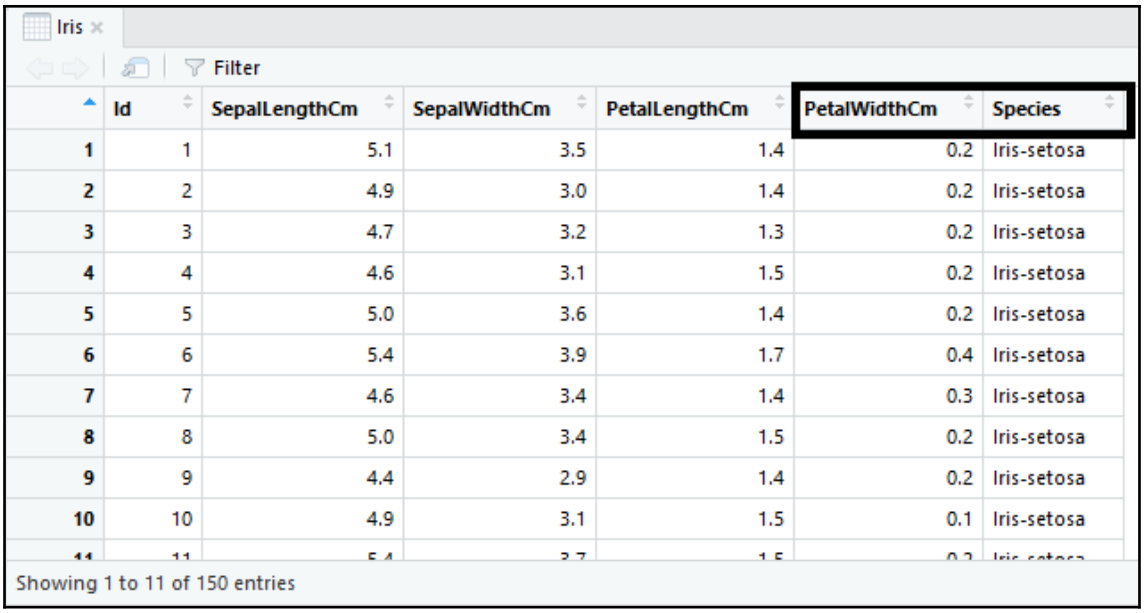
Showing 1 to 12 of 150 entries


8
7
6
5
SepalLengthCm
2.0
2.5
3.0
3.5
4.0
4.5
SepalWidthCm


8
7
6
5
SepalLengthCm
2.0
2.5
3.0
3.5
4.0
4.5
SepalWidthCm
Species
Iris-setosa
Iris-versicolor
Iris-virginica


Histogram of Sepal Width
Frequency
0
5
10
15
20
25
30
35
2.0
2.5
3.0
3.5
4.0
Sepal Width


Histogram of Sepal Width
Frequency
30
20
10
0
2.0
2.5
3.0
3.5
4.0
4.5
Sepal Width
Species
setosa
versicolor
virginica

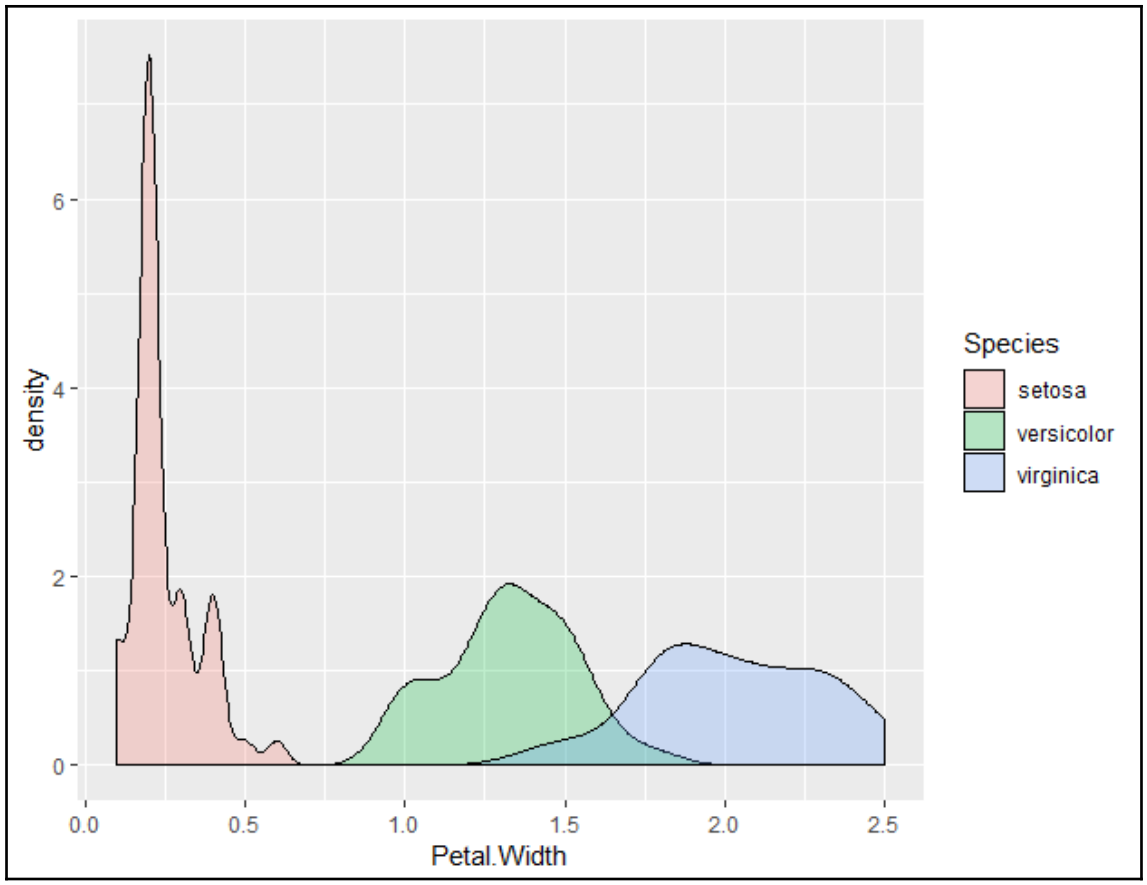


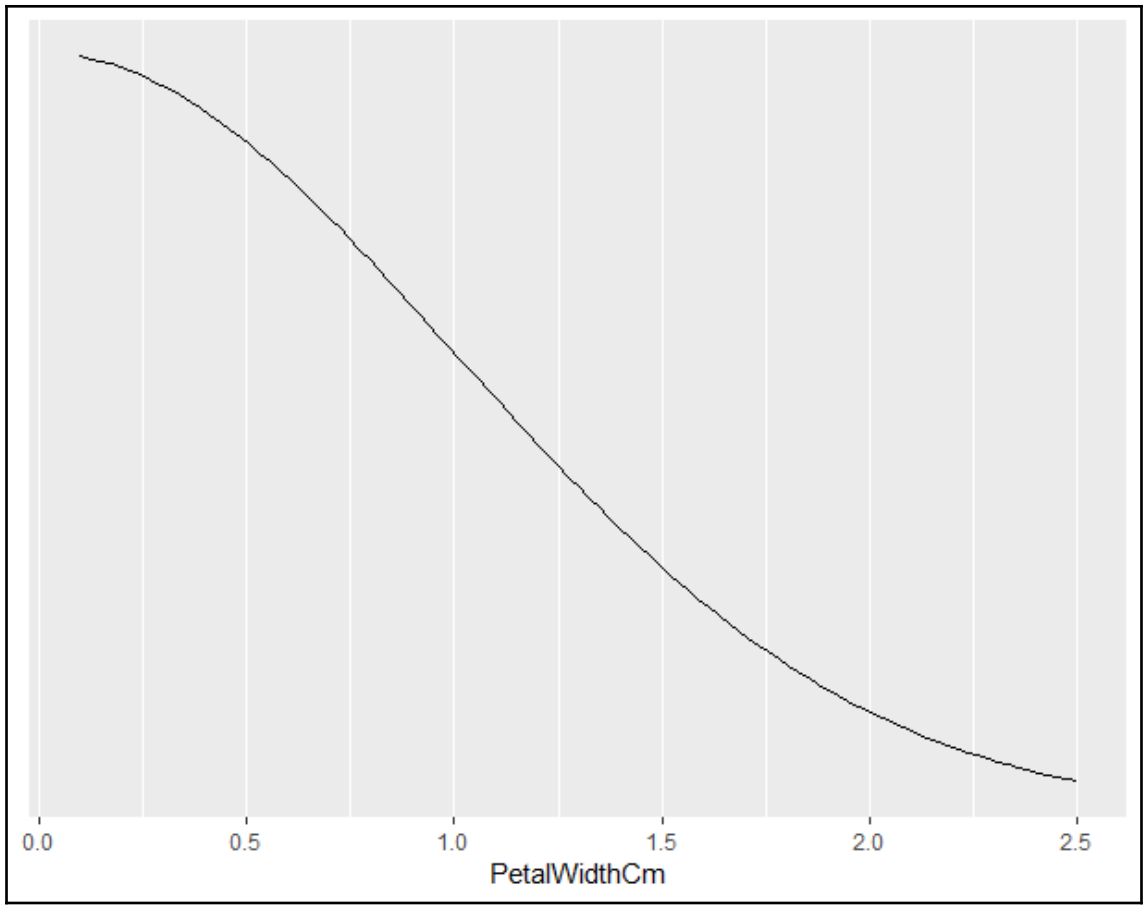
Iris

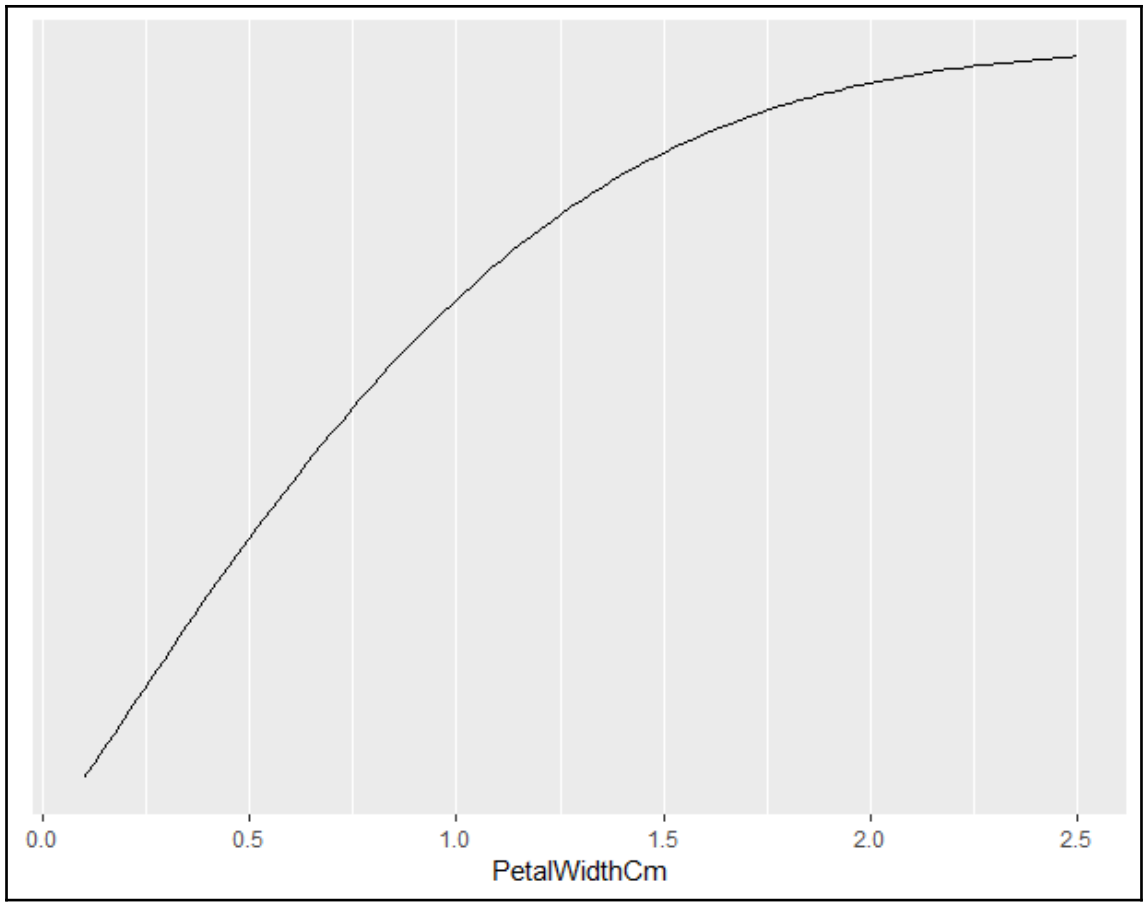
Filter

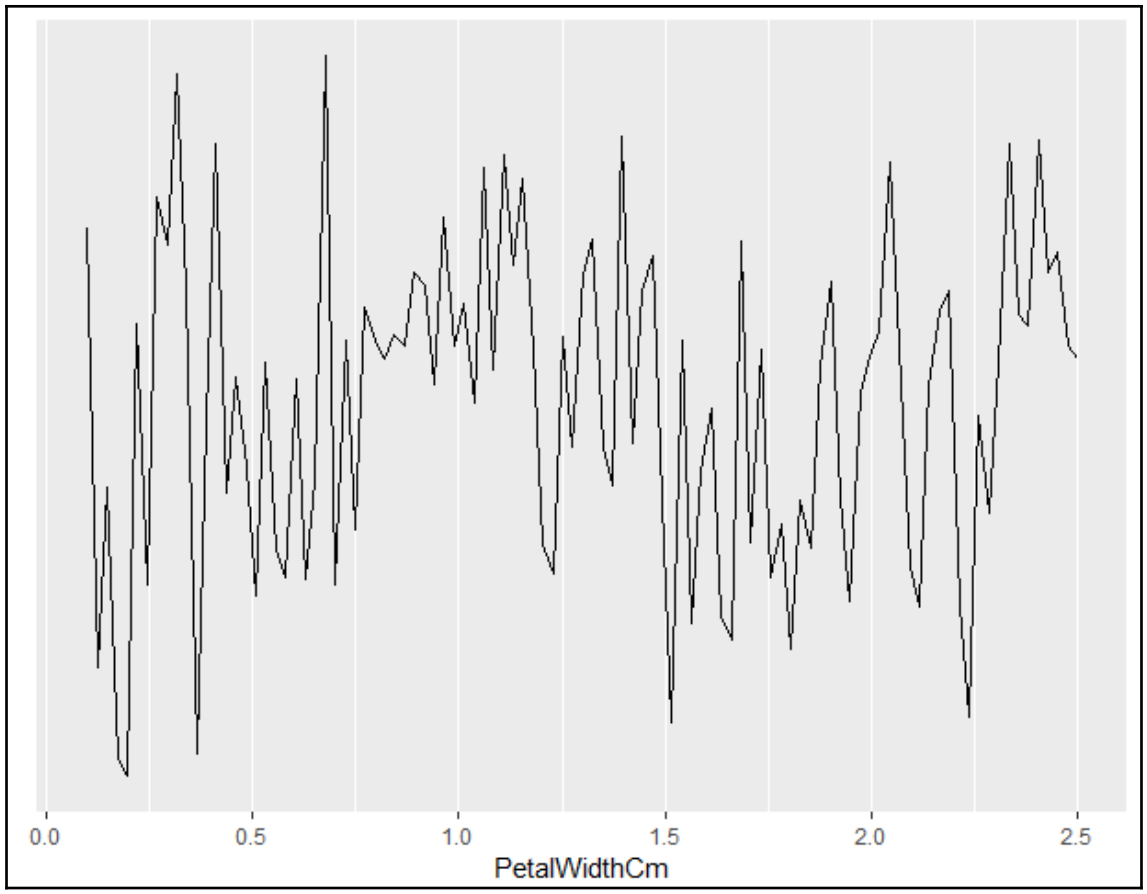
	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species
1	1	5.1	3.5	1.4	0.2	Iris-setosa
2	2	4.9	3.0	1.4	0.2	Iris-setosa
3	3	4.7	3.2	1.3	0.2	Iris-setosa
4	4	4.6	3.1	1.5	0.2	Iris-setosa
5	5	5.0	3.6	1.4	0.2	Iris-setosa
6	6	5.4	3.9	1.7	0.4	Iris-setosa
7	7	4.6	3.4	1.4	0.3	Iris-setosa
8	8	5.0	3.4	1.5	0.2	Iris-setosa
9	9	4.4	2.9	1.4	0.2	Iris-setosa
10	10	4.9	3.1	1.5	0.1	Iris-setosa
11	11	5.4	3.7	1.5	0.2	Iris-setosa

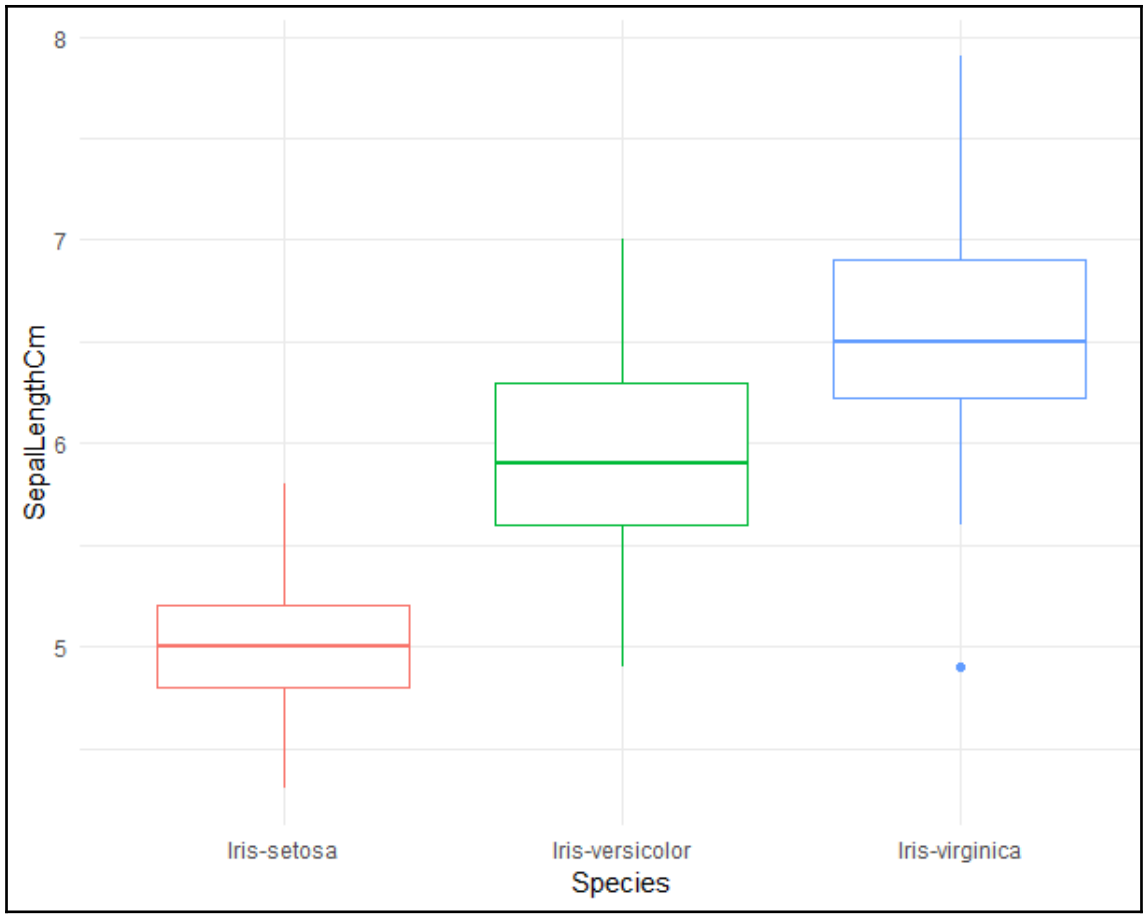
Showing 1 to 11 of 150 entries

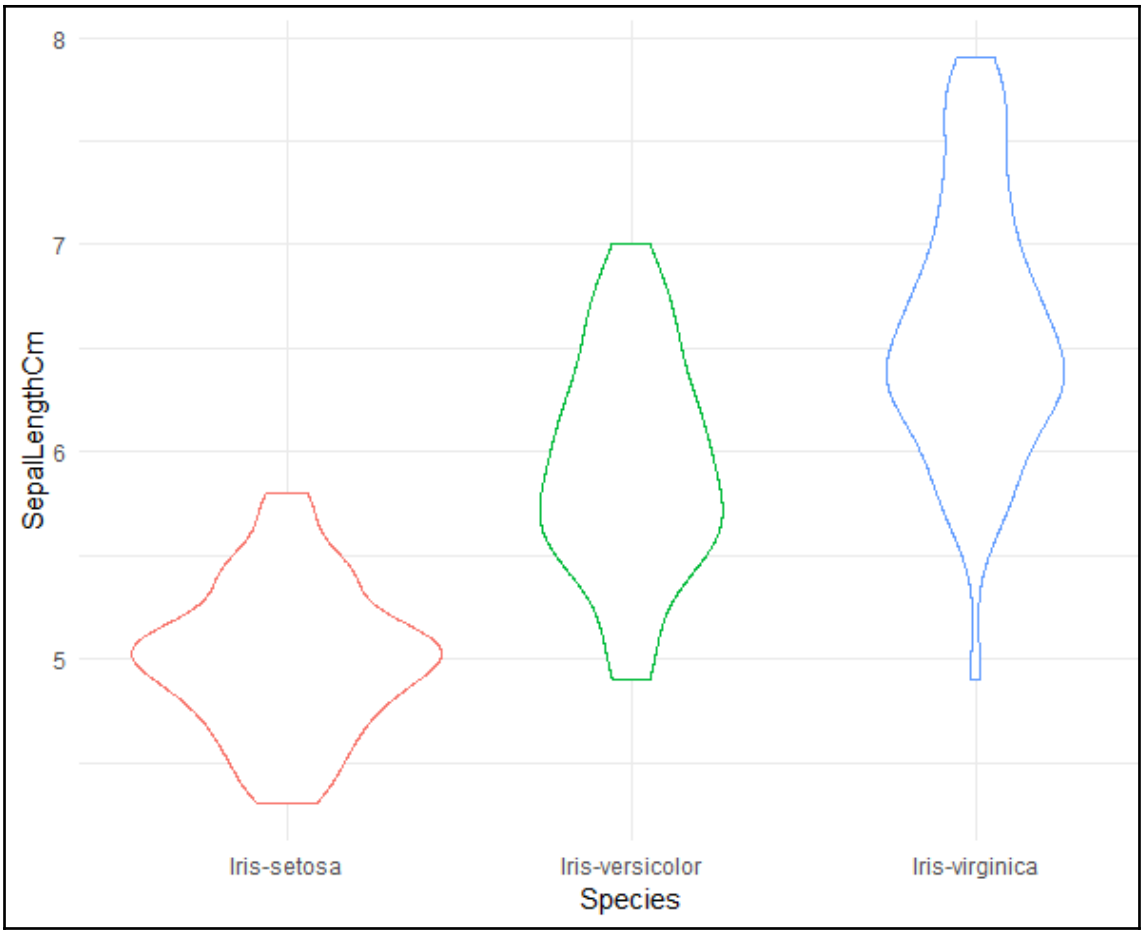

density
6
4
2
0
0.0
0.5
1.0
1.5
2.0
2.5
Petal.Width
Species
setosa
versicolor
virginica

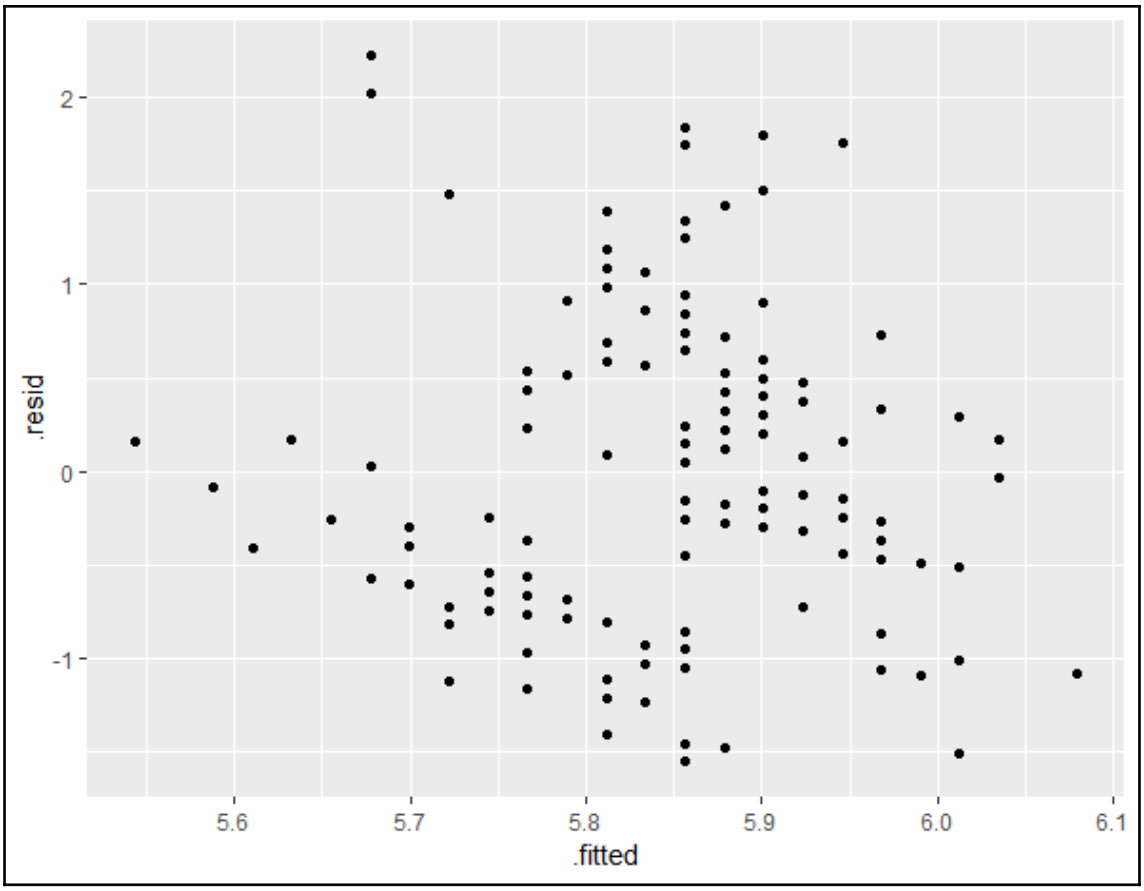

0.0
0.5
1.0
1.5
2.0
2.5
PetalWidthCm


0.0
0.5
1.0
1.5
2.0
2.5
PetalWidthCm


0.0
0.5
1.0
1.5
2.0
2.5
PetalWidthCm


8
7
6
5
SepalLengthCm
Iris-setosa
Iris-versicolor
Iris-virginica
Species


8
7
6
5
SepalLengthCm
Iris-setosa
Iris-versicolor
Iris-virginica
Species


2
1
0
-1
.resid
5.6
5.7
5.8
5.9
6.0
6.1
.fitted

Chapter 5: Creating Aesthetically Pleasing Reports with knitr and R Markdown

```
> install.packages("rmarkdown")
Installing package into 'C:/Users/Radhika/Documents/R/win-library/3.5'
(as 'lib' is unspecified)
also installing the dependencies 'evaluate', 'tinytex'

trying URL 'https://cran.rstudio.com/bin/windows/contrib/3.5/evaluate_0.13.zip'
Content type 'application/zip' length 73984 bytes (72 KB)
downloaded 72 KB

trying URL 'https://cran.rstudio.com/bin/windows/contrib/3.5/tinytex_0.12.zip'
Content type 'application/zip' length 95068 bytes (92 KB)
downloaded 92 KB

trying URL 'https://cran.rstudio.com/bin/windows/contrib/3.5/rmarkdown_1.12.zip'
Content type 'application/zip' length 3564083 bytes (3.4 MB)
downloaded 3.4 MB

package 'evaluate' successfully unpacked and MD5 sums checked
package 'tinytex' successfully unpacked and MD5 sums checked
package 'rmarkdown' successfully unpacked and MD5 sums checked

The downloaded binary packages are in
        C:\Users\Radhika\AppData\Local\Temp\RtmpWQm2mB\downloaded_packages
```

```
> install.packages("tinytex")
Installing package into 'C:/Users/Radhika/Documents/R/win-library/3.5'
(as 'lib' is unspecified)
trying URL 'https://cran.rstudio.com/bin/windows/contrib/3.5/tinytex_0.12.zip'
Content type 'application/zip' length 94822 bytes (92 KB)
downloaded 92 KB

package 'tinytex' successfully unpacked and MD5 sums checked

The downloaded binary packages are in
        C:\Users\Radhika\AppData\Local\Temp\RtmpWQm2mB\downloaded_packages
>
```

```
> install.packages("knitr")
Installing package into 'C:/Users/Radhika/Documents/R/win-library/3.5'
(as 'lib' is unspecified)
also installing the dependency 'xfun'

trying URL 'https://cran.rstudio.com/bin/windows/contrib/3.5/xfun_0.6.zip'
Content type 'application/zip' length 166880 bytes (162 KB)
downloaded 162 KB

trying URL 'https://cran.rstudio.com/bin/windows/contrib/3.5/knitr_1.22.zip'
Content type 'application/zip' length 1470245 bytes (1.4 MB)
downloaded 1.4 MB

package 'xfun' successfully unpacked and MD5 sums checked
package 'knitr' successfully unpacked and MD5 sums checked

The downloaded binary packages are in
        C:\Users\Radhika\AppData\Local\Temp\RtmpWQm2mB\downloaded_packages
>
```

```
---
title: "Introduction"
output: html_document
---

```{r setup, include=FALSE}
knitr::opts_chunk$set(echo = TRUE)
```

R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for
authoring HTML, PDF, and MS Word documents. For more details on using R
Markdown see <http://rmarkdown.rstudio.com>.

When you click the **Knit** button a document will be generated that includes
both content as well as the output of any embedded R code chunks within the
document. You can embed an R code chunk like this:

```{r cars}
```

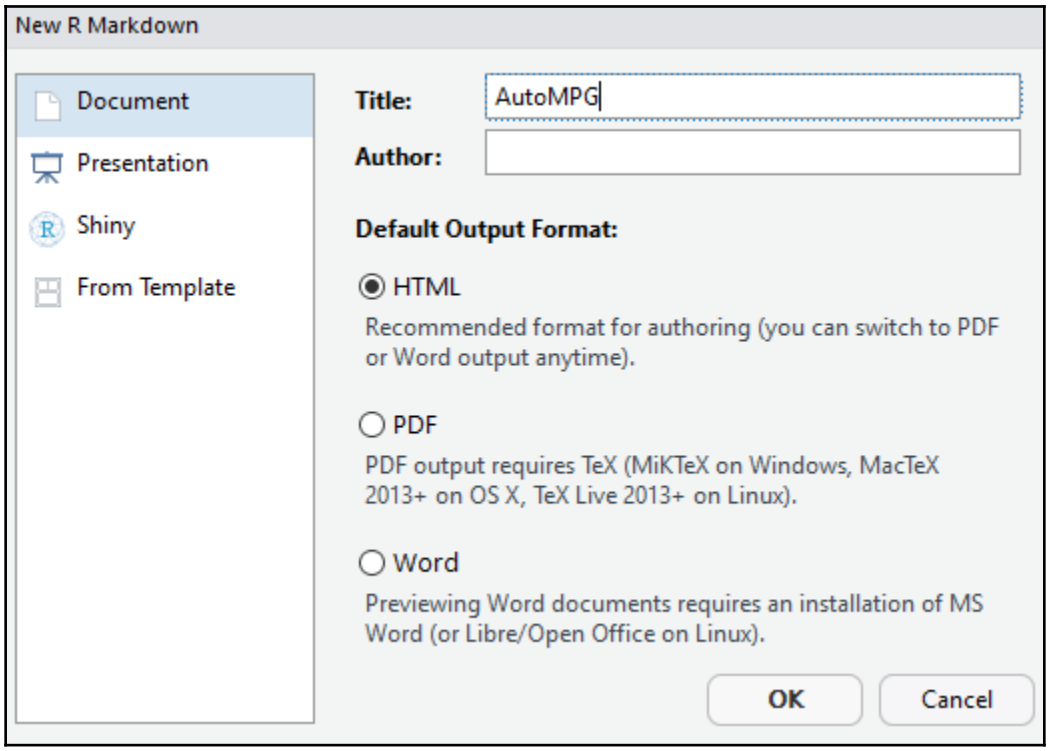
Introduction   R Markdown
```

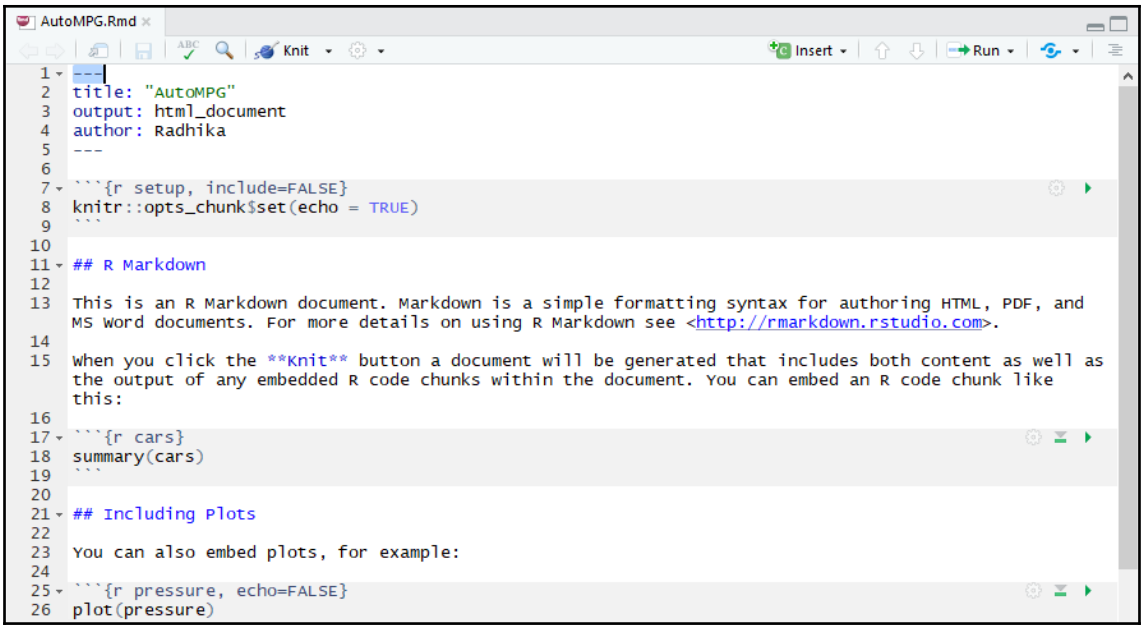
Autompg

Filter

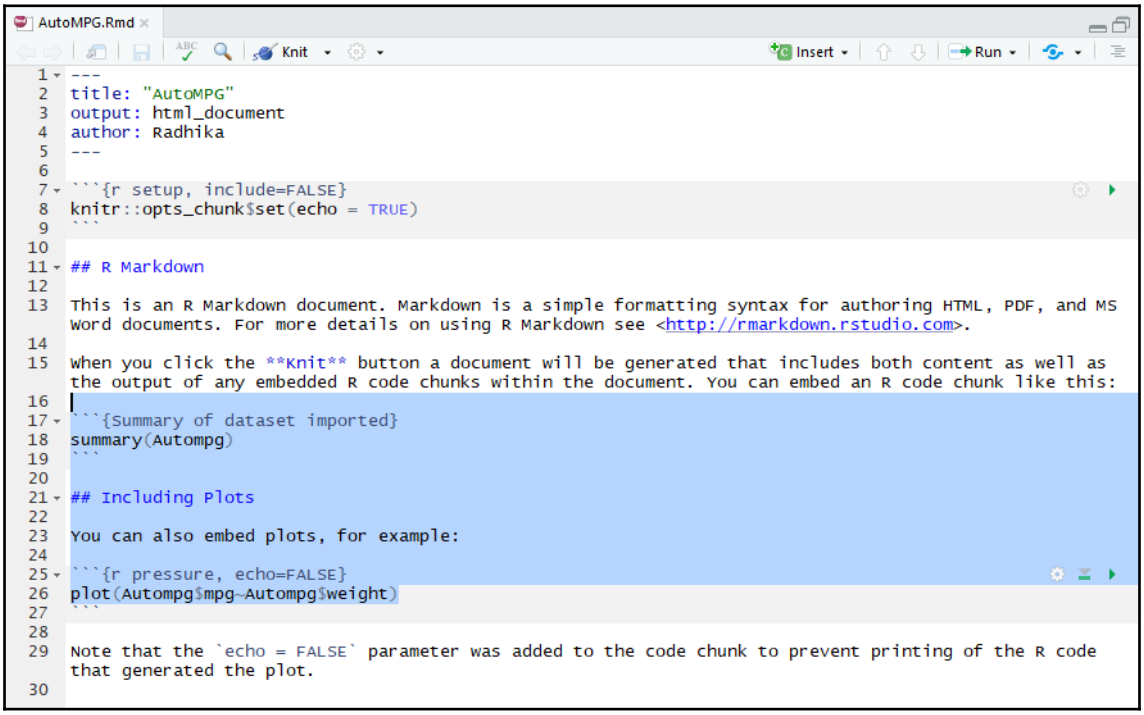
	mpg	cylinders	displacement	horsepower	weight	acceleration	model.year	origin	car.name
1	18	8	307.0	130	3504	12.0	70	1	chevrolet chevelle malibu
2	15	8	350.0	165	3693	11.5	70	1	buick skylark 320
3	18	8	318.0	150	3436	11.0	70	1	plymouth satellite
4	16	8	304.0	150	3433	12.0	70	1	amc rebel sst
5	17	8	302.0	140	3449	10.5	70	1	ford torino
6	15	8	429.0	198	4341	10.0	70	1	ford galaxie 500
7	14	8	454.0	220	4354	9.0	70	1	chevrolet impala
8	14	8	440.0	215	4312	8.5	70	1	plymouth fury iii
9	14	8	455.0	225	4425	10.0	70	1	pontiac catalina
10	15	8	390.0	190	3850	8.5	70	1	amc ambassador dpl
11	15	8	383.0	170	3563	10.0	70	1	dodge challenger se
12	14	8	340.0	160	3609	8.0	70	1	plymouth 'cuda 340

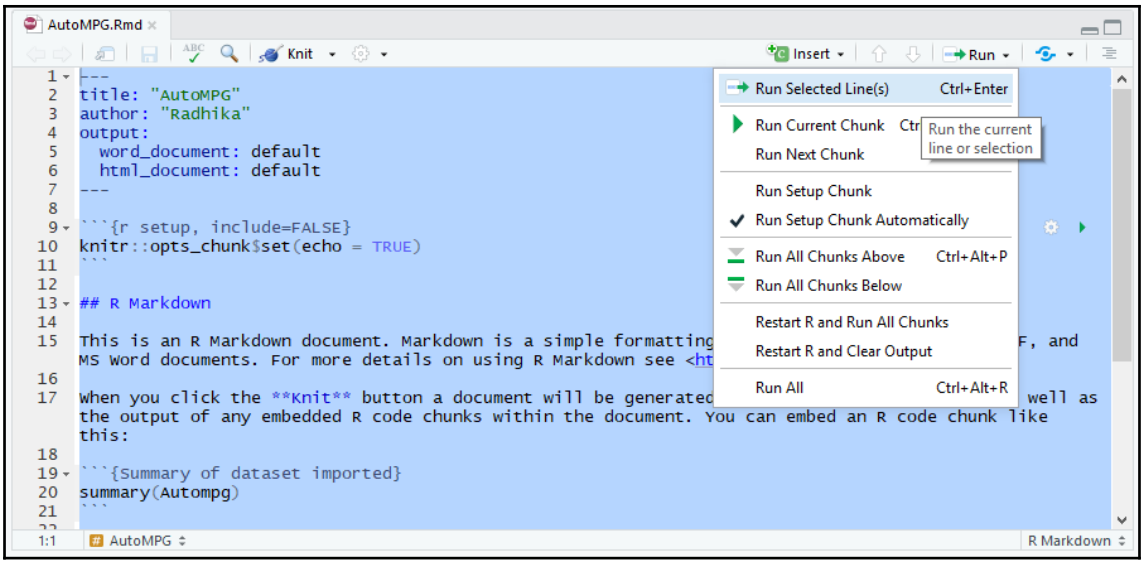
Showing 1 to 12 of 398 entries



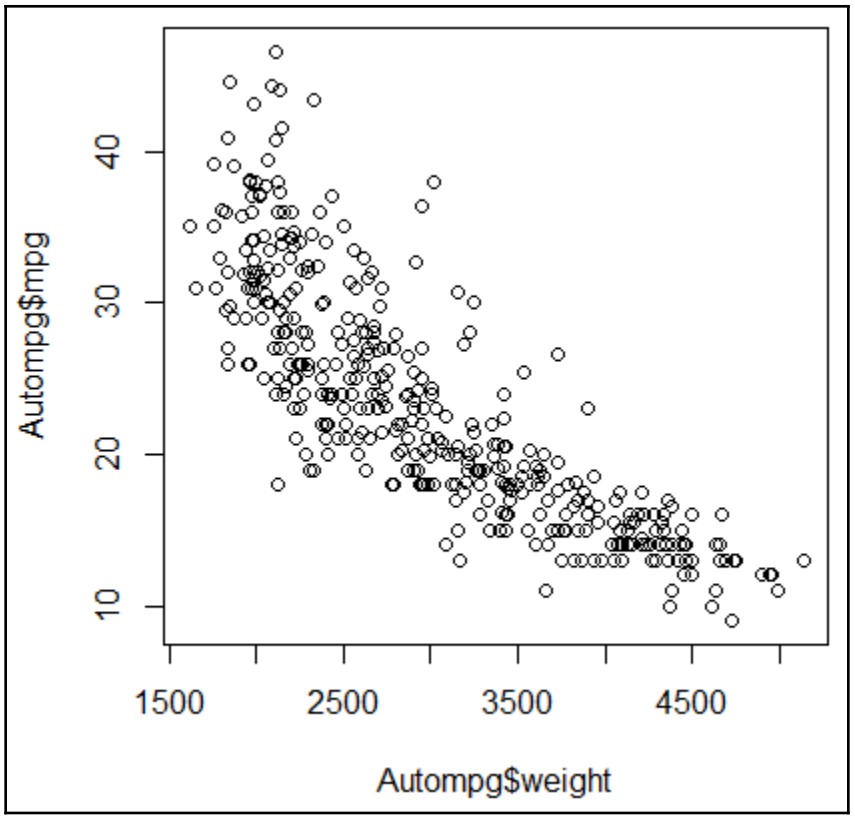

AutoMPG.Rmd
Knit
Insert
Run
---
title: "AutoMPG"
output: html_document
author: Radhika

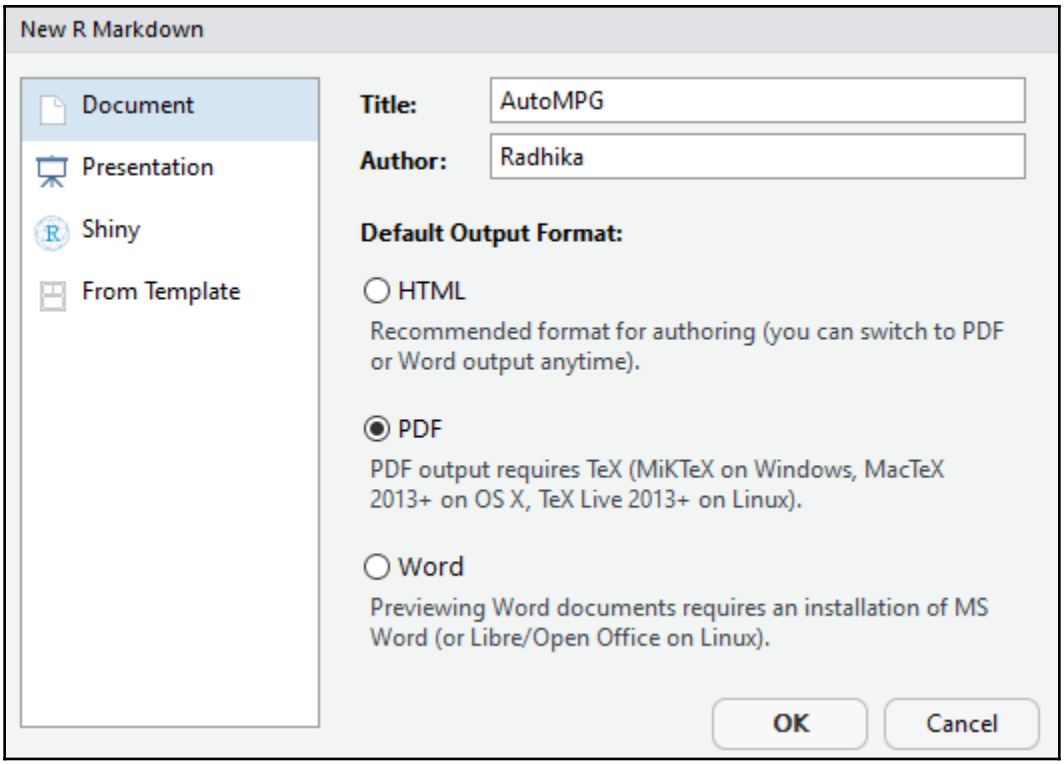
```{r setup, include=FALSE}
knitr::opts_chunk$set(echo = TRUE)
```
R Markdown
This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.
When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:
```{r cars}
summary(cars)
```
Including Plots
You can also embed plots, for example:
```{r pressure, echo=FALSE}
plot(pressure)


AutoMPG.Rmd
Knit
Insert
Run
---
title: "AutoMPG"
output: html_document
author: Radhika
---
```{r setup, include=FALSE}
knitr::opts_chunk$set(echo = TRUE)
```
R Markdown
This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.
When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:
```{Summary of dataset imported}
summary(Autompg)
```
Including Plots
You can also embed plots, for example:
```{r pressure, echo=FALSE}
plot(Autompg$mpg~Autompg$weight)
```
Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.
```


AutoMPG.Rmd
Knit
Insert
Run
Run Selected Line(s) Ctrl+Enter
Run Current Chunk
Run Next Chunk
Run the current line or selection
Run Setup Chunk
Run Setup Chunk Automatically
Run All Chunks Above Ctrl+Alt+P
Run All Chunks Below
Restart R and Run All Chunks
Restart R and Clear Output
Run All Ctrl+Alt+R
---
title: "AutoMPG"
author: "Radhika"
output:
word_document: default
html_document: default

```{r setup, include=FALSE}
knitr::opts_chunk$set(echo = TRUE)
```
R Markdown
This is an R Markdown document. Markdown is a simple formatting
MS Word documents. For more details on using R Markdown see <ht
when you click the **Knit** button a document will be generated
the output of any embedded R code chunks within the document. You can embed an R code chunk like
this:
```{Summary of dataset imported}
summary(Autompg)
```
1:1
AutoMPG
R Markdown


Autompg$mpg
10
20
30
40
1500
2500
3500
4500
Autompg$weight



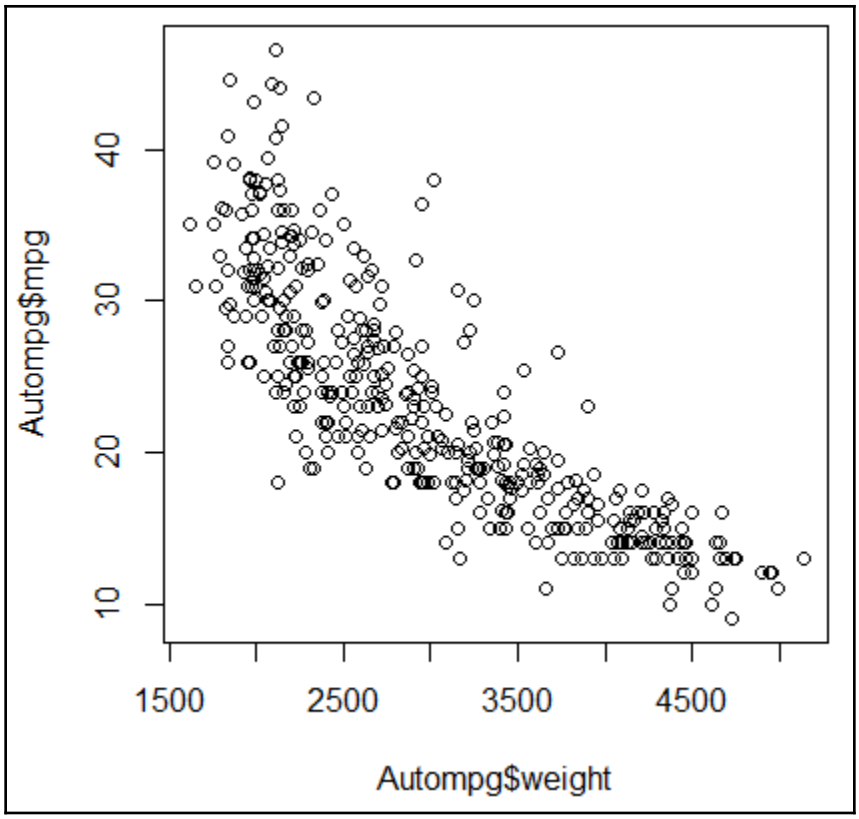

```
---
title: "AutoMPG"
author: "Radhika"
date: "May 7, 2019"
output: pdf_document
---

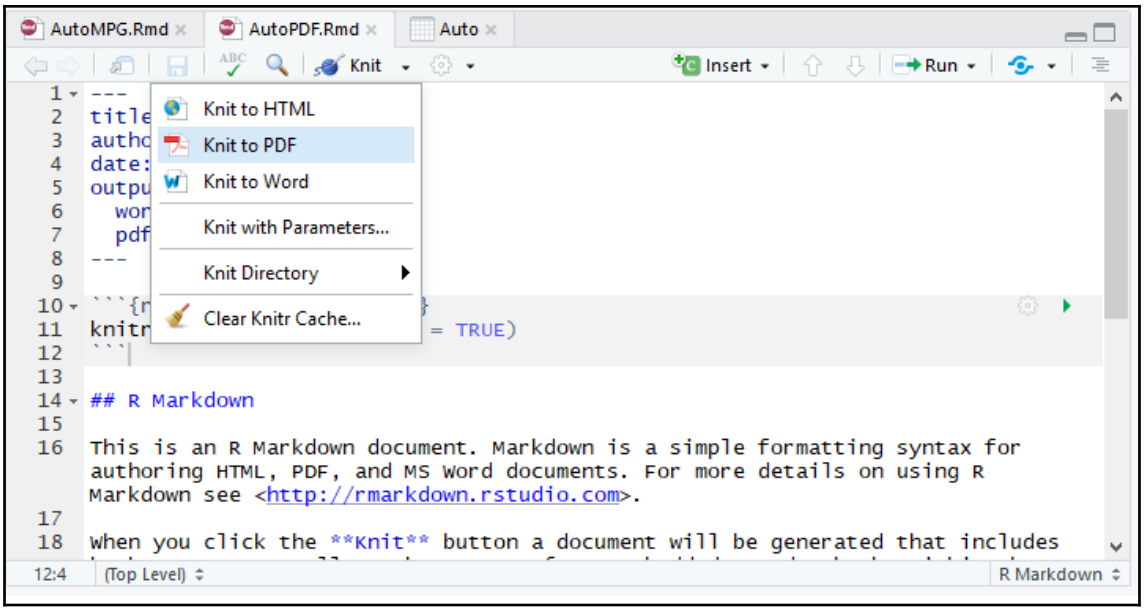
```{r setup, include=FALSE}
knitr::opts_chunk$set(echo = TRUE)
```

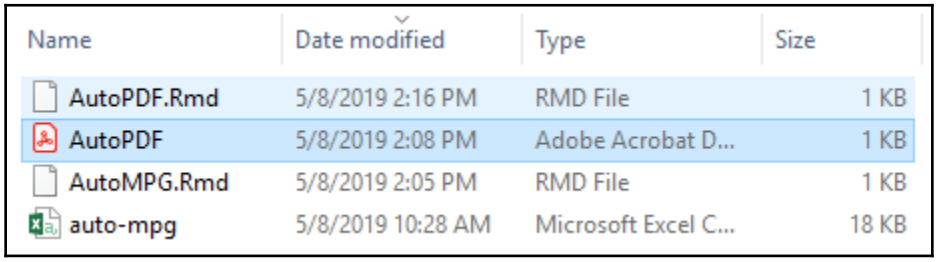
R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for
authoring HTML, PDF, and MS Word documents. For more details on using R Markdown
see <http://rmarkdown.rstudio.com>.

When you click the **Knit** button a document will be generated that includes both
content as well as the output of any embedded R code chunks within the document.
You can embed an R code chunk like this:
```


Autompg$mpg
10
20
30
40
1500
2500
3500
4500
Autompg$weight


AutoMPG.Rmd
AutoPDF.Rmd
Auto
Knit
Insert
Run
Knit to HTML
Knit to PDF
Knit to Word
Knit with Parameters...
Knit Directory
Clear Knitr Cache...
R Markdown
This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.
When you click the **Knit** button a document will be generated that includes
12:4
(Top Level)
R Markdown

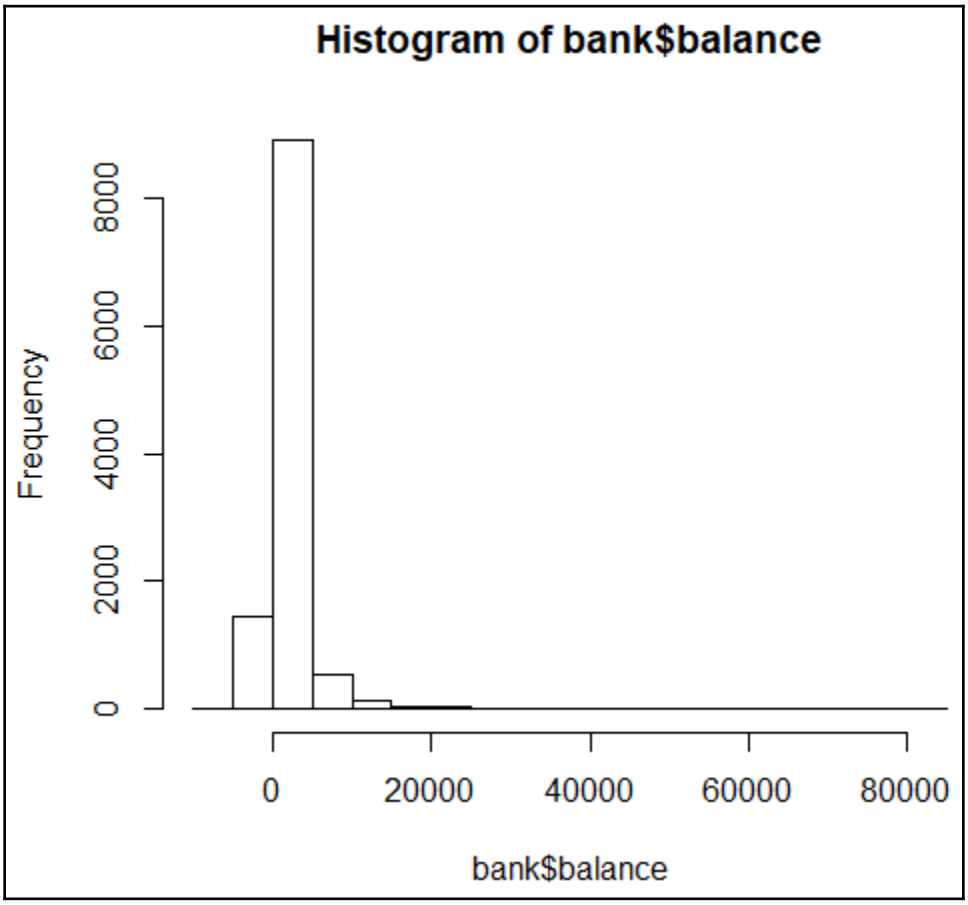

Name
Date modified
Type
Size
AutoPDF.Rmd
5/8/2019 2:16 PM
RMD File
1 KB
AutoPDF
5/8/2019 2:08 PM
Adobe Acrobat D...
1 KB
AutoMPG.Rmd
5/8/2019 2:05 PM
RMD File
1 KB
auto-mpg
5/8/2019 10:28 AM
Microsoft Excel C...
18 KB

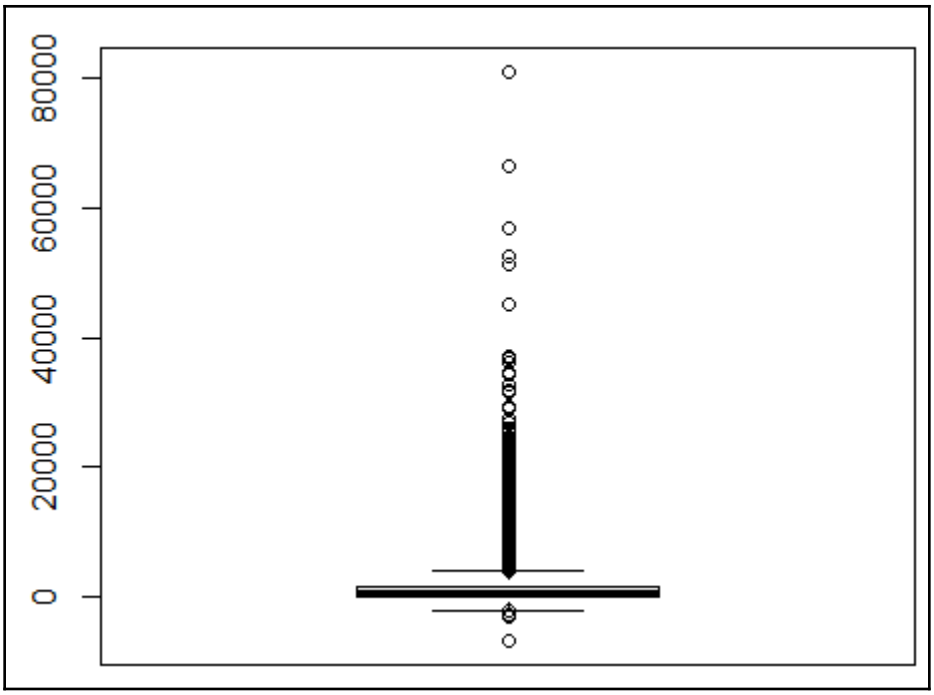
Chapter 6: Univariate and Control Datasets

bank

Filter

	age	job	marital	education	default	balance	housing	loan	contact	day	month	duration	campaign	pdays	previous	poutcome	deposit
1	59	admin.	married	secondary	no	2343	yes	no	unknown	5	may	1042	1	-1	0	unknown	yes
2	56	admin.	married	secondary	no	45	no	no	unknown	5	may	1467	1	-1	0	unknown	yes
3	41	technician	married	secondary	no	1270	yes	no	unknown	5	may	1389	1	-1	0	unknown	yes
4	55	services	married	secondary	no	2476	yes	no	unknown	5	may	579	1	-1	0	unknown	yes
5	54	admin.	married	tertiary	no	184	no	no	unknown	5	may	673	2	-1	0	unknown	yes
6	42	management	single	tertiary	no	0	yes	yes	unknown	5	may	562	2	-1	0	unknown	yes
7	56	management	married	tertiary	no	830	yes	yes	unknown	6	may	1201	1	-1	0	unknown	yes
8	60	retired	divorced	secondary	no	545	yes	no	unknown	6	may	1030	1	-1	0	unknown	yes
9	37	technician	married	secondary	no	1	yes	no	unknown	6	may	608	1	-1	0	unknown	yes
10	28	services	single	secondary	no	5090	yes	no	unknown	6	may	1297	3	-1	0	unknown	yes
11	38	admin.	single	secondary	no	100	yes	no	unknown	7	may	786	1	-1	0	unknown	yes

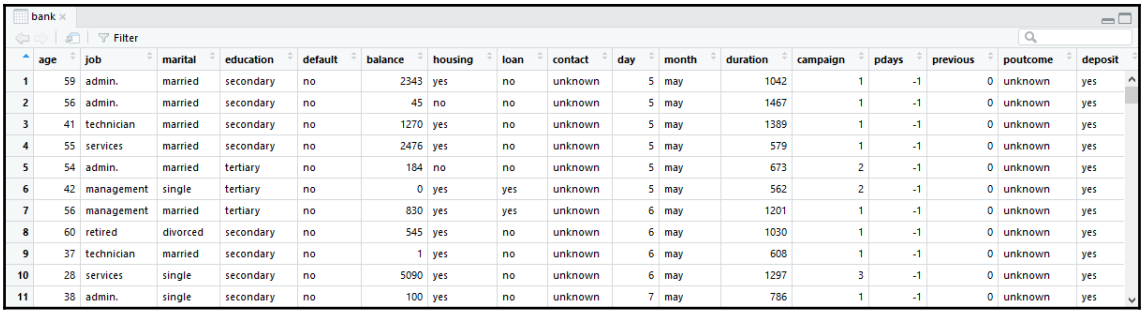




bank × bank1 ×

Filter

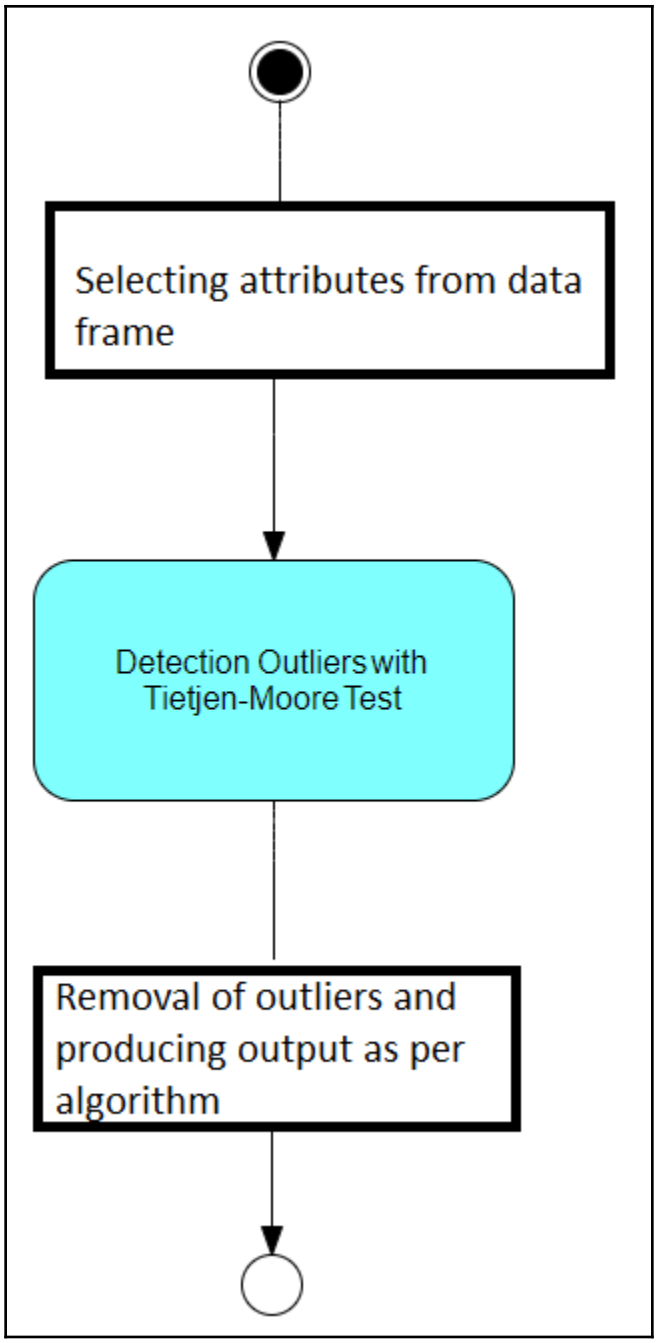
	age	job	marital	education	default	housing	contact	balance	day	month	duration
6	42	management	single	tertiary	no	yes	unknown	yes	5	may	562
7	56	management	married	tertiary	no	yes	unknown	yes	6	may	1201
13	29	management	married	tertiary	no	yes	unknown	yes	7	may	1689
20	49	admin.	divorced	secondary	no	yes	unknown	yes	8	may	513
41	60	blue-collar	married	primary	no	yes	unknown	yes	13	may	1015
52	39	management	divorced	tertiary	no	yes	unknown	yes	14	may	1328
53	59	retired	married	secondary	no	yes	unknown	yes	14	may	1125
67	49	unknown	married	primary	no	yes	unknown	yes	15	may	520
77	39	technician	married	tertiary	no	yes	unknown	yes	16	may	813
81	29	admin.	single	secondary	no	yes	unknown	yes	16	may	803

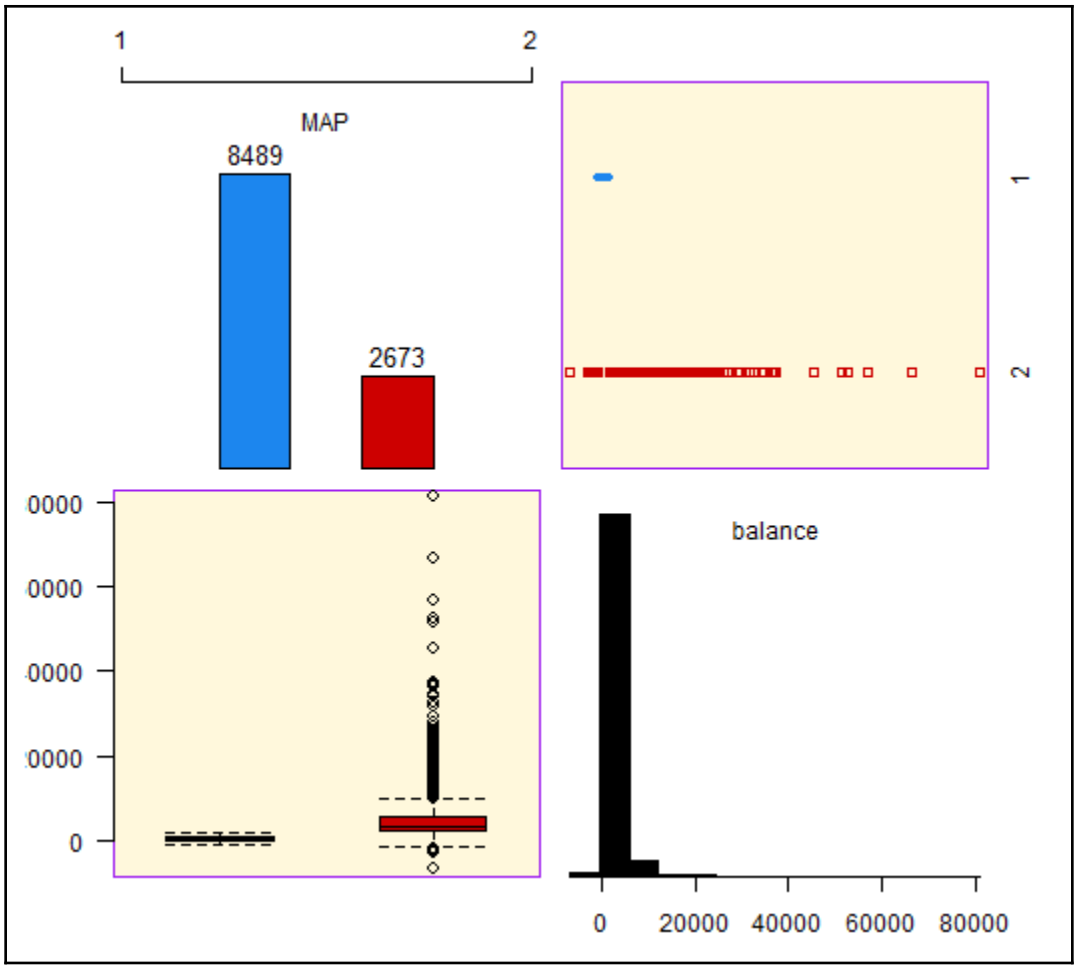


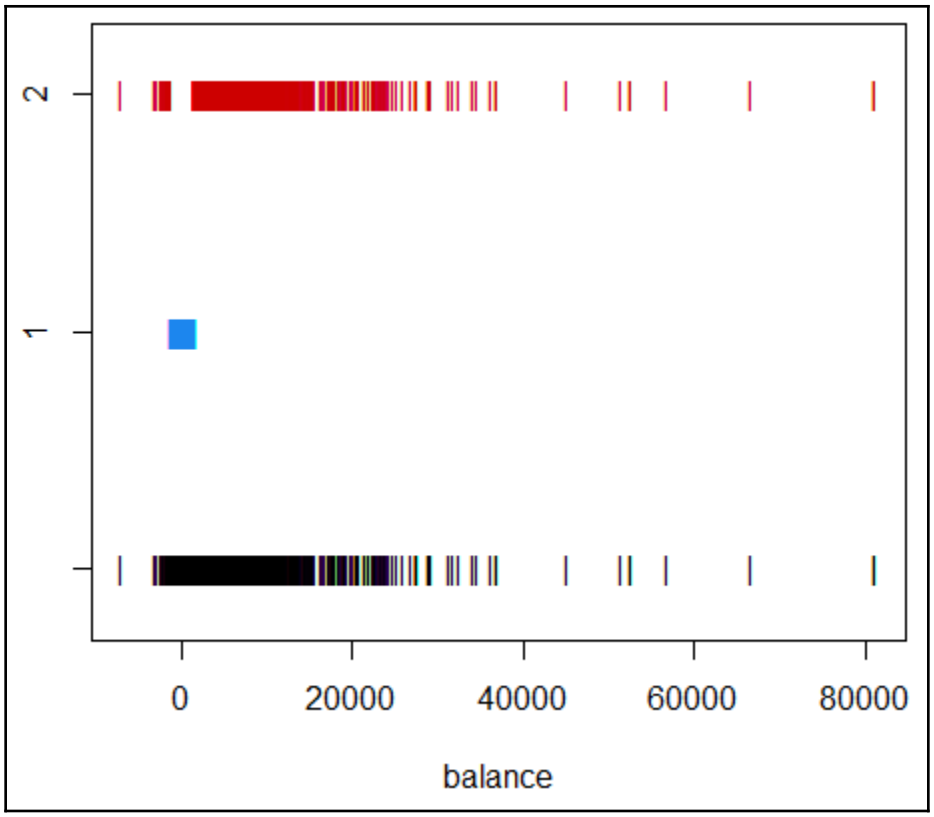
bank

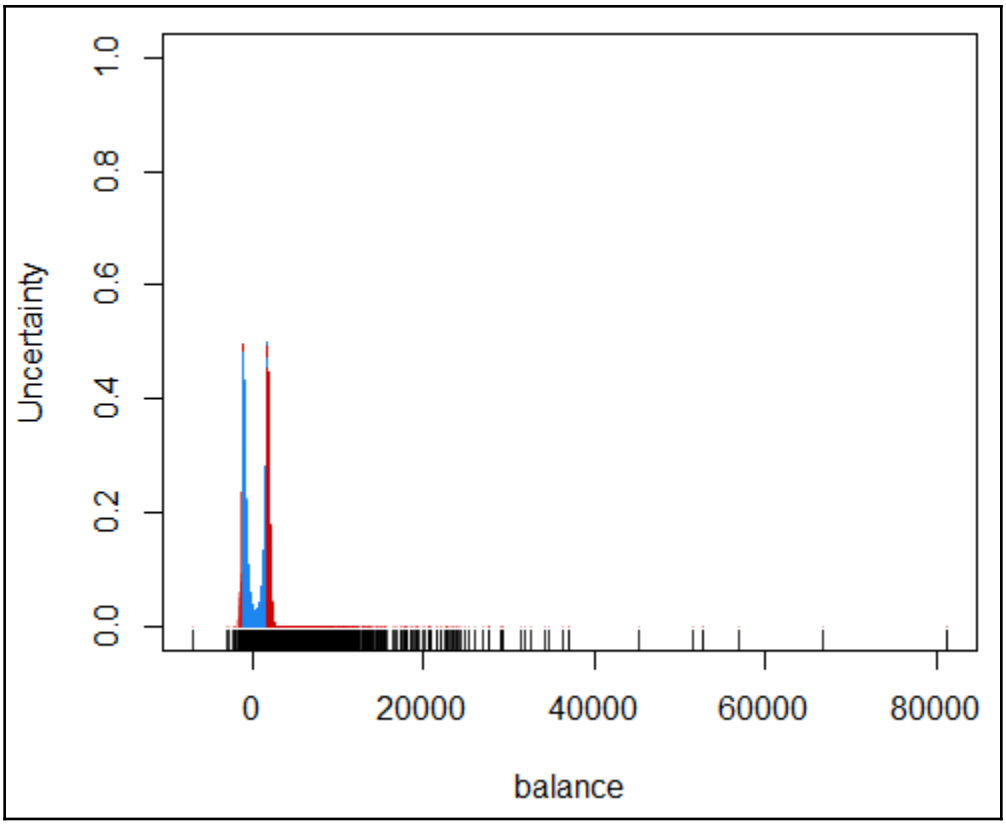
Filter

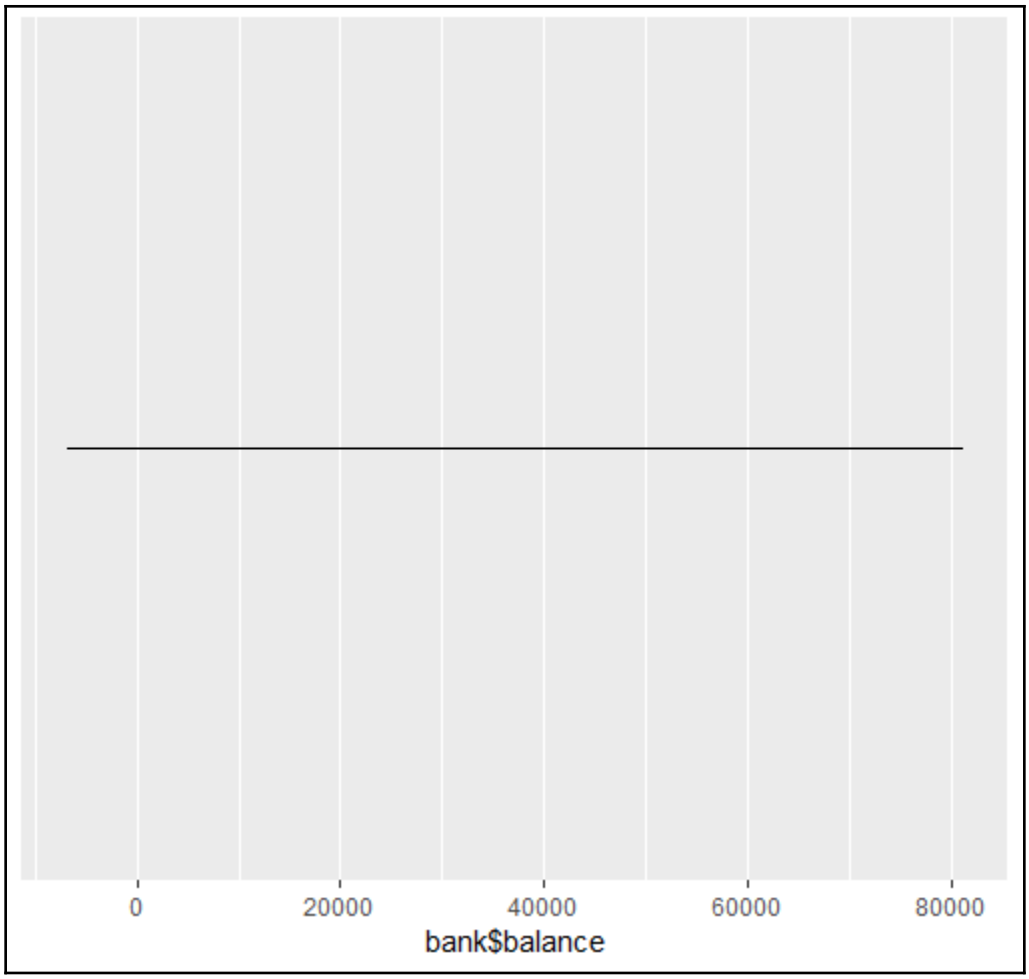
	age	job	marital	education	default	balance	housing	loan	contact	day	month	duration	campaign	pdays	previous	poutcome	deposit
1	59	admin.	married	secondary	no	2343	yes	no	unknown	5	may	1042	1	-1	0	unknown	yes
2	56	admin.	married	secondary	no	45	no	no	unknown	5	may	1467	1	-1	0	unknown	yes
3	41	technician	married	secondary	no	1270	yes	no	unknown	5	may	1389	1	-1	0	unknown	yes
4	55	services	married	secondary	no	2476	yes	no	unknown	5	may	579	1	-1	0	unknown	yes
5	54	admin.	married	tertiary	no	184	no	no	unknown	5	may	673	2	-1	0	unknown	yes
6	42	management	single	tertiary	no	0	yes	yes	unknown	5	may	562	2	-1	0	unknown	yes
7	56	management	married	tertiary	no	830	yes	yes	unknown	6	may	1201	1	-1	0	unknown	yes
8	60	retired	divorced	secondary	no	545	yes	no	unknown	6	may	1030	1	-1	0	unknown	yes
9	37	technician	married	secondary	no	1	yes	no	unknown	6	may	608	1	-1	0	unknown	yes
10	28	services	single	secondary	no	5090	yes	no	unknown	6	may	1297	3	-1	0	unknown	yes
11	38	admin.	single	secondary	no	100	yes	no	unknown	7	may	786	1	-1	0	unknown	yes

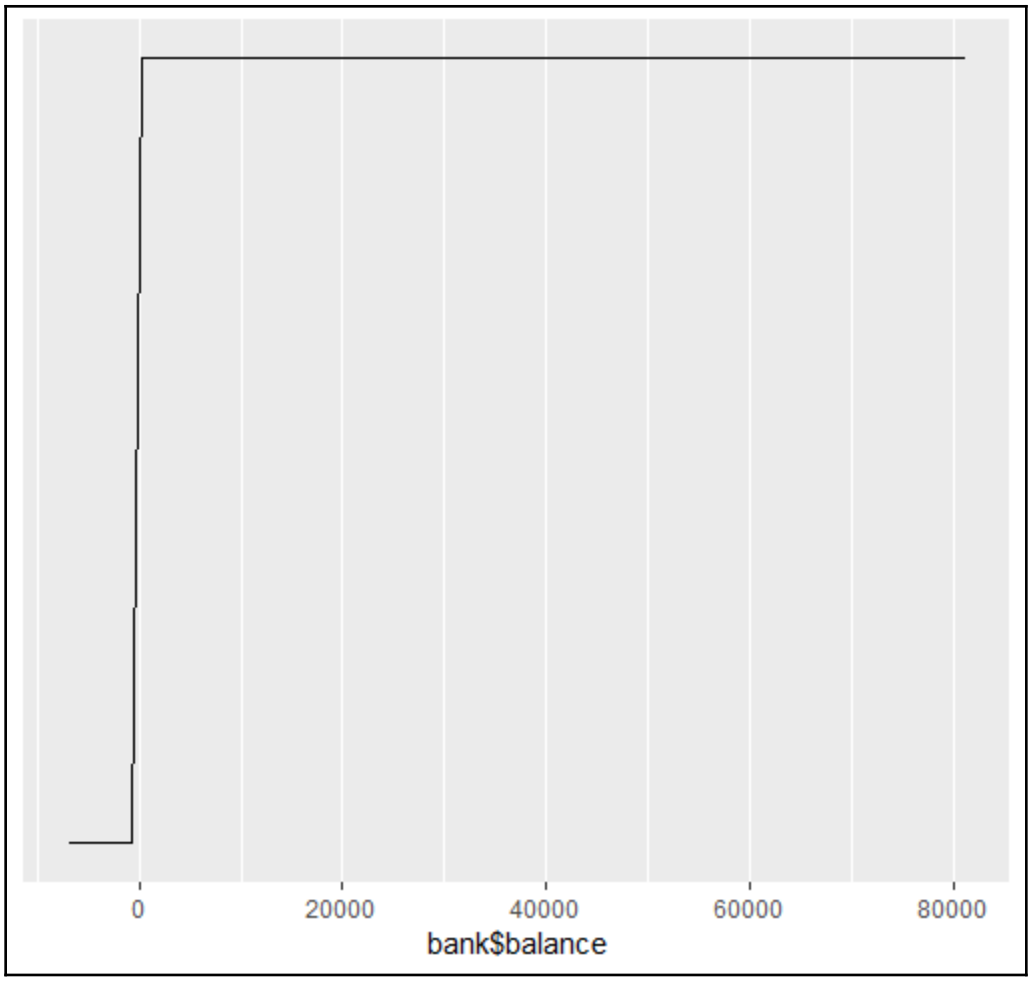

Selecting attributes from data frame
Detection Outliers with Tietjen-Moore Test
Removal of outliers and producing output as per algorithm

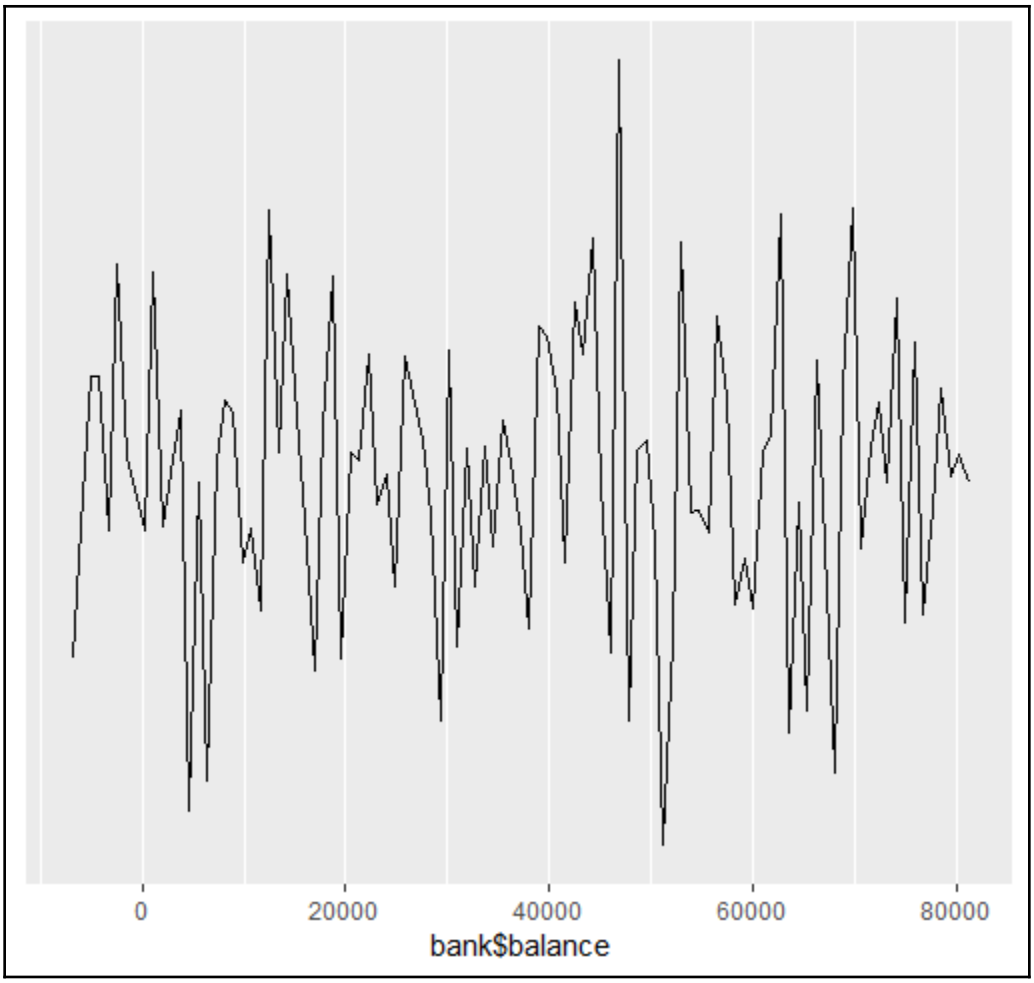

1
2
MAP
8489
2673
1
2
0000
0000
0000
0000
0
balance
0
20000
40000
60000
80000


2
1
0
20000
40000
60000
80000
balance


Uncertainty
0.0
0.2
0.4
0.6
0.8
1.0
0
20000
40000
60000
80000
balance


0
20000
40000
60000
80000
bank$balance


0
20000
40000
60000
80000
bank$balance


0
20000
40000
60000
80000
bank$balance

shapiro.test {stats} R Documentation

Shapiro-Wilk Normality Test

Description

Performs the Shapiro-Wilk test of normality.

Usage

```
shapiro.test(x)
```

Arguments

`x`	a numeric vector of data values. Missing values are allowed, but the number of non-missing values must be between 3 and 5000.

Value

A list with class `"htest"` containing the following components:

`statistic`	the value of the Shapiro-Wilk statistic.
`p.value`	an approximate p-value for the test. This is said in Royston (1995) to be adequate

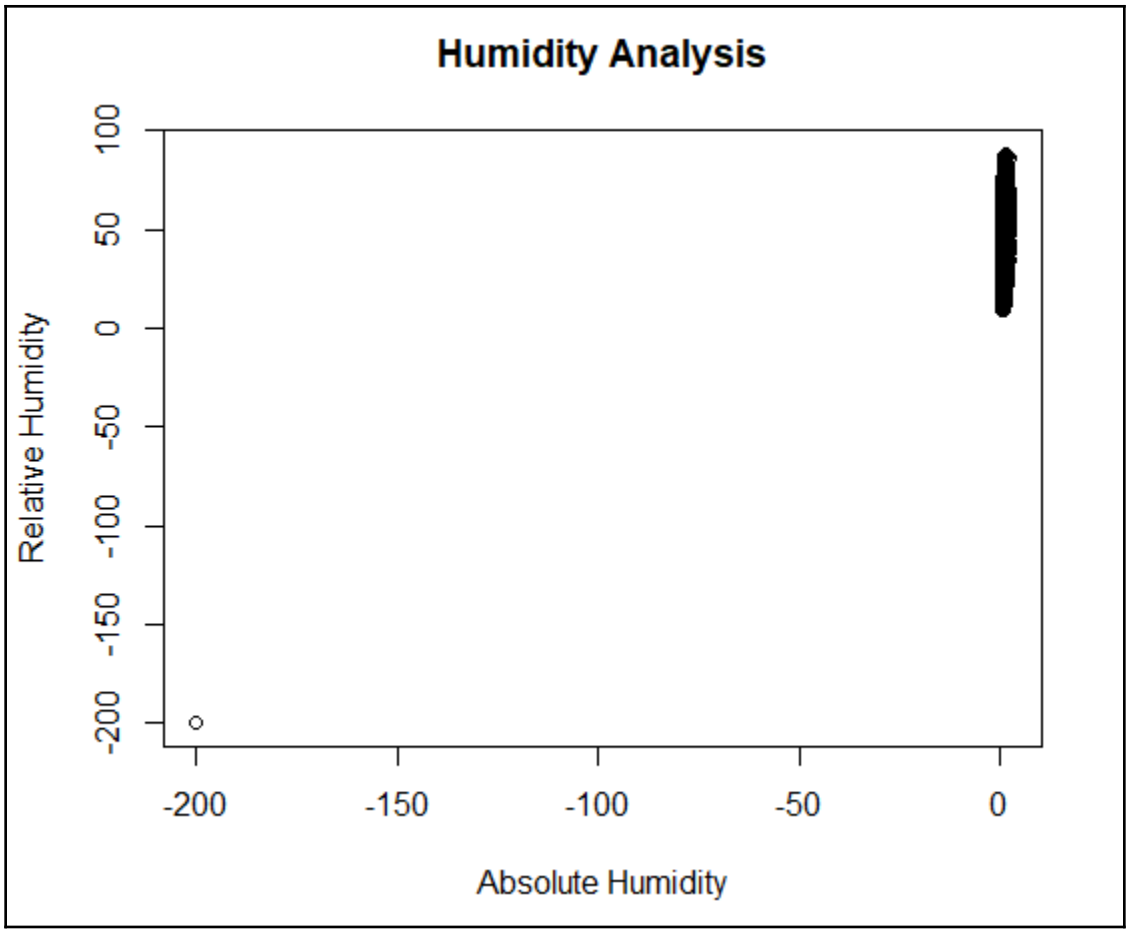
Chapter 7: Time Series Datasets

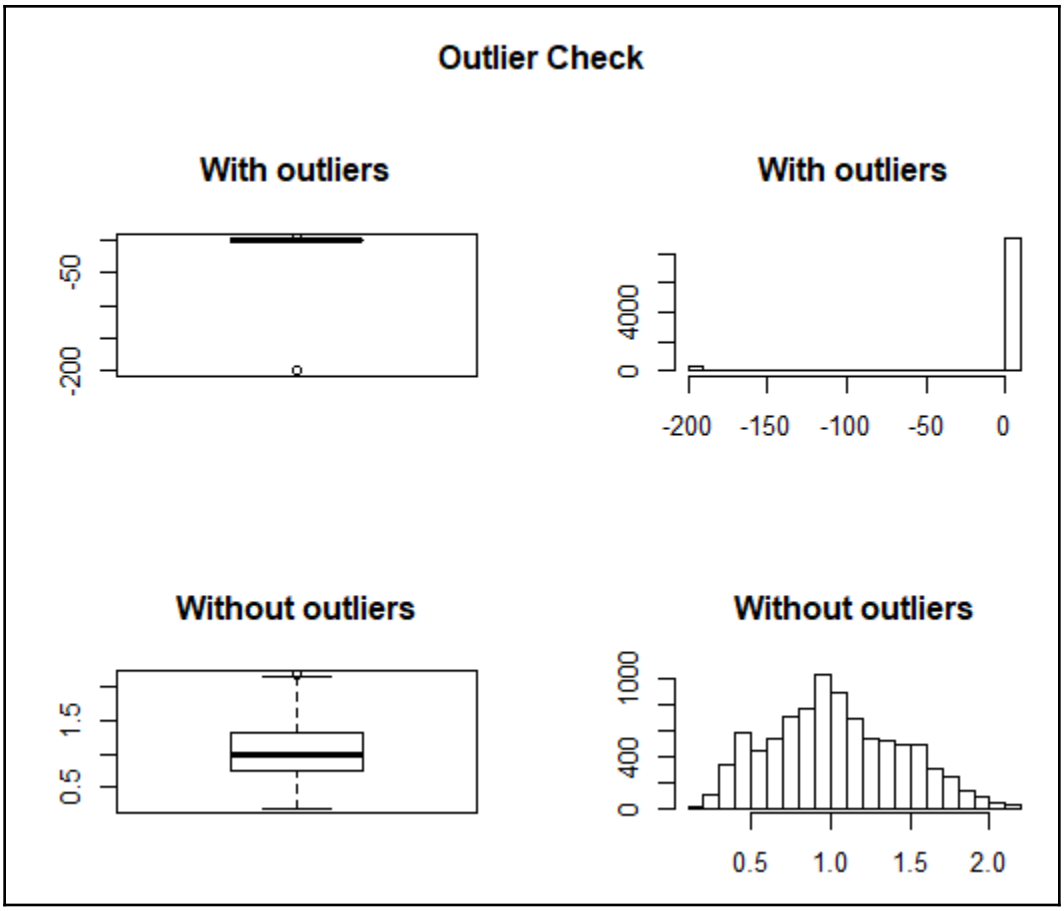
AirQualityUCI

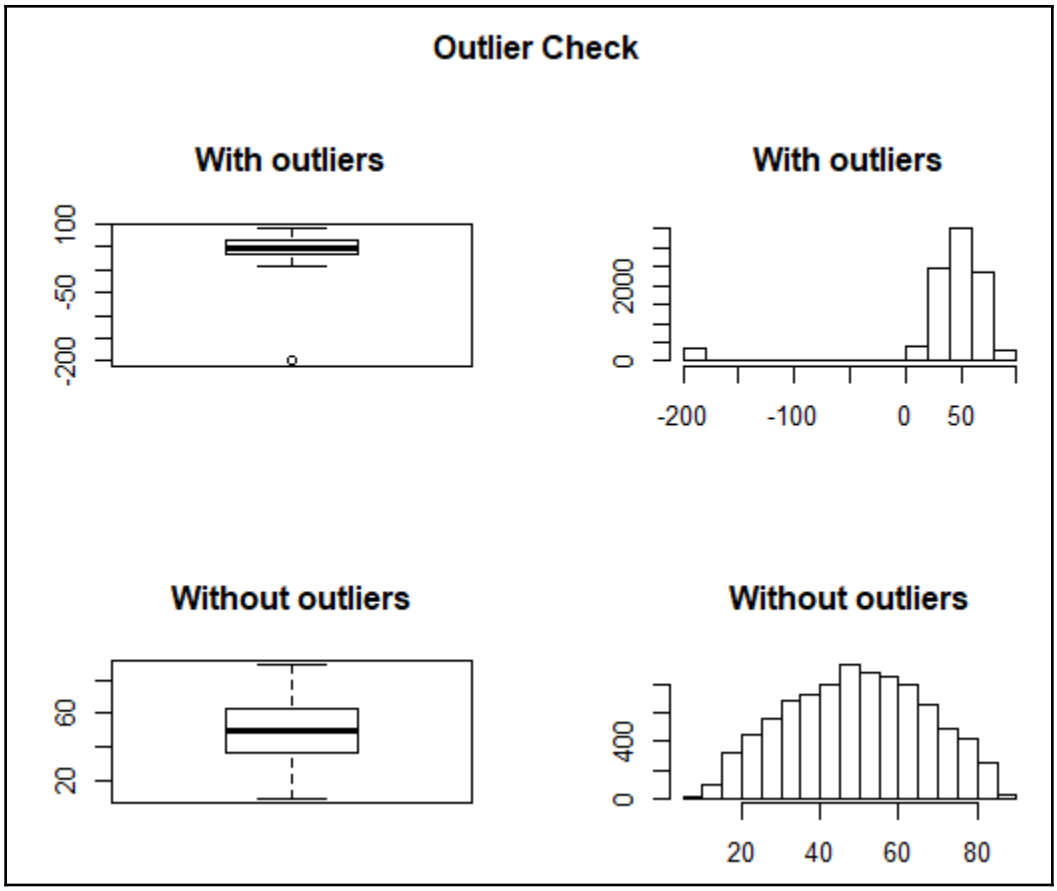
Filter

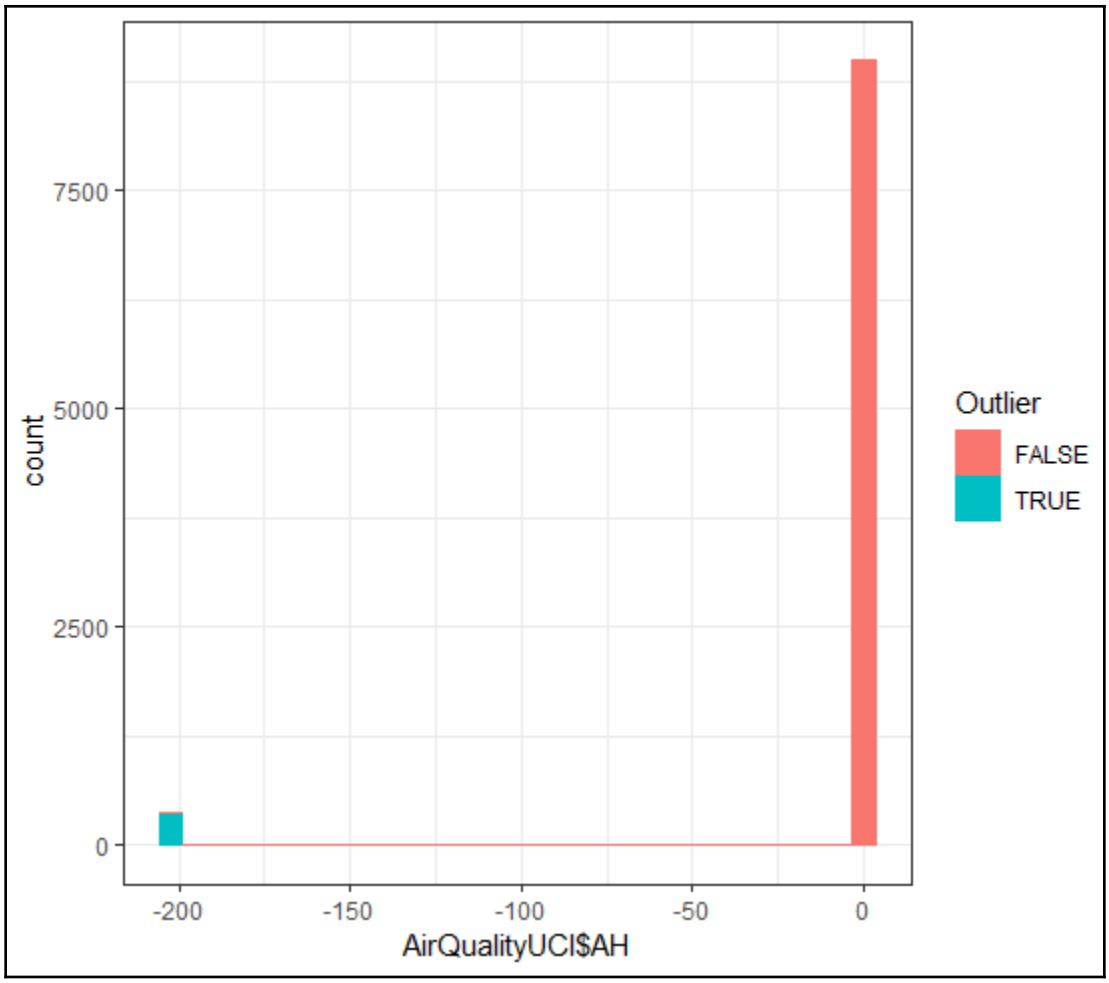
	Date	Time	CO(GT)	PT08.S1(CO)	NMHC(GT)	C6H6(GT)	PT08.S2(NMHC)	NOx(GT)	PT08.S3(NOx)	NO2(GT)	PT08.S4(NO2)
1	2004-03-10	1899-12-31 18:00:00	2.6	1360.000	150	11.881723	1045.50	166	1056.2500	113	1692.000
2	2004-03-10	1899-12-31 19:00:00	2.0	1292.250	112	9.397165	954.75	103	1173.7500	92	1558.750
3	2004-03-10	1899-12-31 20:00:00	2.2	1402.000	88	8.997817	939.25	131	1140.0000	114	1554.500
4	2004-03-10	1899-12-31 21:00:00	2.2	1375.500	80	9.228796	948.25	172	1092.0000	122	1583.750
5	2004-03-10	1899-12-31 22:00:00	1.6	1272.250	51	6.518224	835.50	131	1205.0000	116	1490.000
6	2004-03-10	1899-12-31 23:00:00	1.2	1197.000	38	4.741012	750.25	89	1336.5000	96	1393.000
7	2004-03-11	1899-12-31 00:00:00	1.2	1185.000	31	3.624399	689.50	62	1461.7500	77	1332.750
8	2004-03-11	1899-12-31 01:00:00	1.0	1136.250	31	3.326677	672.00	62	1453.2500	76	1332.750
9	2004-03-11	1899-12-31 02:00:00	0.9	1094.000	24	2.339416	608.50	45	1579.0000	60	1276.000
10	2004-03-11	1899-12-31 03:00:00	0.6	1009.750	19	1.696658	560.75	-200	1705.0000	-200	1234.750

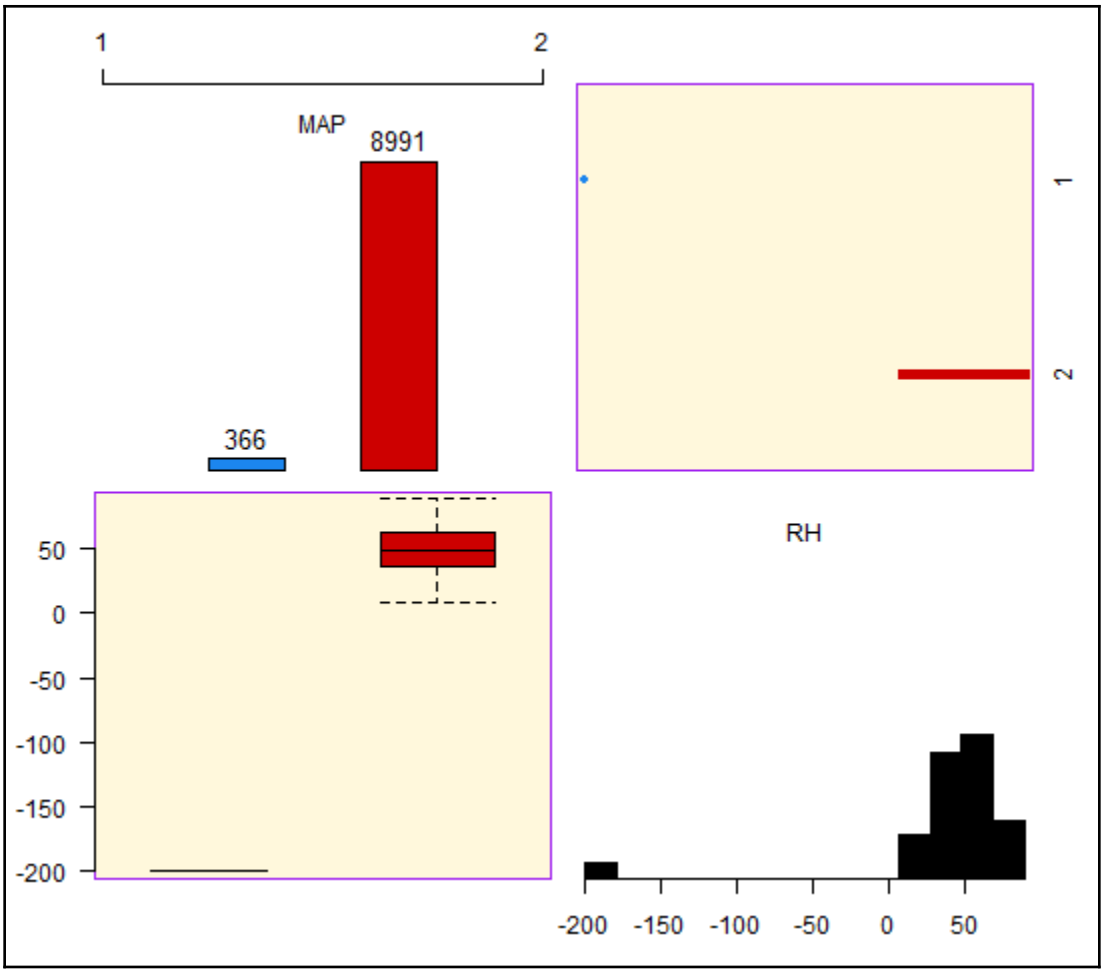
A	B
Date	Time
3/10/2004	18:00:00
3/10/2004	19:00:00
3/10/2004	20:00:00
3/10/2004	21:00:00
3/10/2004	22:00:00
3/10/2004	23:00:00
3/11/2004	0:00:00

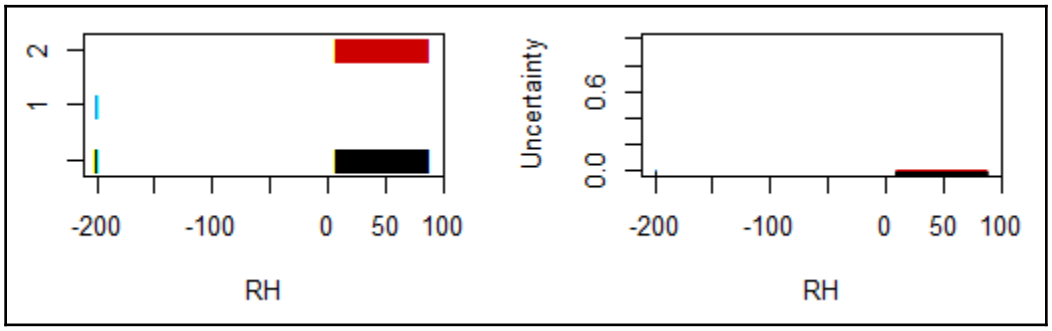

Humidity Analysis
Relative Humidity
100
50
0
-50
-100
-150
-200
-200
-150
-100
-50
0
Absolute Humidity

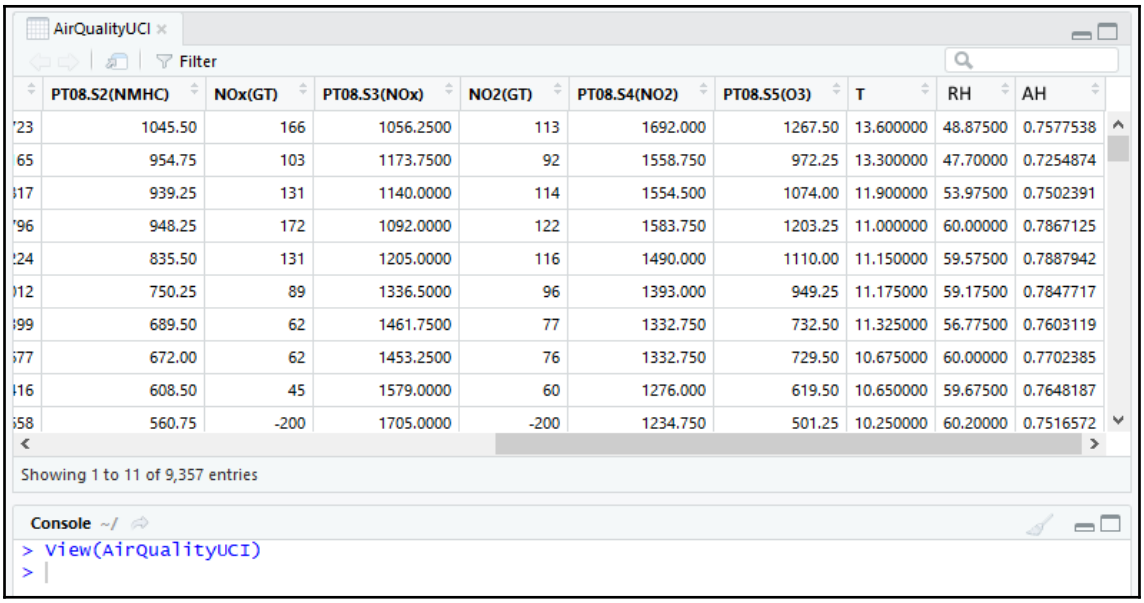

Outlier Check
With outliers
-50
-200
With outliers
0
4000
-200
-150
-100
-50
0
Without outliers
0.5
1.5
Without outliers
0
400
1000
0.5
1.0
1.5
2.0

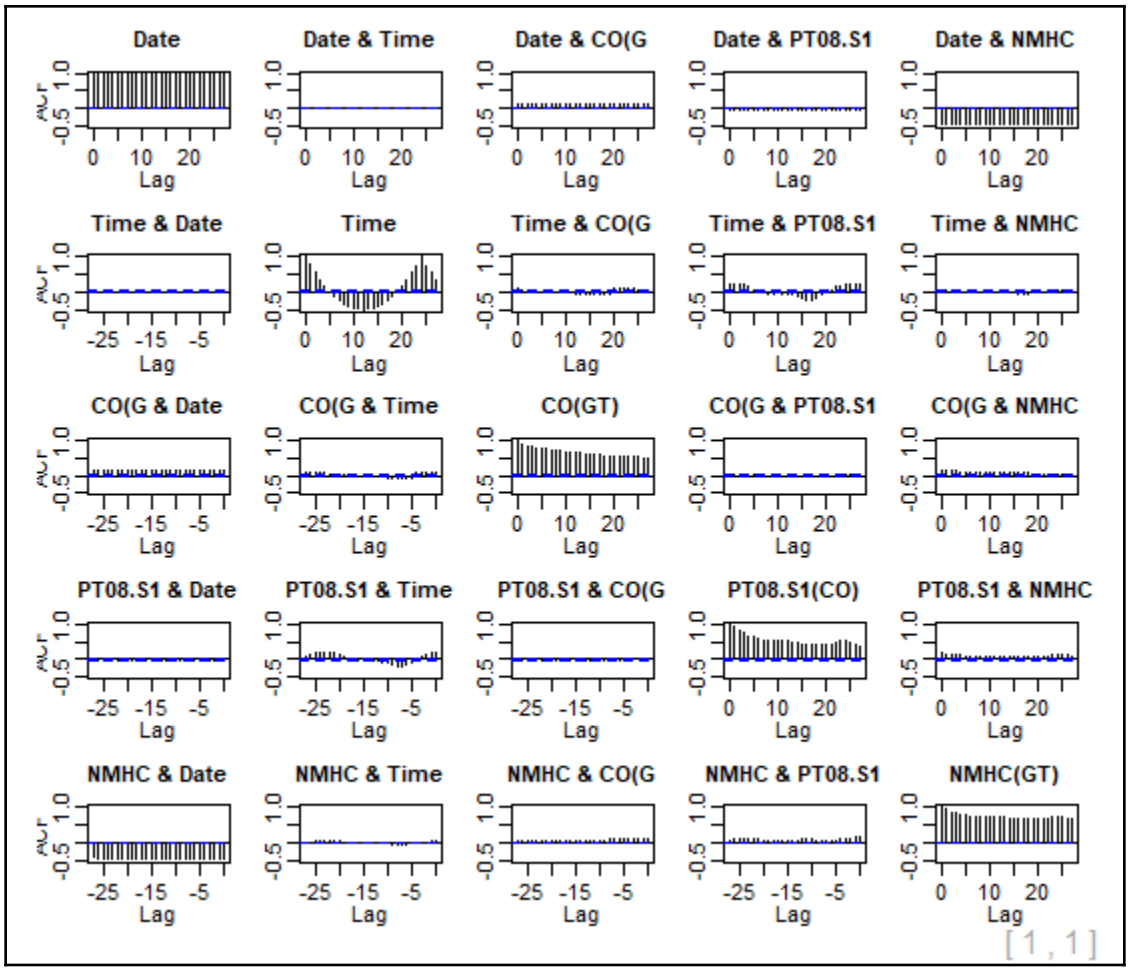

Outlier Check
With outliers
100
-50
-200
With outliers
2000
0
-200
-100
0
50
Without outliers
60
20
Without outliers
400
0
20
40
60
80

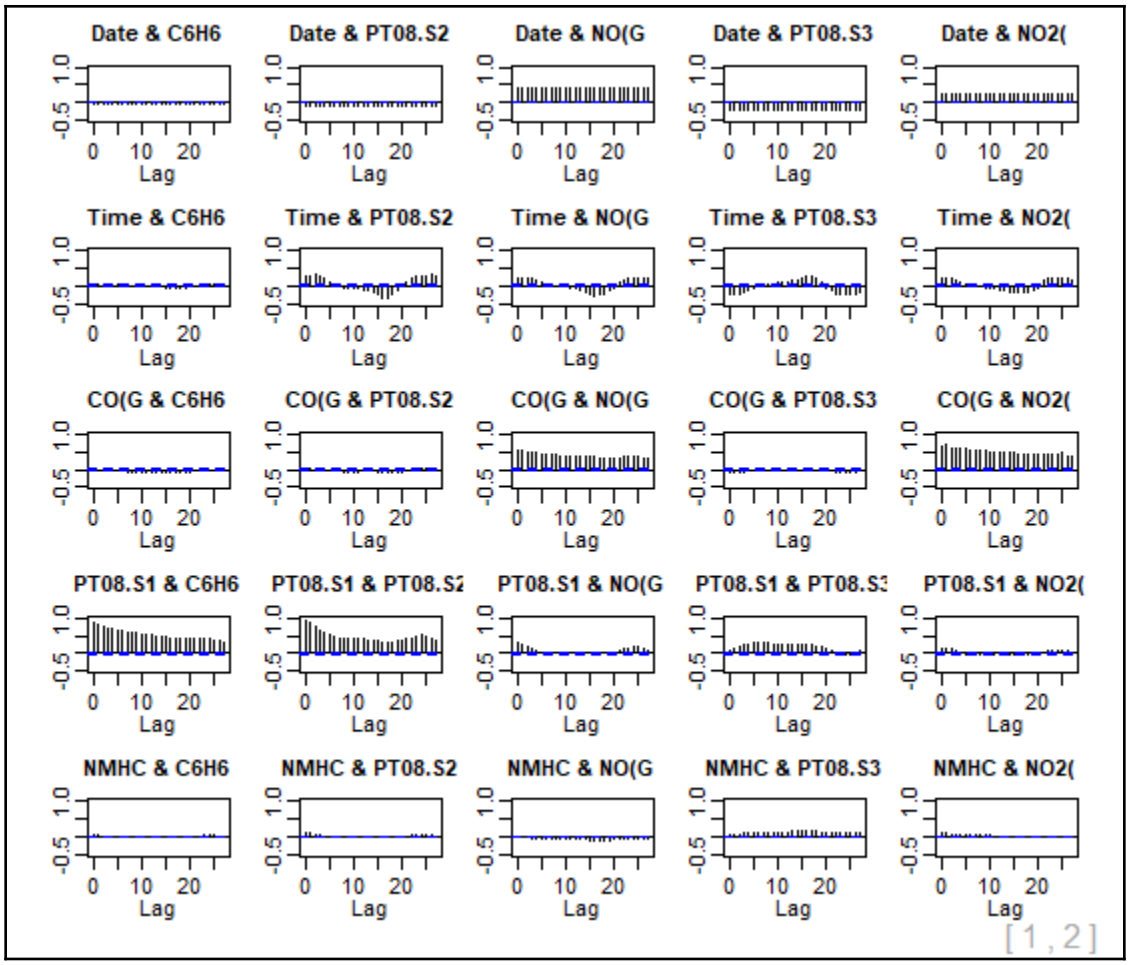

7500
5000
2500
0
count
-200
-150
-100
-50
0
AirQualityUCI$AH
Outlier
FALSE
TRUE

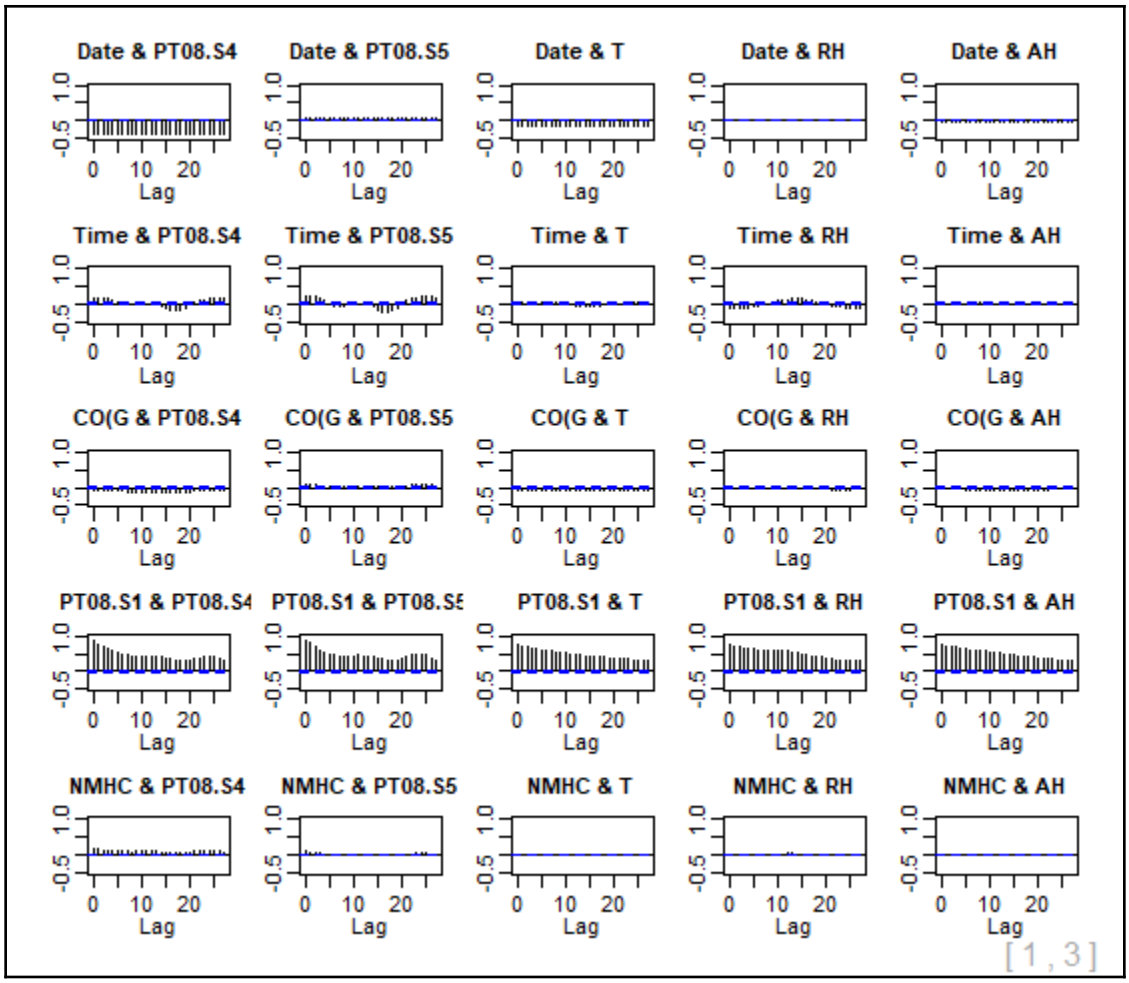

1
2
MAP
8991
366
50
0
-50
-100
-150
-200
RH
1
2
-200
-150
-100
-50
0
50

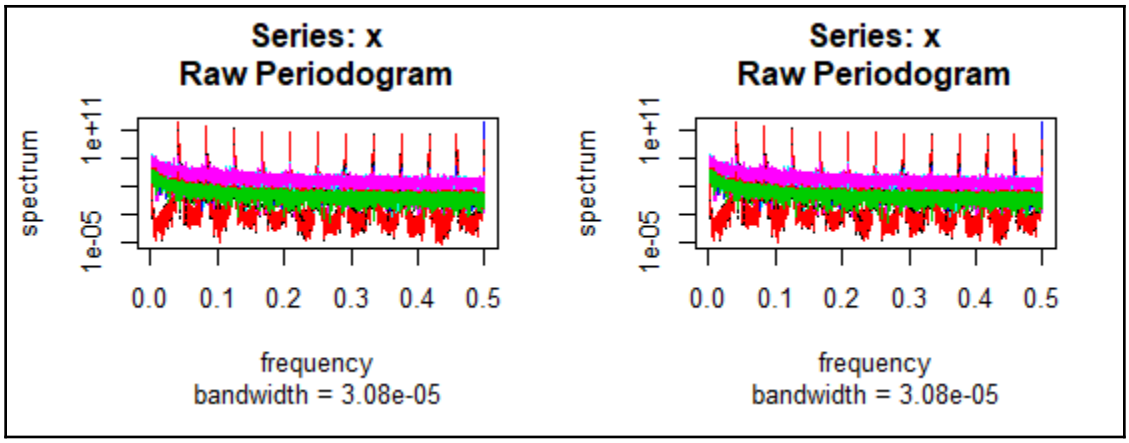

1
2
-200
-100
0
50
100
RH
Uncertainty
0.0
0.6
-200
-100
0
50
100
RH

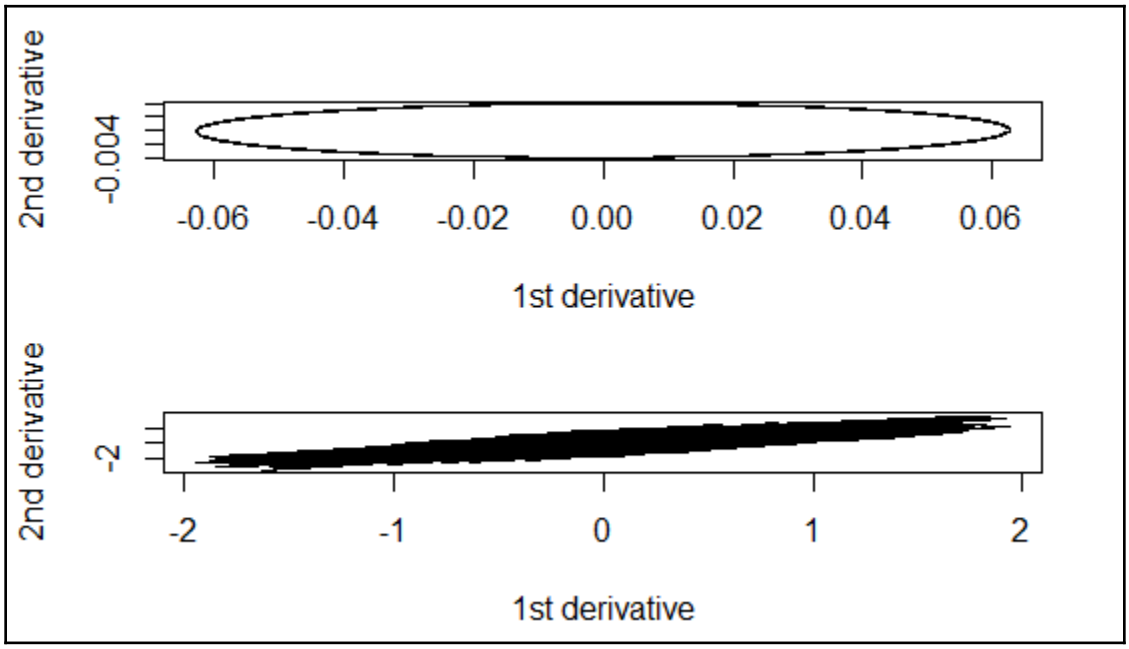

AirQualityUCI
Filter
PT08.S2(NMHC) NOx(GT) PT08.S3(NOx) NO2(GT) PT08.S4(NO2) PT08.S5(O3) T RH AH
723 1045.50 166 1056.2500 113 1692.000 1267.50 13.600000 48.87500 0.7577538
65 954.75 103 1173.7500 92 1558.750 972.25 13.300000 47.70000 0.7254874
317 939.25 131 1140.0000 114 1554.500 1074.00 11.900000 53.97500 0.7502391
796 948.25 172 1092.0000 122 1583.750 1203.25 11.000000 60.00000 0.7867125
224 835.50 131 1205.0000 116 1490.000 1110.00 11.150000 59.57500 0.7887942
012 750.25 89 1336.5000 96 1393.000 949.25 11.175000 59.17500 0.7847717
699 689.50 62 1461.7500 77 1332.750 732.50 11.325000 56.77500 0.7603119
577 672.00 62 1453.2500 76 1332.750 729.50 10.675000 60.00000 0.7702385
916 608.50 45 1579.0000 60 1276.000 619.50 10.650000 59.67500 0.7648187
558 560.75 -200 1705.0000 -200 1234.750 501.25 10.250000 60.20000 0.7516572
Showing 1 to 11 of 9,357 entries
Console ~/
> View(AirQualityUCI)
>


Date
Date & Time
Date & CO(G
Date & PT08.S1
Date & NMHC
Time & Date
Time
Time & CO(G
Time & PT08.S1
Time & NMHC
CO(G & Date
CO(G & Time
CO(GT)
CO(G & PT08.S1
CO(G & NMHC
PT08.S1 & Date
PT08.S1 & Time
PT08.S1 & CO(G
PT08.S1(CO)
PT08.S1 & NMHC
NMHC & Date
NMHC & Time
NMHC & CO(G
NMHC & PT08.S1
NMHC(GT)
ACF
Lag
[1,1]


Date & C6H6
Date & PT08.S2
Date & NO(G
Date & PT08.S3
Date & NO2(
Time & C6H6
Time & PT08.S2
Time & NO(G
Time & PT08.S3
Time & NO2(
CO(G & C6H6
CO(G & PT08.S2
CO(G & NO(G
CO(G & PT08.S3
CO(G & NO2(
PT08.S1 & C6H6
PT08.S1 & PT08.S2
PT08.S1 & NO(G
PT08.S1 & PT08.S3
PT08.S1 & NO2(
NMHC & C6H6
NMHC & PT08.S2
NMHC & NO(G
NMHC & PT08.S3
NMHC & NO2(
-0.5 1.0
0 10 20
Lag
[1,2]


Date & PT08.S4
Date & PT08.S5
Date & T
Date & RH
Date & AH
Time & PT08.S4
Time & PT08.S5
Time & T
Time & RH
Time & AH
CO(G & PT08.S4
CO(G & PT08.S5
CO(G & T
CO(G & RH
CO(G & AH
PT08.S1 & PT08.S4
PT08.S1 & PT08.S5
PT08.S1 & T
PT08.S1 & RH
PT08.S1 & AH
NMHC & PT08.S4
NMHC & PT08.S5
NMHC & T
NMHC & RH
NMHC & AH
-0.5 1.0
0 10 20
Lag
[1,3]


Series: x
Raw Periodogram
spectrum
1e+11
1e-05
0.0
0.1
0.2
0.3
0.4
0.5
frequency
bandwidth = 3.08e-05
Series: x
Raw Periodogram
spectrum
1e+11
1e-05
0.0
0.1
0.2
0.3
0.4
0.5
frequency
bandwidth = 3.08e-05


2nd derivative
-0.004
-0.06
-0.04
-0.02
0.00
0.02
0.04
0.06
1st derivative
2nd derivative
-2
-2
-1
0
1
2
1st derivative

Chapter 8: Multivariate Datasets

longley

Filter

	GNP.deflator	GNP	Unemployed	Armed.Forces	Population	Year	Employed
1947	83.0	234.289	235.6	159.0	107.608	1947	60.323
1948	88.5	259.426	232.5	145.6	108.632	1948	61.122
1949	88.2	258.054	368.2	161.6	109.773	1949	60.171
1950	89.5	284.599	335.1	165.0	110.929	1950	61.187
1951	96.2	328.975	209.9	309.9	112.075	1951	63.221
1952	98.1	346.999	193.2	359.4	113.270	1952	63.639
1953	99.0	365.385	187.0	354.7	115.094	1953	64.989
1954	100.0	363.112	357.8	335.0	116.219	1954	63.761
1955	101.2	397.469	290.4	304.8	117.388	1955	66.019
1956	104.6	419.180	282.2	285.7	118.734	1956	67.857

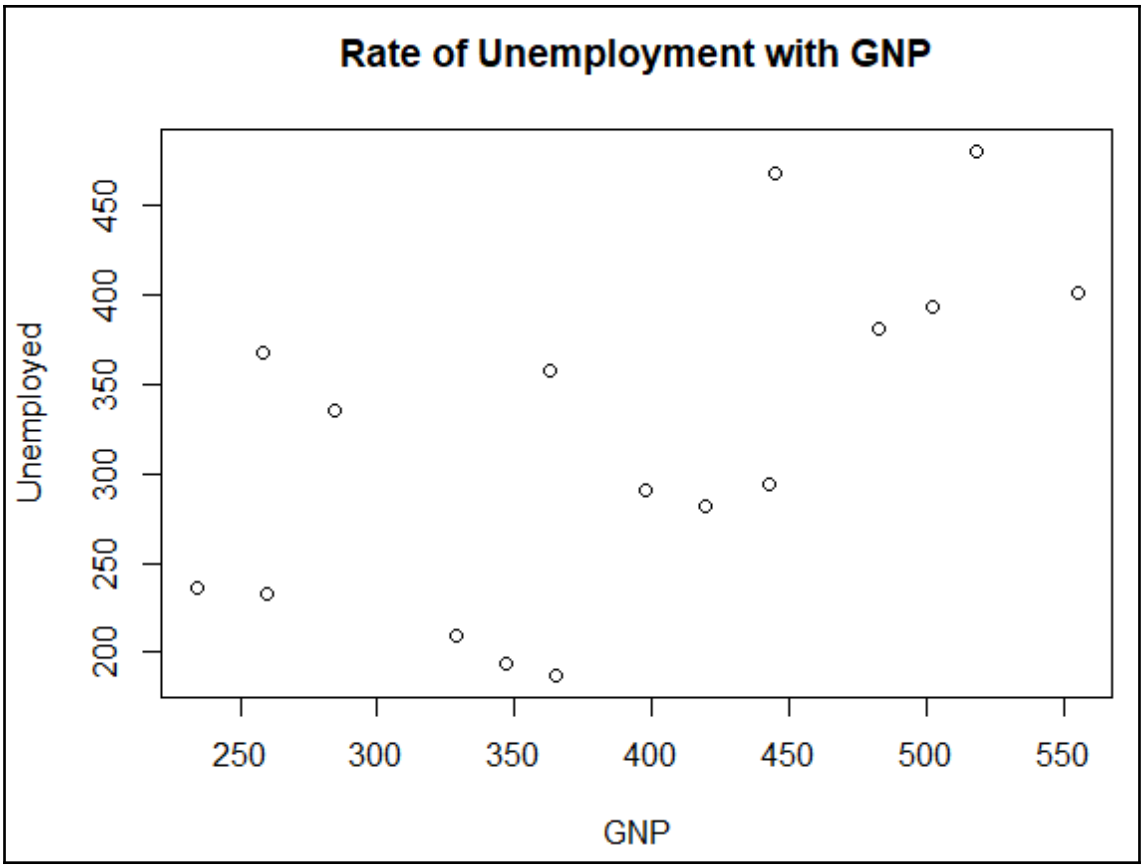
Showing 1 to 11 of 16 entries

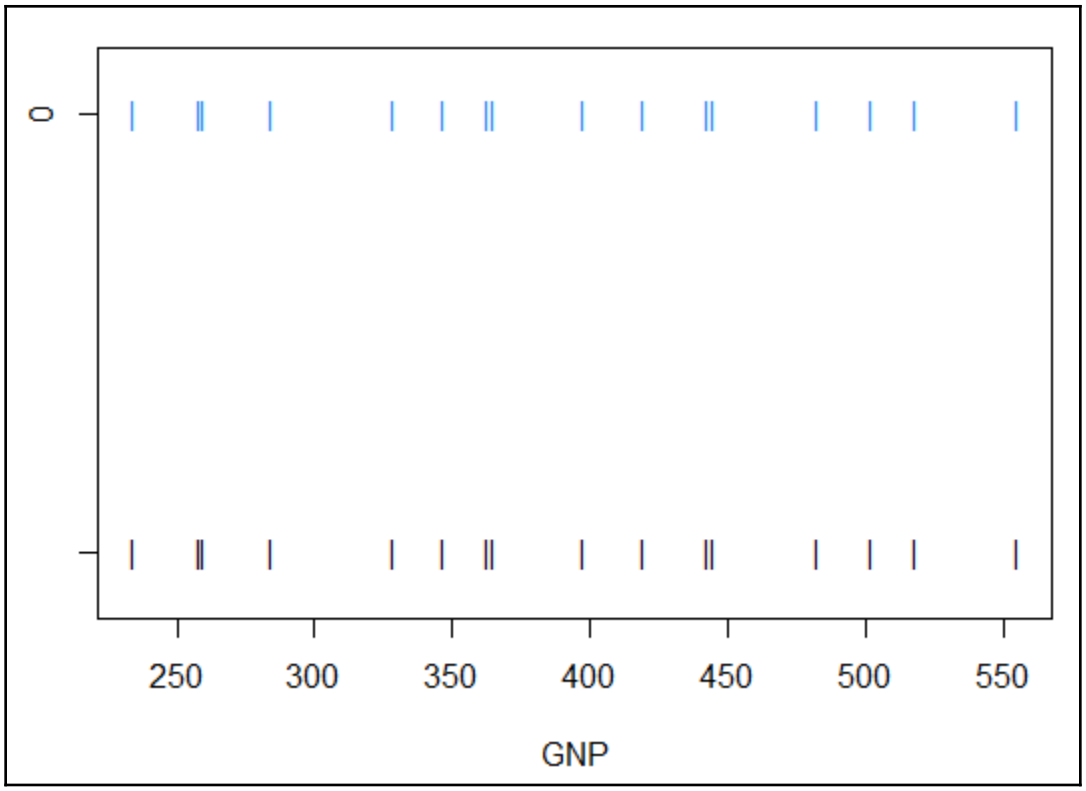
longley

Filter

	GNP Deflator	GNP	Unemployed	Armed Forces	Population	Year	Employed
1947	83.0	234.289	235.6	159.0	107.608	1947	60.323
1948	88.5	259.426	232.5	145.6	108.632	1948	61.122
1949	88.2	258.054	368.2	161.6	109.773	1949	60.171
1950	89.5	284.599	335.1	165.0	110.929	1950	61.187
1951	96.2	328.975	209.9	309.9	112.075	1951	63.221
1952	98.1	346.999	193.2	359.4	113.270	1952	63.639
1953	99.0	365.385	187.0	354.7	115.094	1953	64.989
1954	100.0	363.112	357.8	335.0	116.219	1954	63.761
1955	101.2	397.469	290.4	304.8	117.388	1955	66.019
1956	104.6	419.180	282.2	285.7	118.734	1956	67.857

Showing 1 to 10 of 16 entries


Rate of Unemployment with GNP
Unemployed
200
250
300
350
400
450
250
300
350
400
450
500
550
GNP

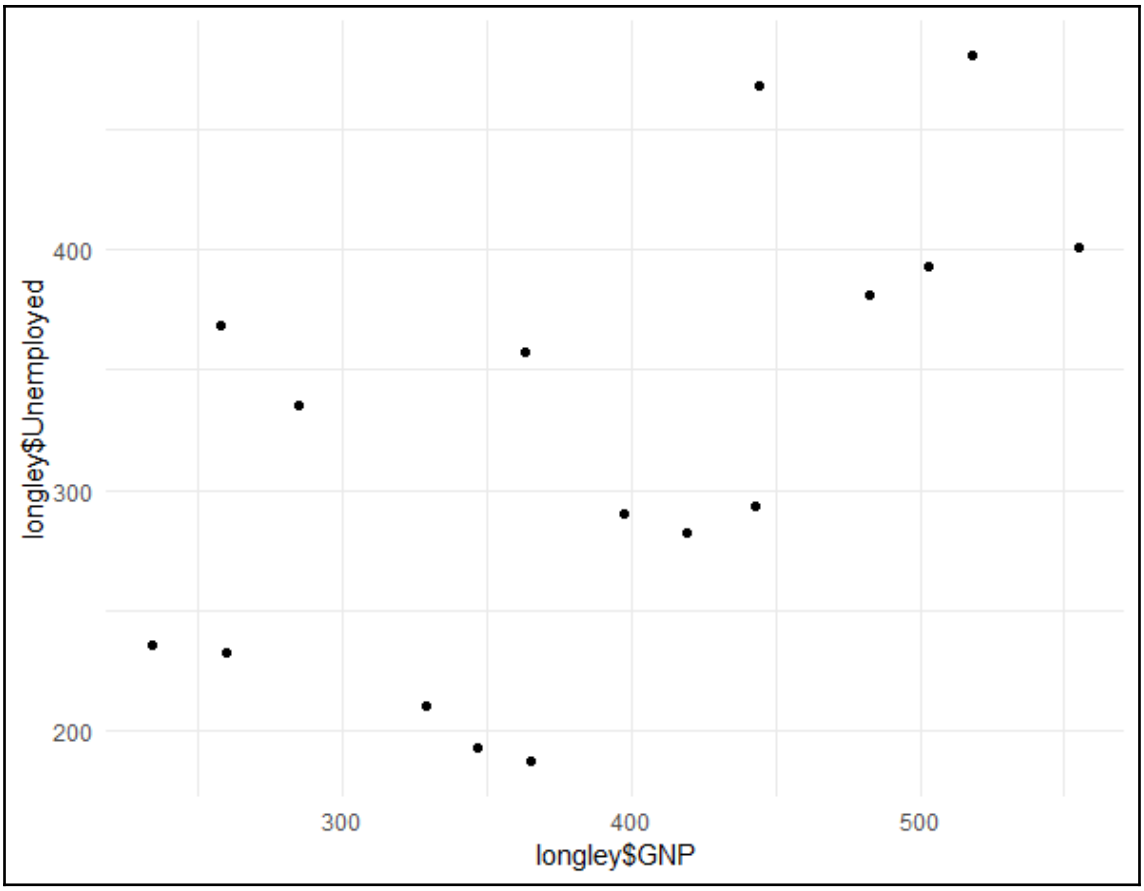


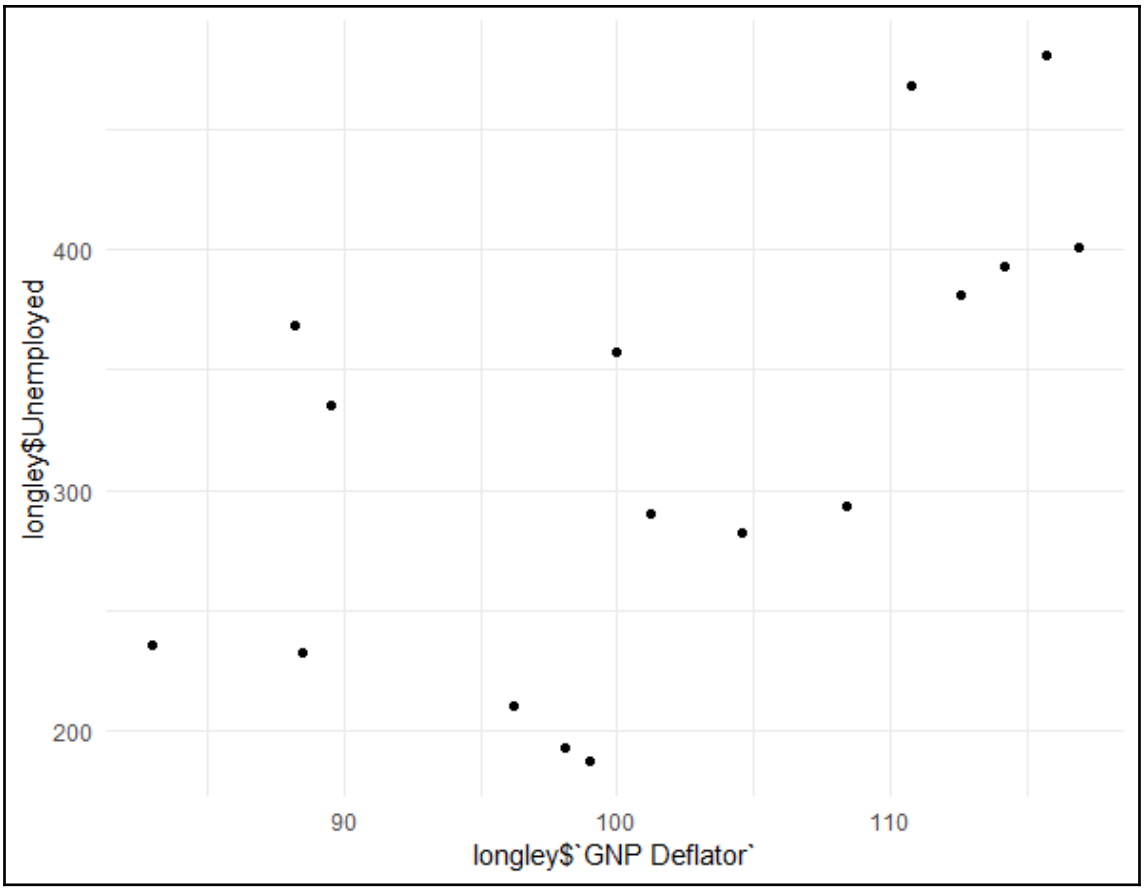
longley

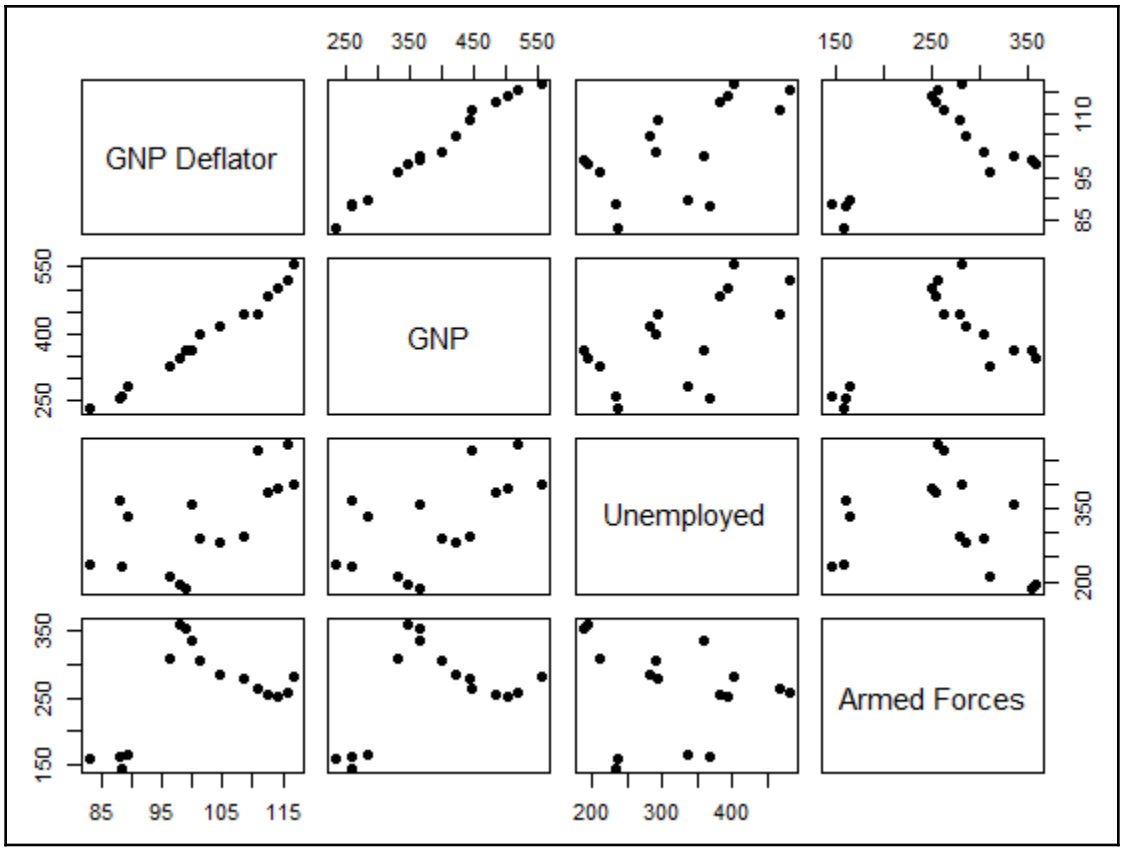
Filter

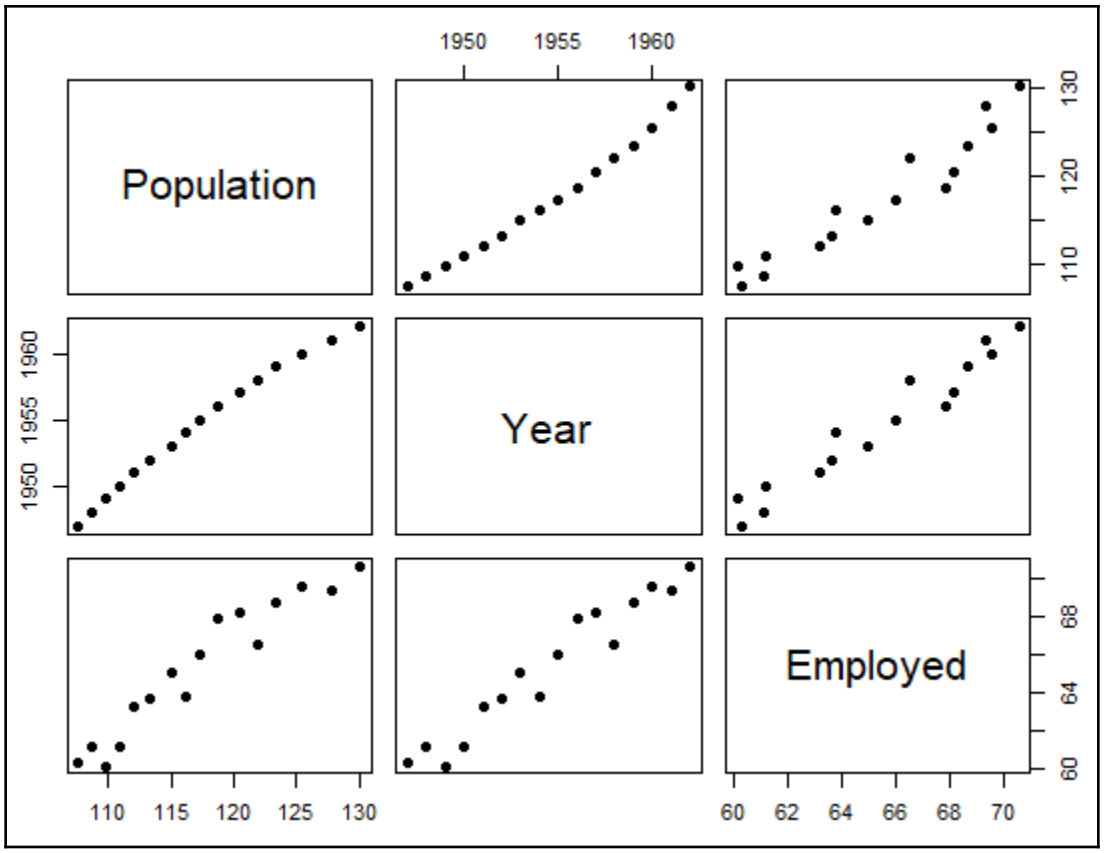
	GNP Deflator	GNP	Unemployed	Armed Forces	Population	Year	Employed
1947	83.0	234.289	235.6	159.0	107.608	1947	60.323
1948	88.5	259.426	232.5	145.6	108.632	1948	61.122
1949	88.2	258.054	368.2	161.6	109.773	1949	60.171
1950	89.5	284.599	335.1	165.0	110.929	1950	61.187
1951	96.2	328.975	209.9	309.9	112.075	1951	63.221
1952	98.1	346.999	193.2	359.4	113.270	1952	63.639
1953	99.0	365.385	187.0	354.7	115.094	1953	64.989
1954	100.0	363.112	357.8	335.0	116.219	1954	63.761
1955	101.2	397.469	290.4	304.8	117.388	1955	66.019
1956	104.6	419.180	282.2	285.7	118.734	1956	67.857

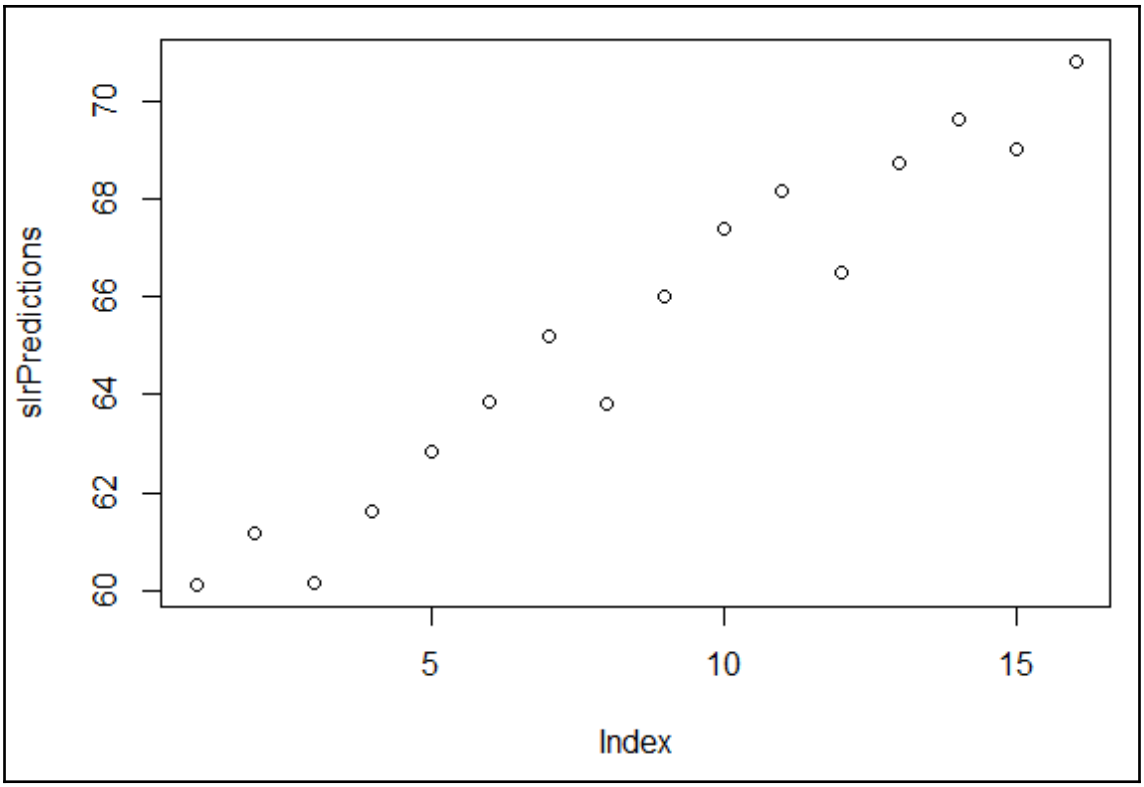
Showing 1 to 10 of 16 entries

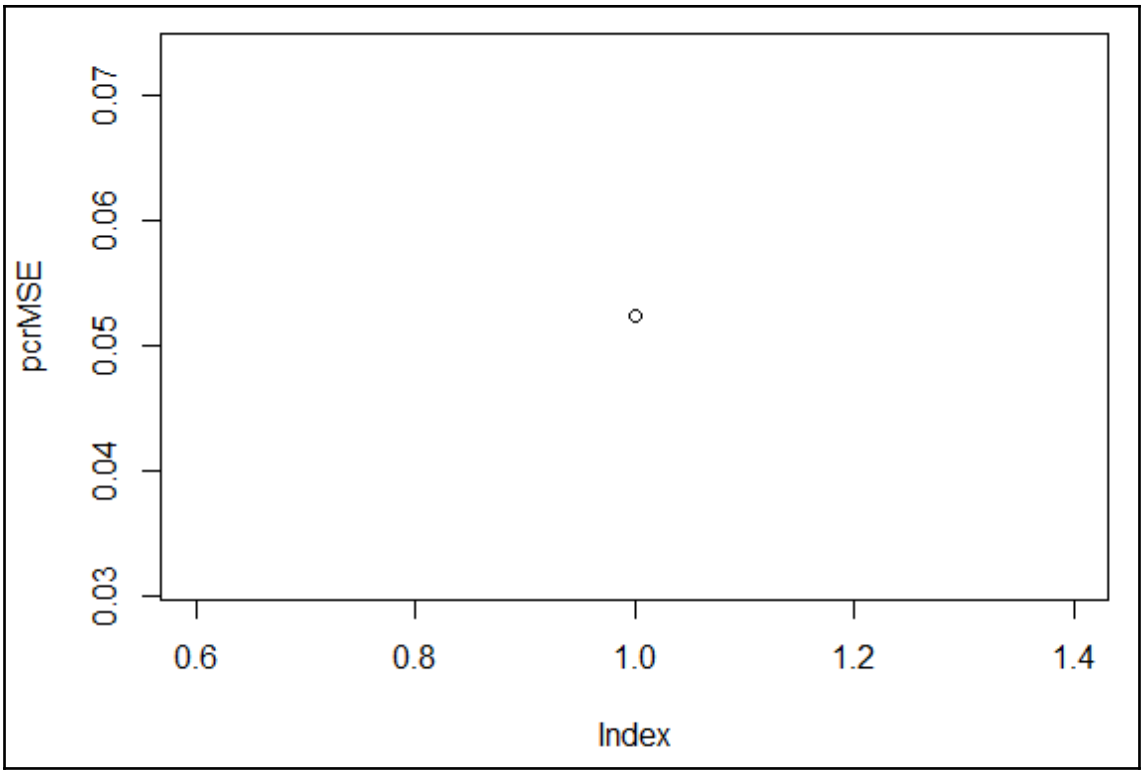

longley$Unemployed
400
300
200
300
400
500
longley$GNP

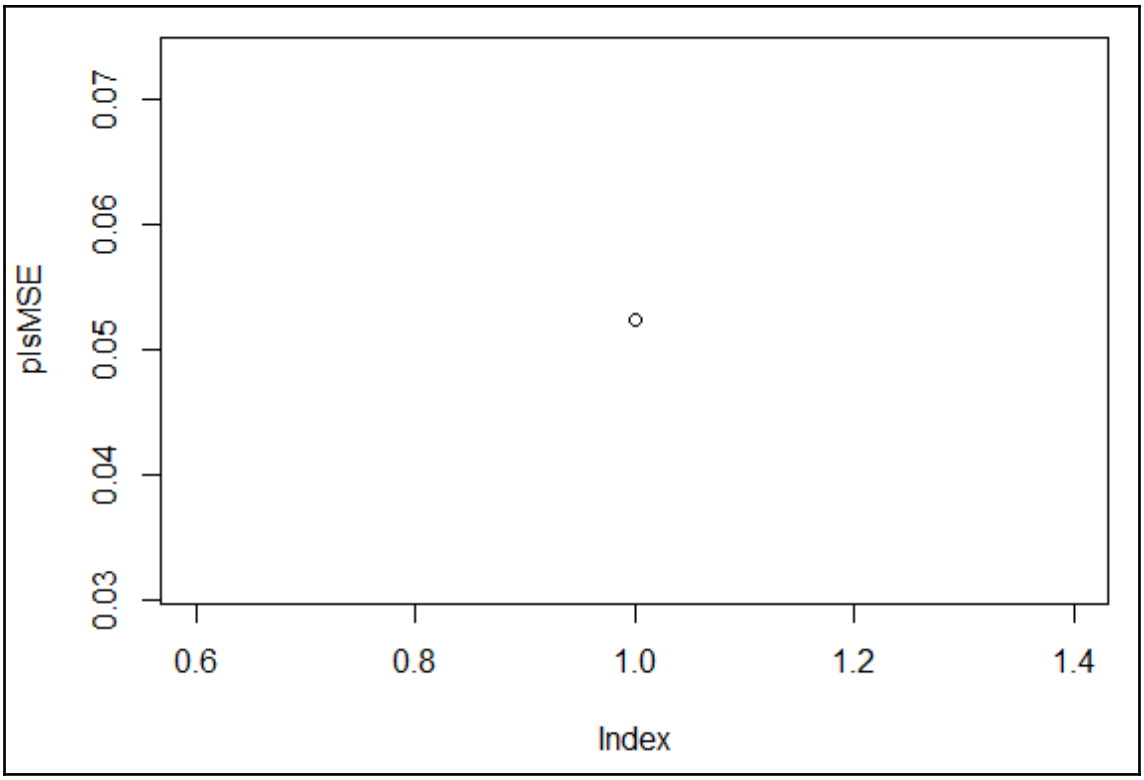

longley$Unemployed
400
300
200
90
100
110
longley$`GNP Deflator`


GNP Deflator
GNP
Unemployed
Armed Forces
250 350 450 550
150 250 350
85 95 110
250 400 550
200 350
150 250 350
85 95 105 115
200 300 400

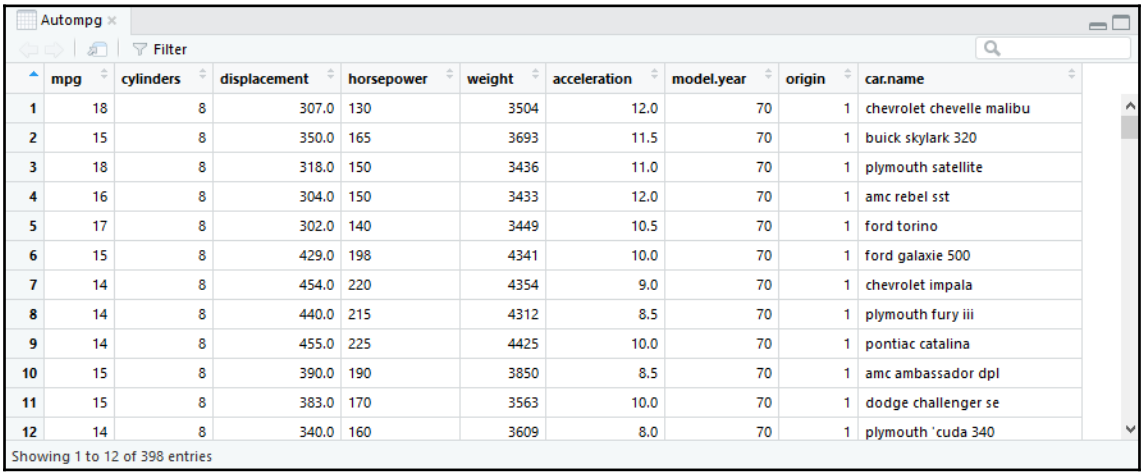

Population
Year
Employed
1950
1955
1960
110
115
120
125
130
60
62
64
66
68
70


slrPredictions
60
62
64
66
68
70
5
10
15
Index


pcrMSE
0.03
0.04
0.05
0.06
0.07
0.6
0.8
1.0
1.2
1.4
Index


plsMSE
0.03
0.04
0.05
0.06
0.07
0.6
0.8
1.0
1.2
1.4
Index

Chapter 9: Multi-Factor Datasets

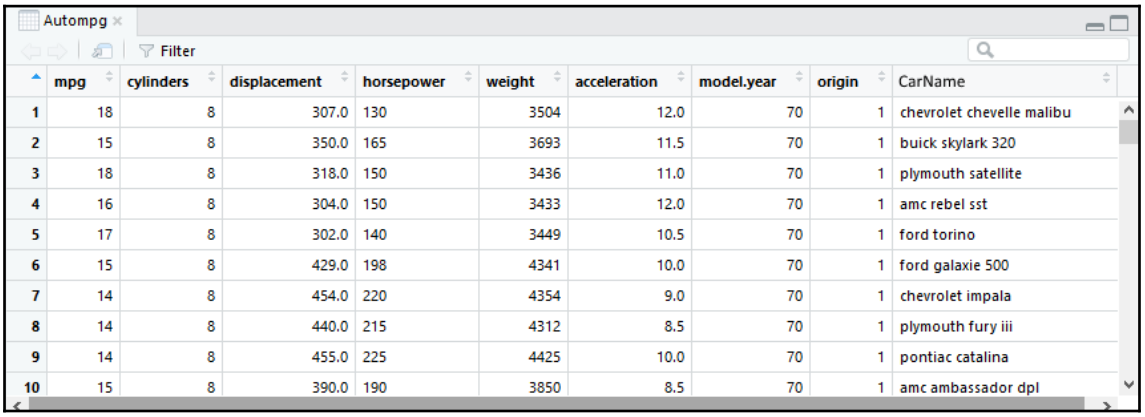


Autompg

Filter

	mpg	cylinders	displacement	horsepower	weight	acceleration	model.year	origin	car.name
1	18	8	307.0	130	3504	12.0	70	1	chevrolet chevelle malibu
2	15	8	350.0	165	3693	11.5	70	1	buick skylark 320
3	18	8	318.0	150	3436	11.0	70	1	plymouth satellite
4	16	8	304.0	150	3433	12.0	70	1	amc rebel sst
5	17	8	302.0	140	3449	10.5	70	1	ford torino
6	15	8	429.0	198	4341	10.0	70	1	ford galaxie 500
7	14	8	454.0	220	4354	9.0	70	1	chevrolet impala
8	14	8	440.0	215	4312	8.5	70	1	plymouth fury iii
9	14	8	455.0	225	4425	10.0	70	1	pontiac catalina
10	15	8	390.0	190	3850	8.5	70	1	amc ambassador dpl
11	15	8	383.0	170	3563	10.0	70	1	dodge challenger se
12	14	8	340.0	160	3609	8.0	70	1	plymouth 'cuda 340

Showing 1 to 12 of 398 entries



Autompg

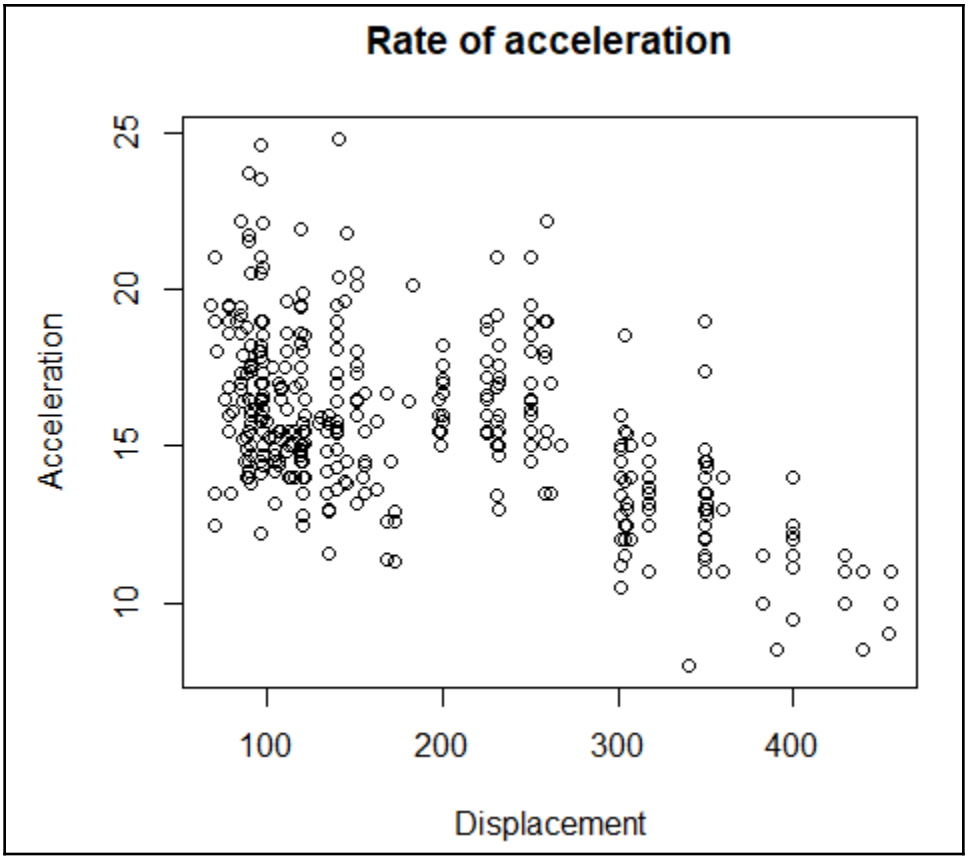
Filter

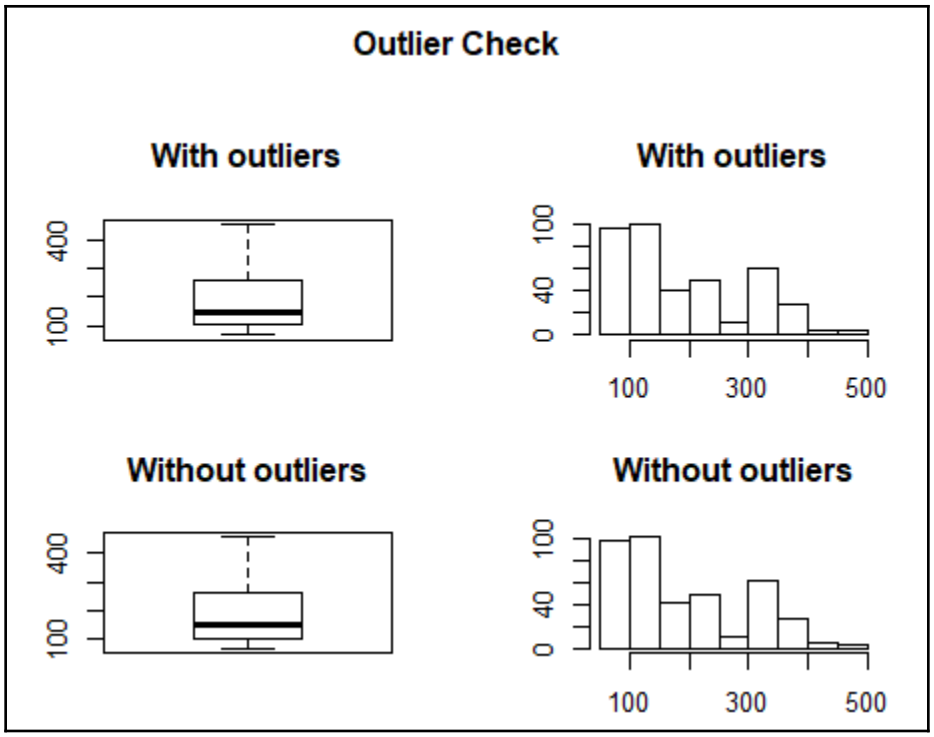
	mpg	cylinders	displacement	horsepower	weight	acceleration	model.year	origin	CarName
1	18	8	307.0	130	3504	12.0	70	1	chevrolet chevelle malibu
2	15	8	350.0	165	3693	11.5	70	1	buick skylark 320
3	18	8	318.0	150	3436	11.0	70	1	plymouth satellite
4	16	8	304.0	150	3433	12.0	70	1	amc rebel sst
5	17	8	302.0	140	3449	10.5	70	1	ford torino
6	15	8	429.0	198	4341	10.0	70	1	ford galaxie 500
7	14	8	454.0	220	4354	9.0	70	1	chevrolet impala
8	14	8	440.0	215	4312	8.5	70	1	plymouth fury iii
9	14	8	455.0	225	4425	10.0	70	1	pontiac catalina
10	15	8	390.0	190	3850	8.5	70	1	amc ambassador dpl

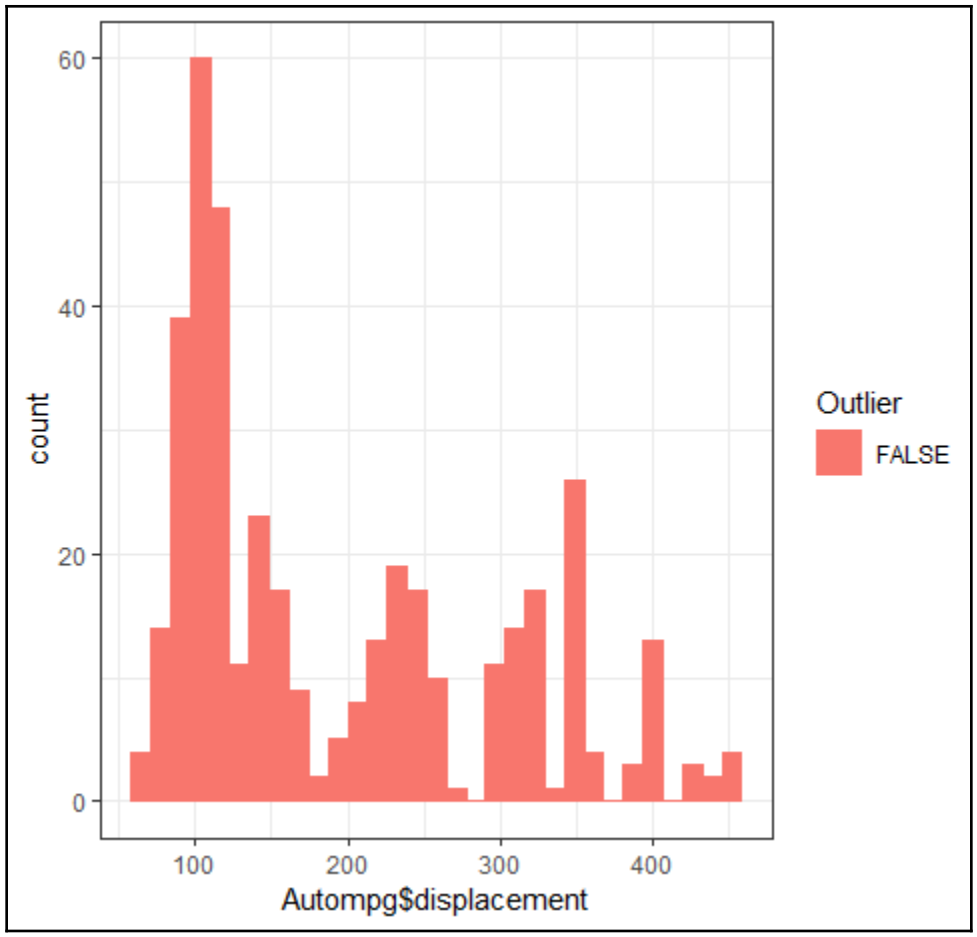
who × Autompg ×

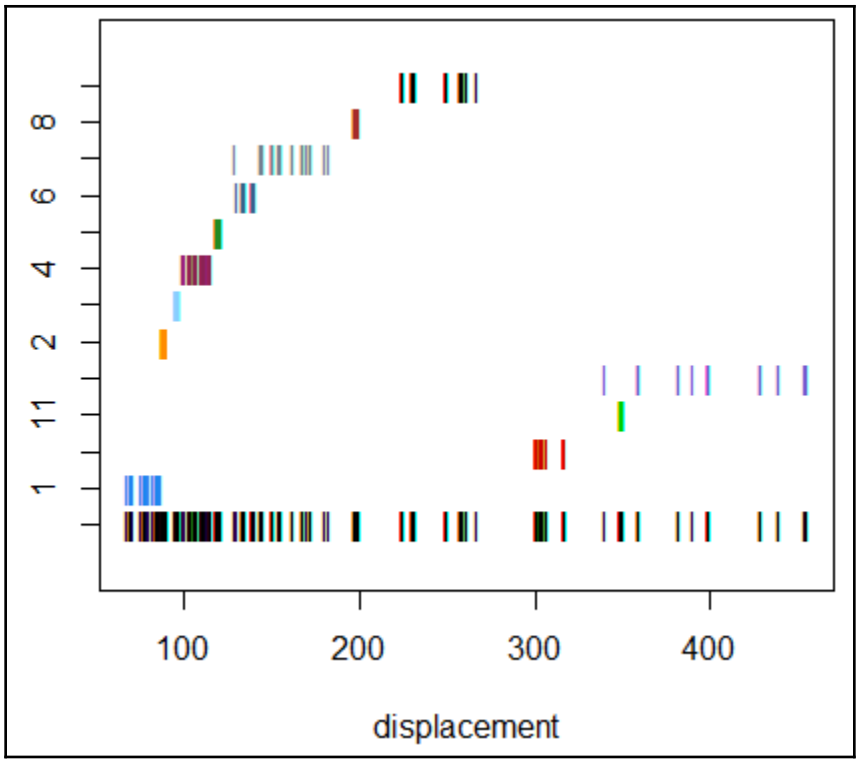
Filter

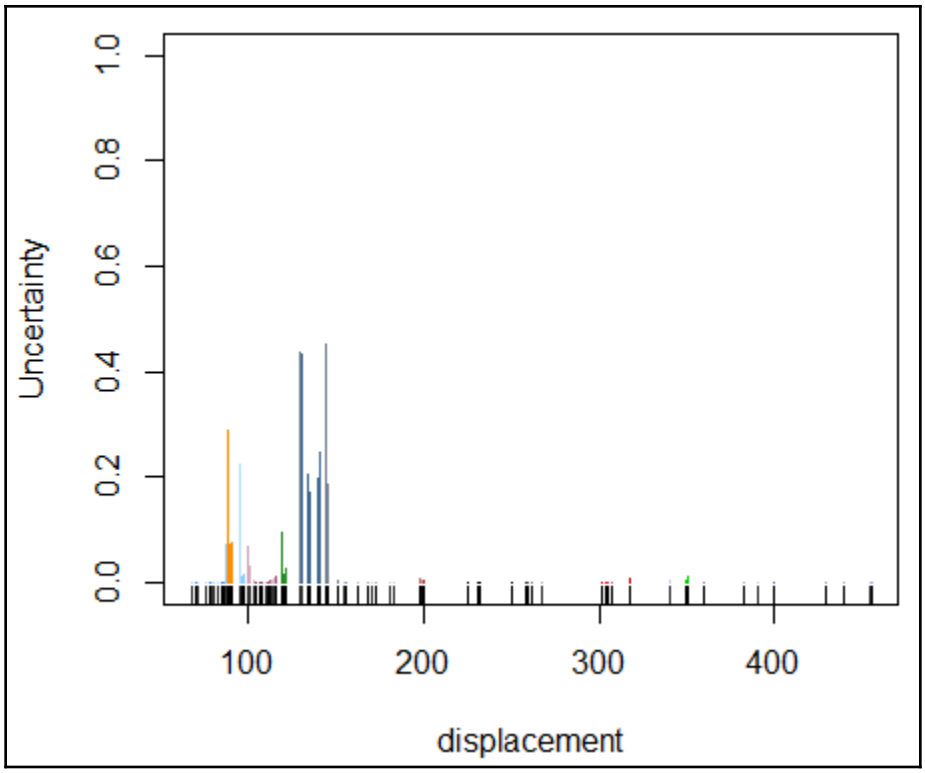
	mpg	cylinders	displacement	horsepower	weight	acceleration	model.year	origin	CarName
1	18	8	307.0	130	3504	12.0	70	1	chevrolet chevelle malibu
2	15	8	350.0	165	3693	11.5	70	1	buick skylark 320
3	18	8	318.0	150	3436	11.0	70	1	plymouth satellite
4	16	8	304.0	150	3433	12.0	70	1	amc rebel sst
5	17	8	302.0	140	3449	10.5	70	1	ford torino
6	15	8	429.0	198	4341	10.0	70	1	ford galaxie 500
7	14	8	454.0	220	4354	9.0	70	1	chevrolet impala
8	14	8	440.0	215	4312	8.5	70	1	plymouth fury iii
9	14	8	455.0	225	4425	10.0	70	1	pontiac catalina
10	15	8	390.0	190	3850	8.5	70	1	amc ambassador dpl

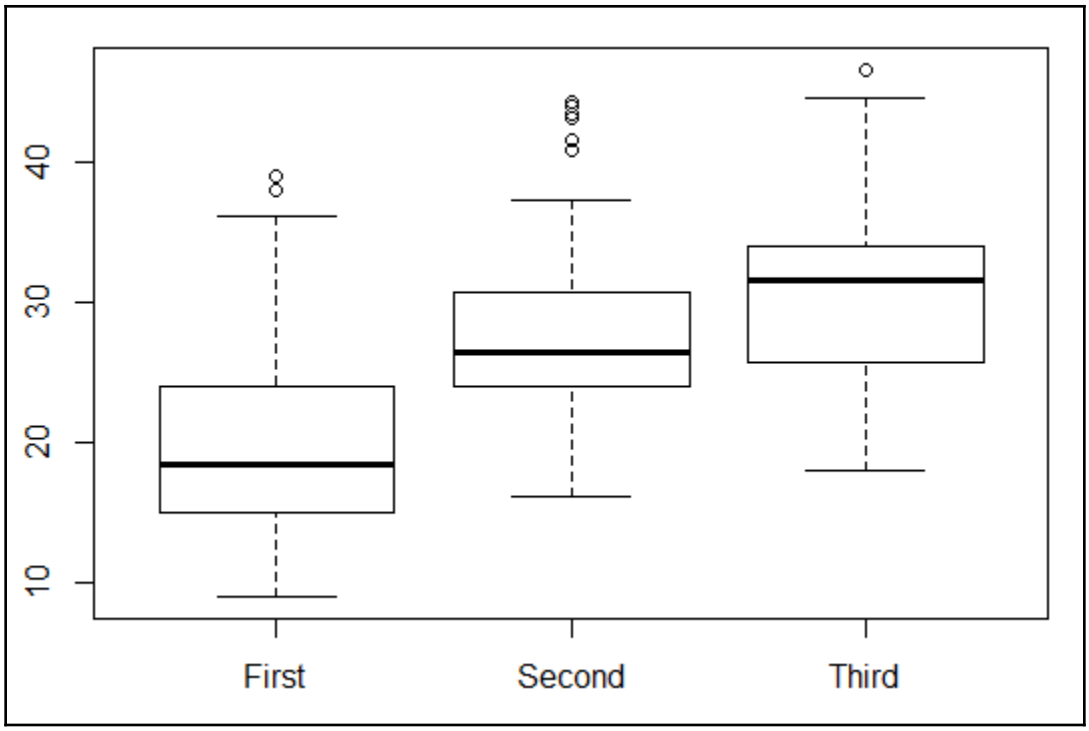


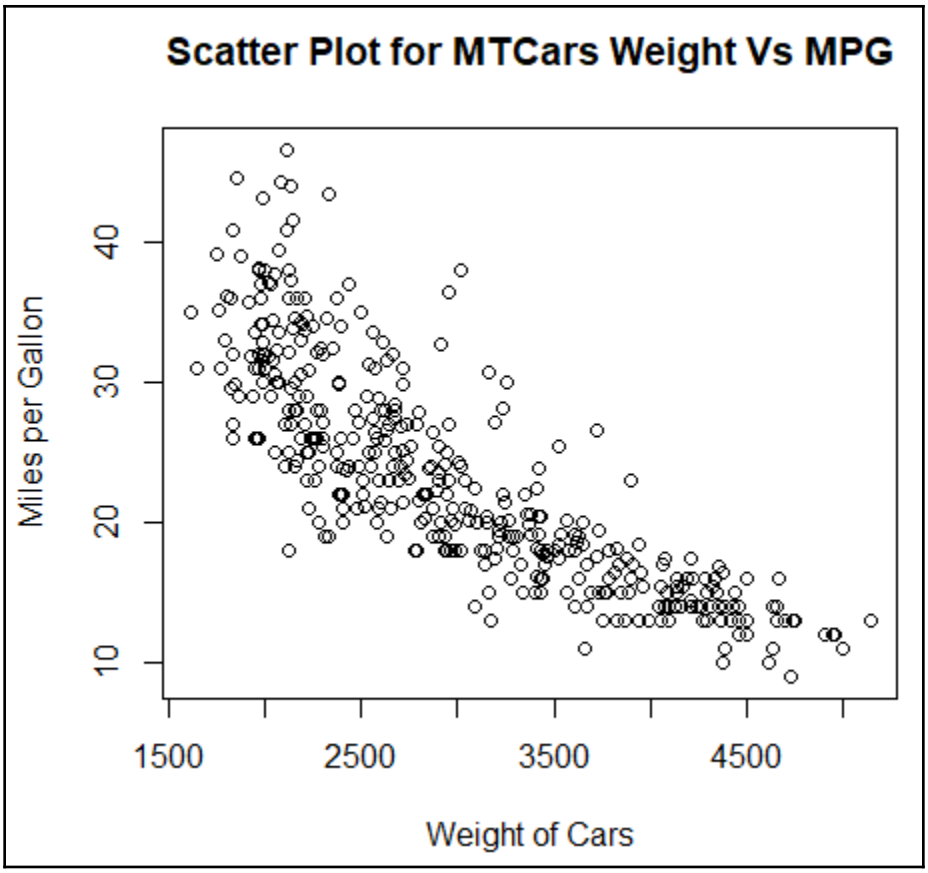

Outlier Check
With outliers
With outliers
100
400
0
40
100
100
300
500
Without outliers
Without outliers
100
400
0
40
100
100
300
500

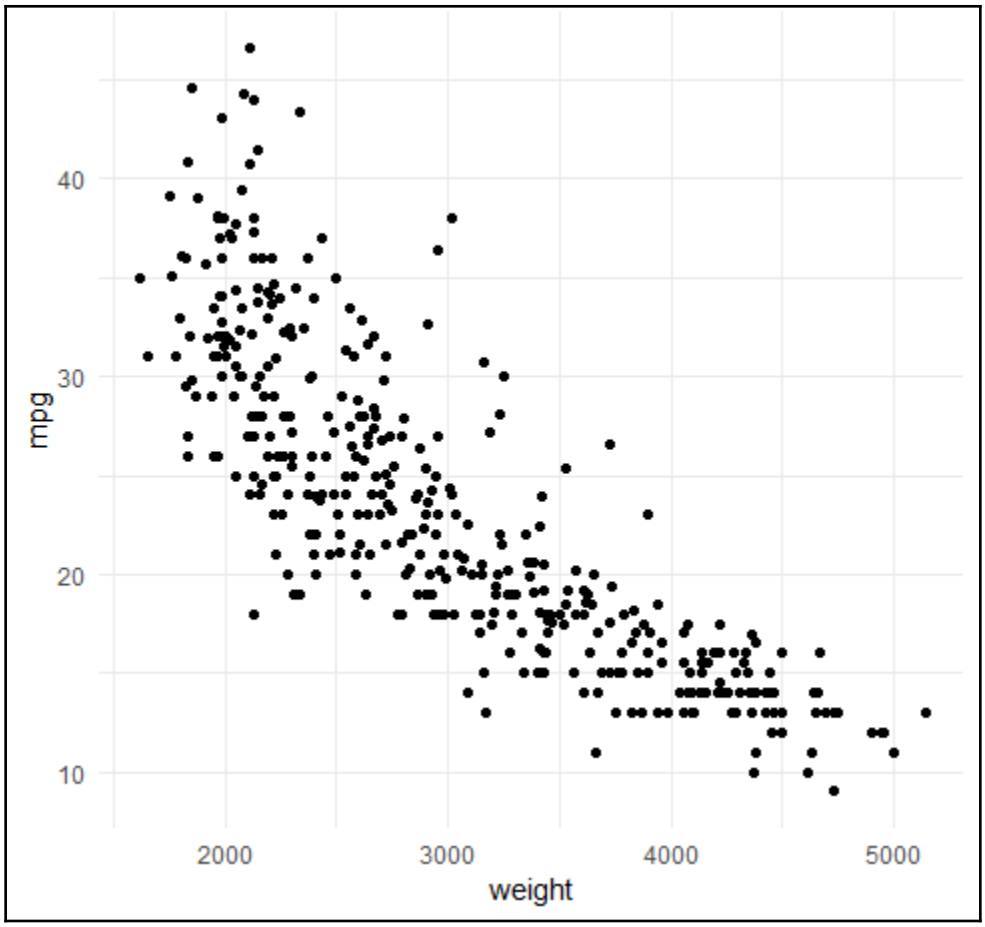

60
40
20
0
count
100
200
300
400
Autompg$displacement
Outlier
FALSE

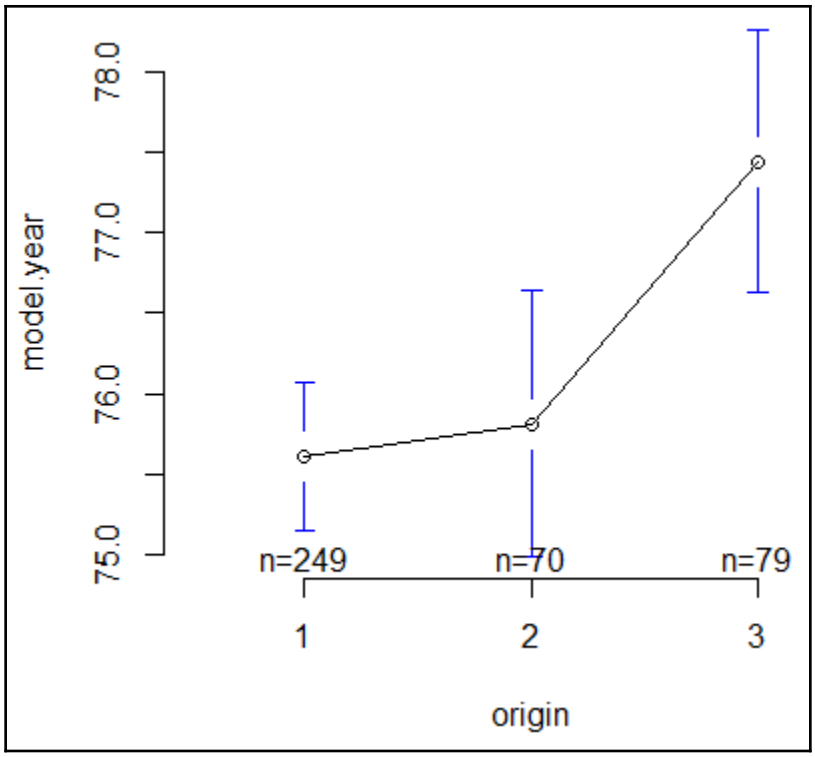

1
11
2
4
6
8
100
200
300
400
displacement

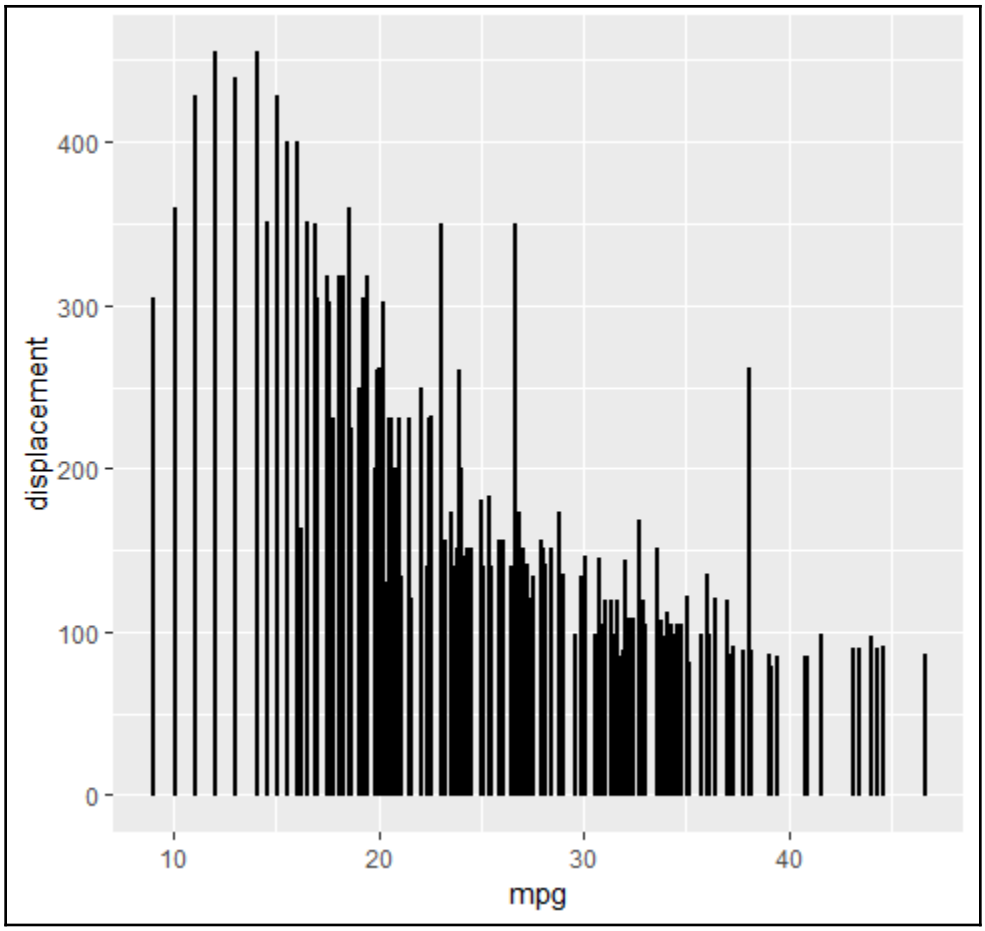

Uncertainty
0.0
0.2
0.4
0.6
0.8
1.0
100
200
300
400
displacement

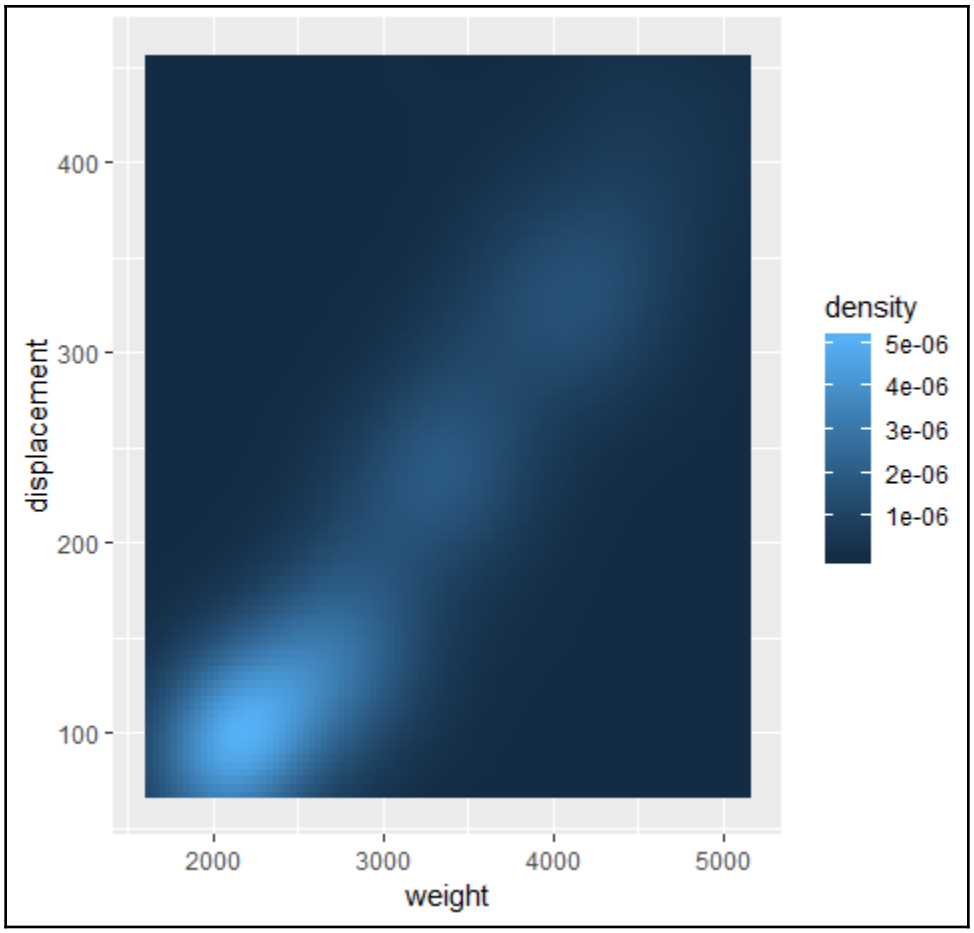

40
30
20
10
First
Second
Third

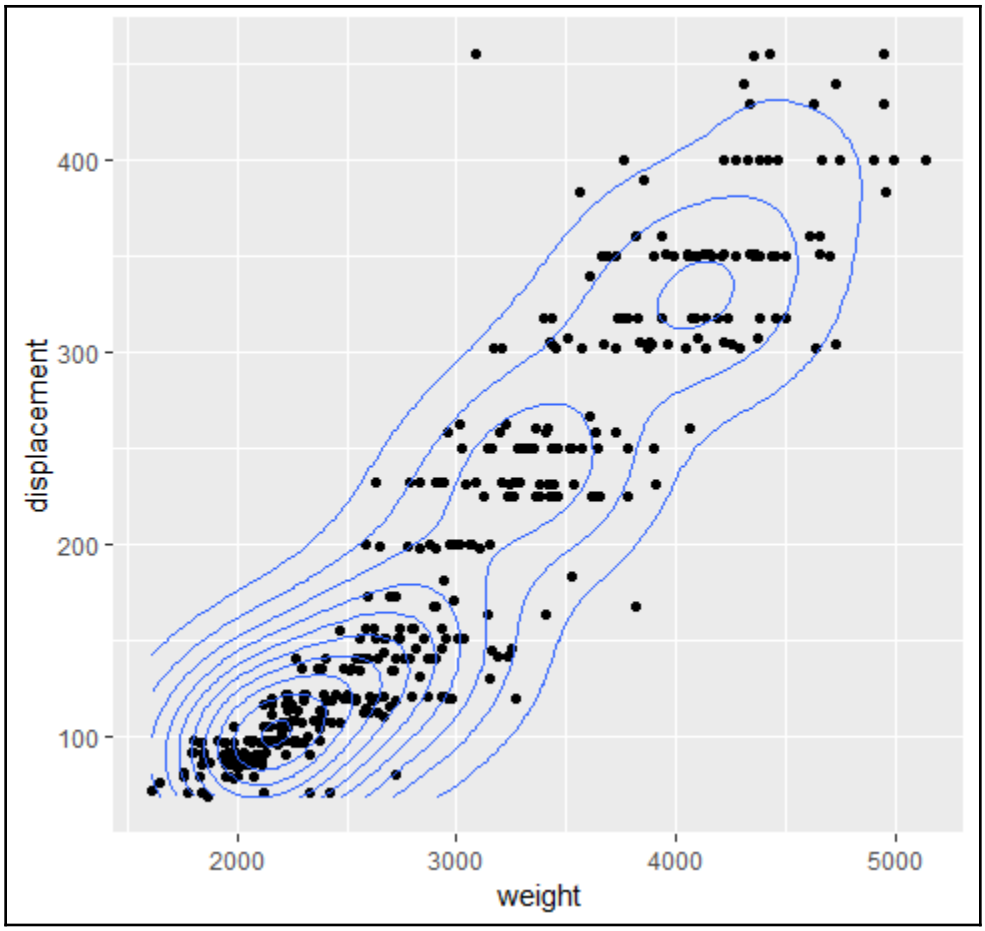

Scatter Plot for MTCars Weight Vs MPG
Miles per Gallon
10
20
30
40
1500
2500
3500
4500
Weight of Cars

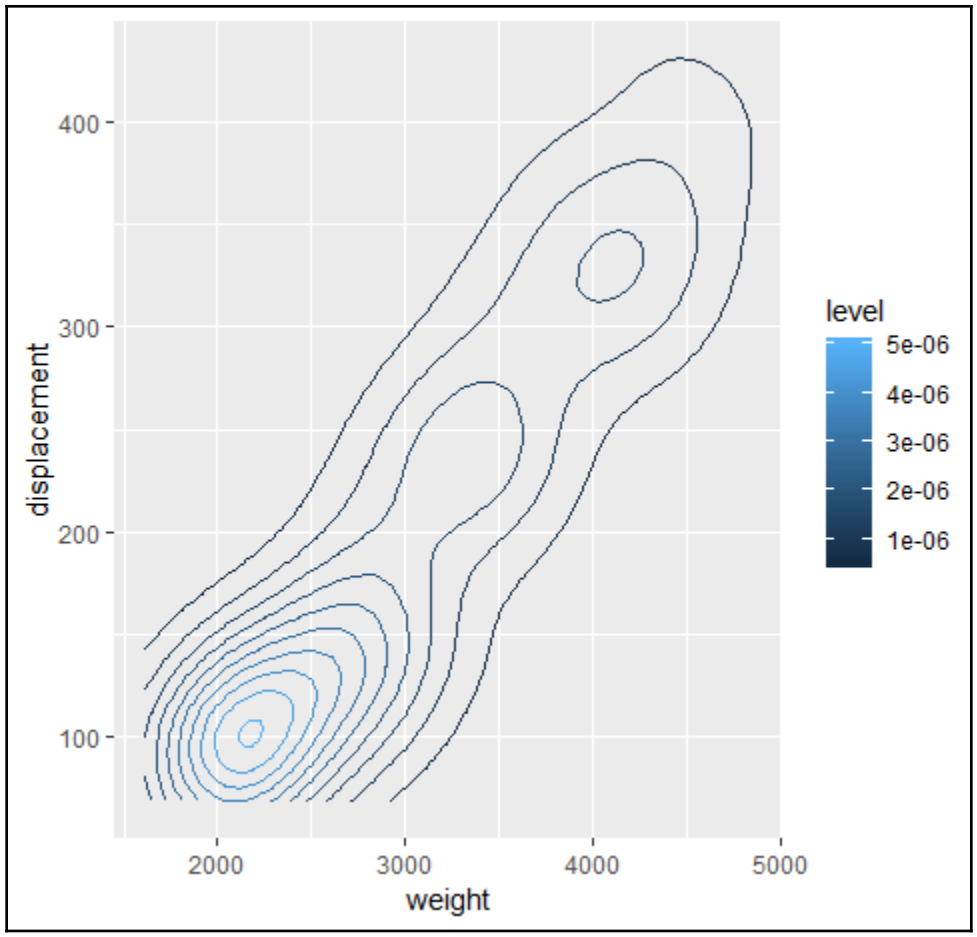

40
30
20
10
mpg
2000
3000
4000
5000
weight


model.year
75.0
76.0
77.0
78.0
n=249
n=70
n=79
1
2
3
origin


displacement
400
300
200
100
0
10
20
30
40
mpg


400
300
200
100
displacement
2000
3000
4000
5000
weight
density
5e-06
4e-06
3e-06
2e-06
1e-06


400
300
200
100
displacement
2000
3000
4000
5000
weight


400
300
200
100
displacement
2000
3000
4000
5000
weight
level
5e-06
4e-06
3e-06
2e-06
1e-06

Chapter 10: Handling Optimization and Regression Data Problems

GlassDataset

Filter

	Id	RI	Na	Mg	Al	Si	K	Ca	Ba	Fe	Type
1	1	1.52101	13.64	4.49	1.10	71.78	0.06	8.75	0.00	0.00	1
2	2	1.51761	13.89	3.60	1.36	72.73	0.48	7.83	0.00	0.00	1
3	3	1.51618	13.53	3.55	1.54	72.99	0.39	7.78	0.00	0.00	1
4	4	1.51766	13.21	3.69	1.29	72.61	0.57	8.22	0.00	0.00	1
5	5	1.51742	13.27	3.62	1.24	73.08	0.55	8.07	0.00	0.00	1
6	6	1.51596	12.79	3.61	1.62	72.97	0.64	8.07	0.00	0.26	1
7	7	1.51743	13.30	3.60	1.14	73.09	0.58	8.17	0.00	0.00	1
8	8	1.51756	13.15	3.61	1.05	73.24	0.57	8.24	0.00	0.00	1
9	9	1.51918	14.04	3.58	1.37	72.08	0.56	8.30	0.00	0.00	1
10	10	1.51755	13.00	3.60	1.36	72.99	0.57	8.40	0.00	0.11	1
11	11	1.51571	12.72	3.46	1.56	73.20	0.67	8.09	0.00	0.24	1

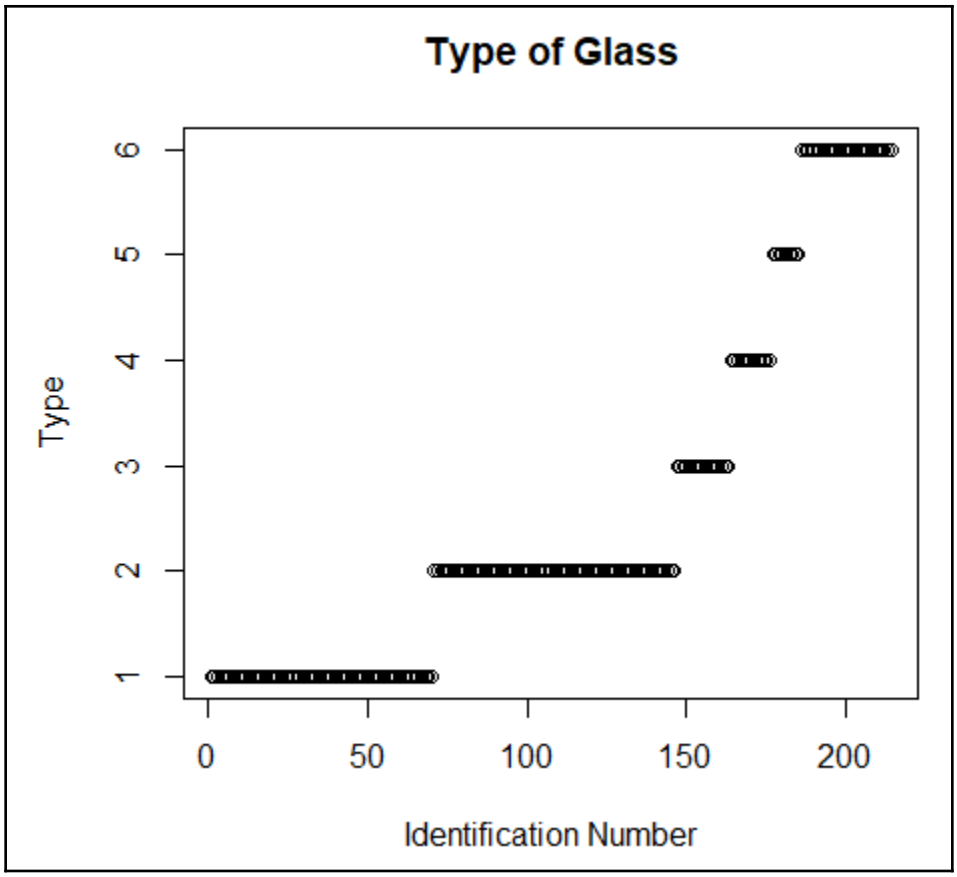
Showing 1 to 11 of 214 entries

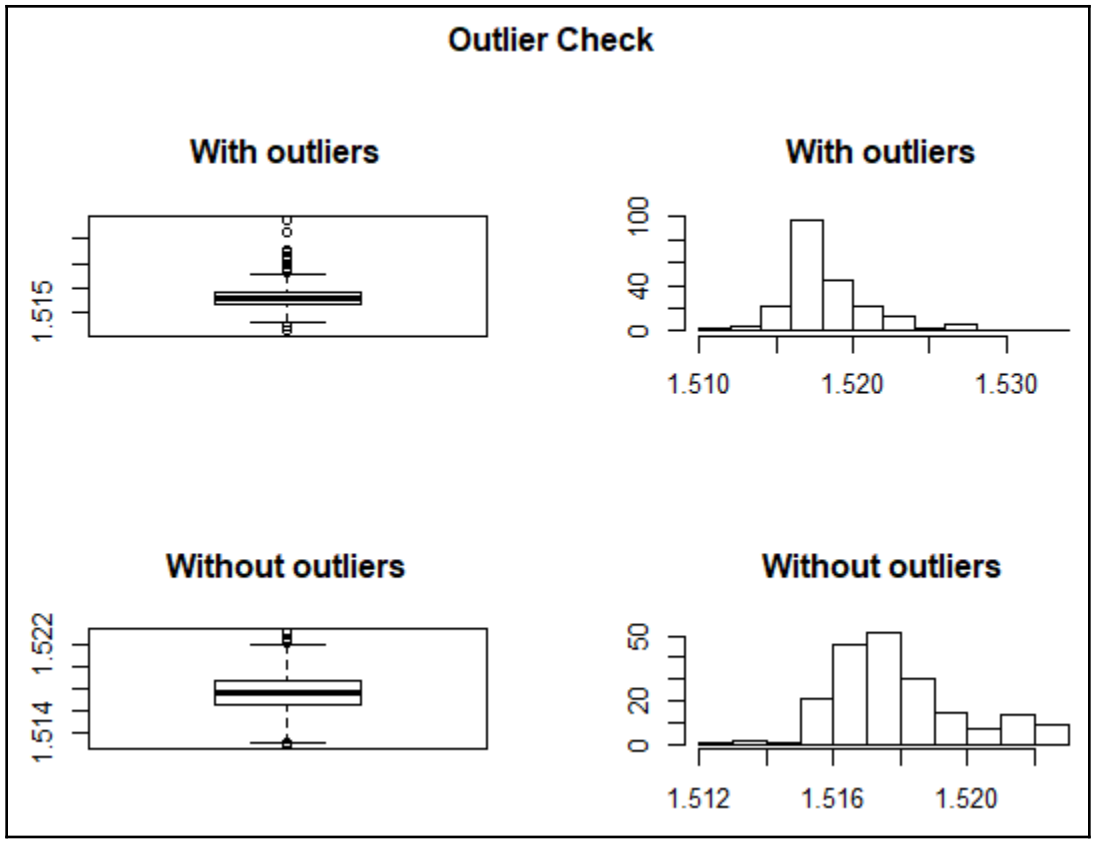
GlassDataset

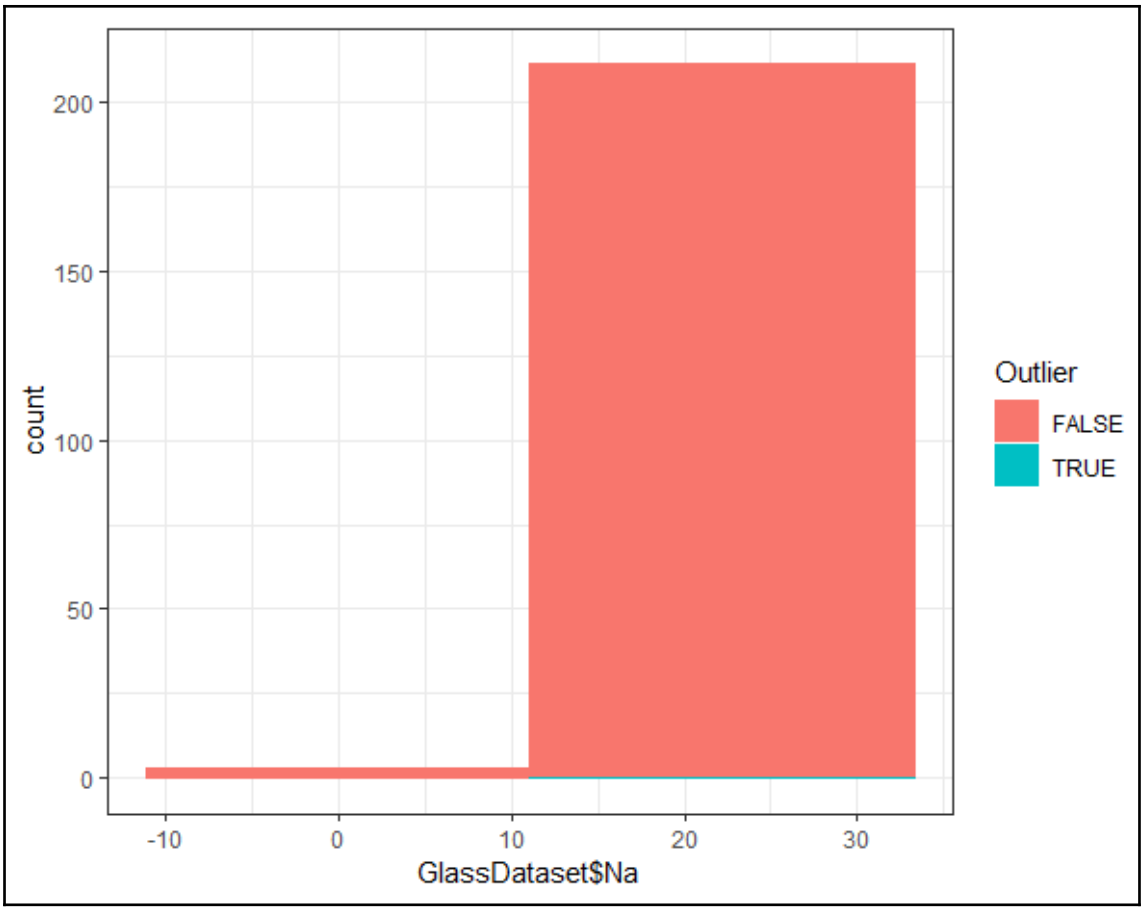
Filter

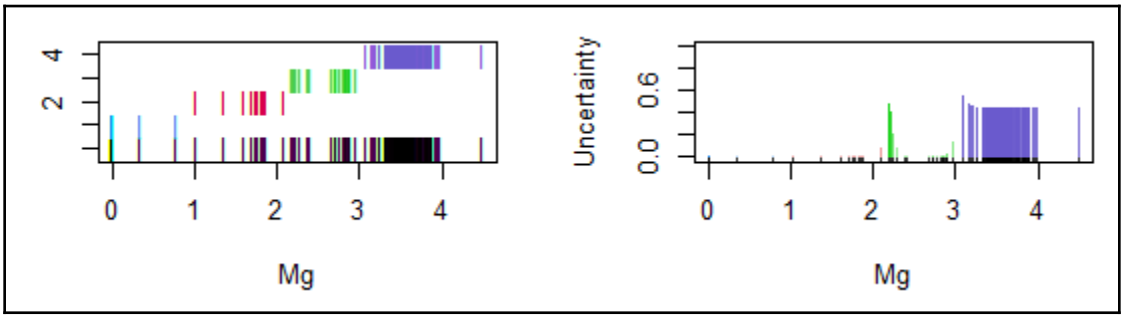
	Id	RI	Na	Mg	Al	Si	K	Ca	Ba	Fe	Type
1	1	1.52101	13.64	4.49	1.10	71.78	0.06	8.75	0.00	0.00	1
2	2	1.51761	13.89	3.60	1.36	72.73	0.48	7.83	0.00	0.00	1
3	3	1.51618	13.53	3.55	1.54	72.99	0.39	7.78	0.00	0.00	1
4	4	1.51766	13.21	3.69	1.29	72.61	0.57	8.22	0.00	0.00	1
5	5	1.51742	13.27	3.62	1.24	73.08	0.55	8.07	0.00	0.00	1
6	6	1.51596	12.79	3.61	1.62	72.97	0.64	8.07	0.00	0.26	1
7	7	1.51743	13.30	3.60	1.14	73.09	0.58	8.17	0.00	0.00	1
8	8	1.51756	13.15	3.61	1.05	73.24	0.57	8.24	0.00	0.00	1
9	9	1.51918	14.04	3.58	1.37	72.08	0.56	8.30	0.00	0.00	1
10	10	1.51755	13.00	3.60	1.36	72.99	0.57	8.40	0.00	0.11	1
11	11	1.51571	12.72	3.46	1.56	73.20	0.67	8.09	0.00	0.24	1

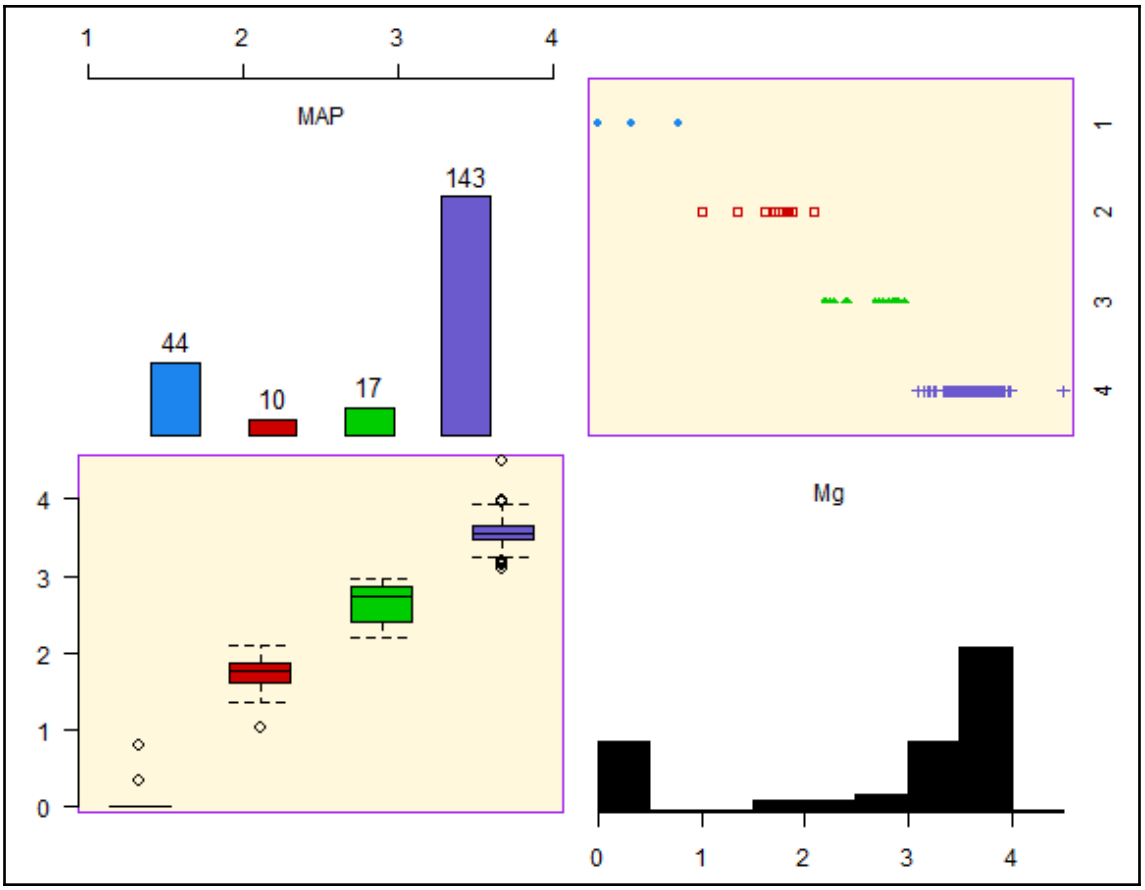
Showing 1 to 11 of 214 entries

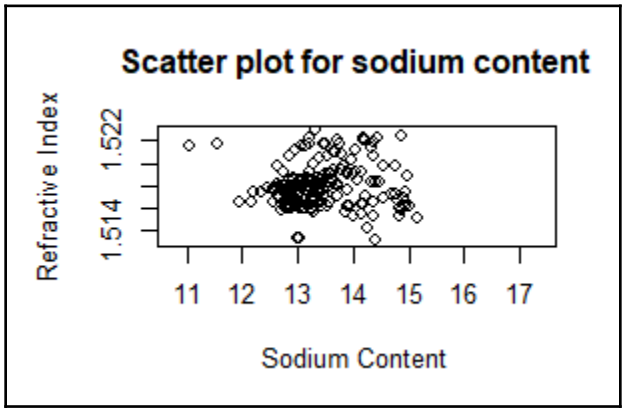

Type of Glass
Type
1
2
3
4
5
6
0
50
100
150
200
Identification Number

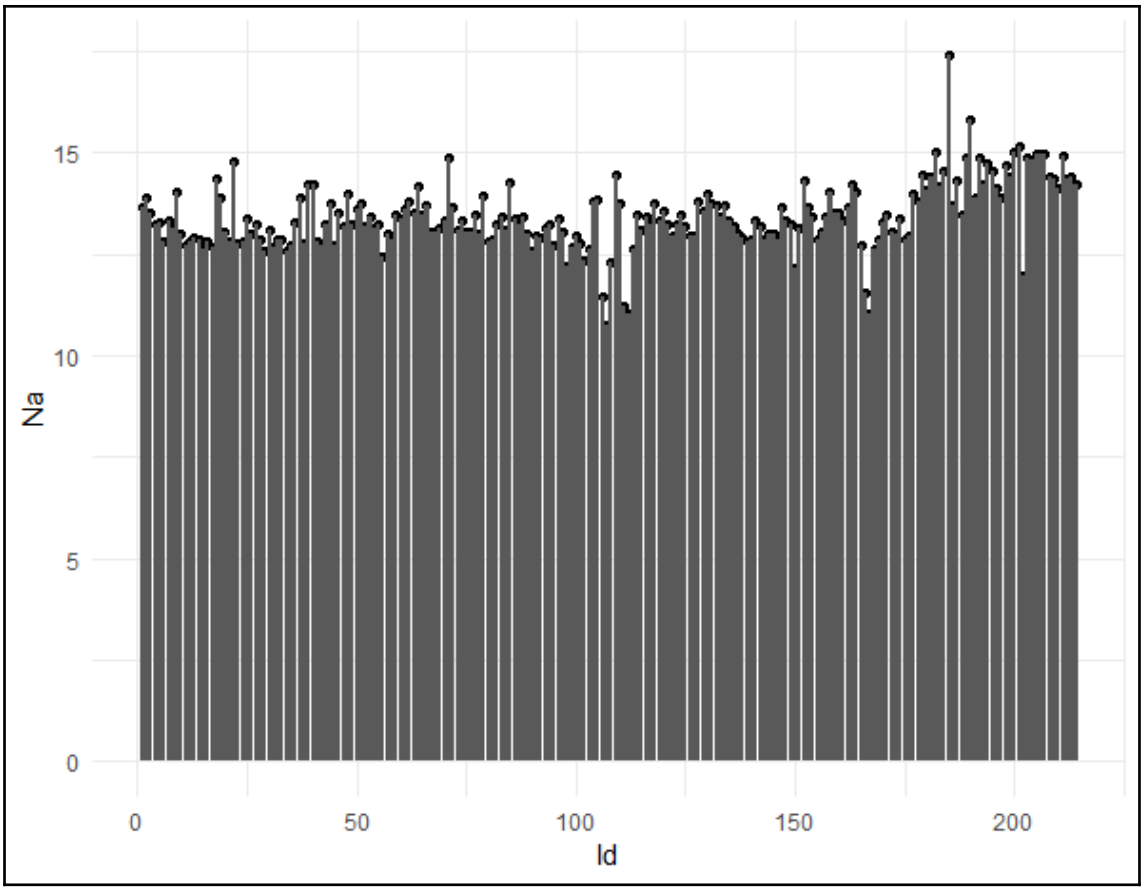

Outlier Check
With outliers
1.515
With outliers
0
40
100
1.510
1.520
1.530
Without outliers
1.514
1.522
Without outliers
0
20
50
1.512
1.516
1.520

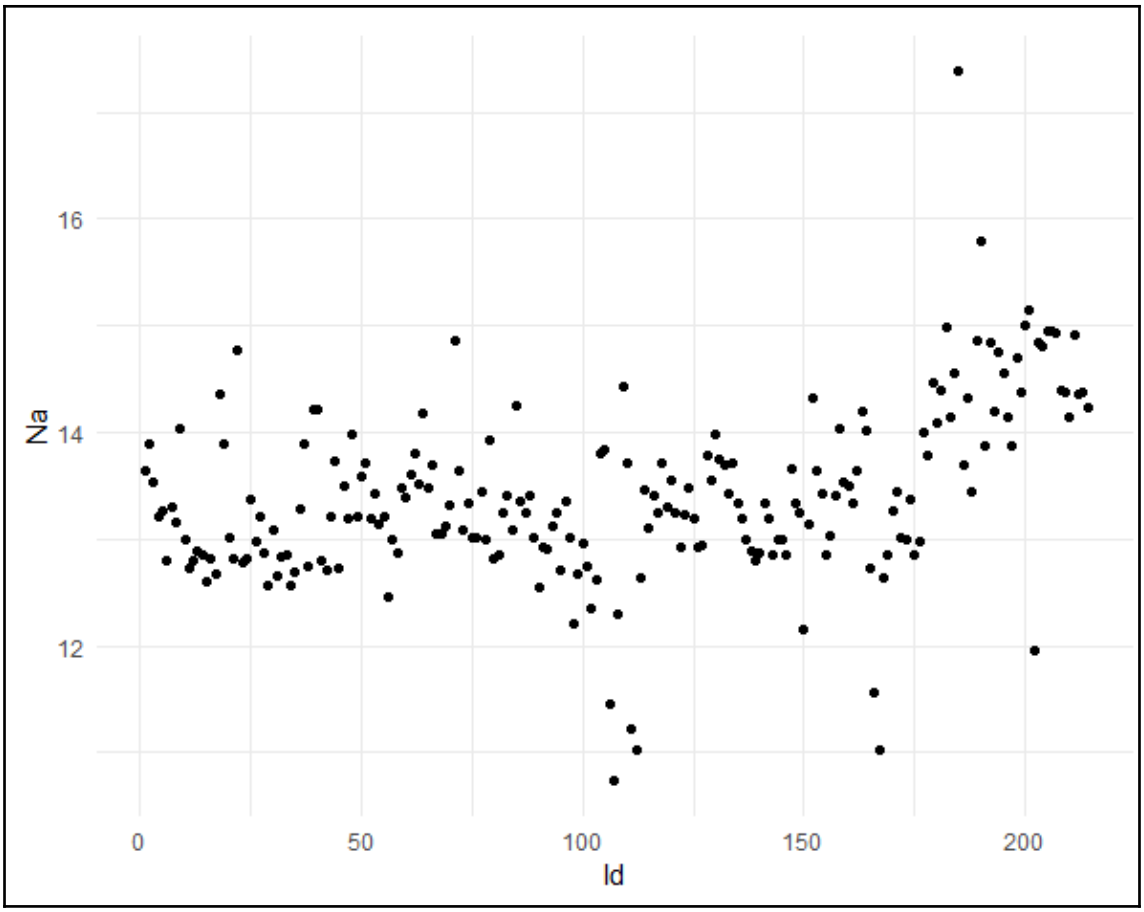

200
150
100
50
0
count
-10
0
10
20
30
GlassDataset$Na
Outlier
FALSE
TRUE

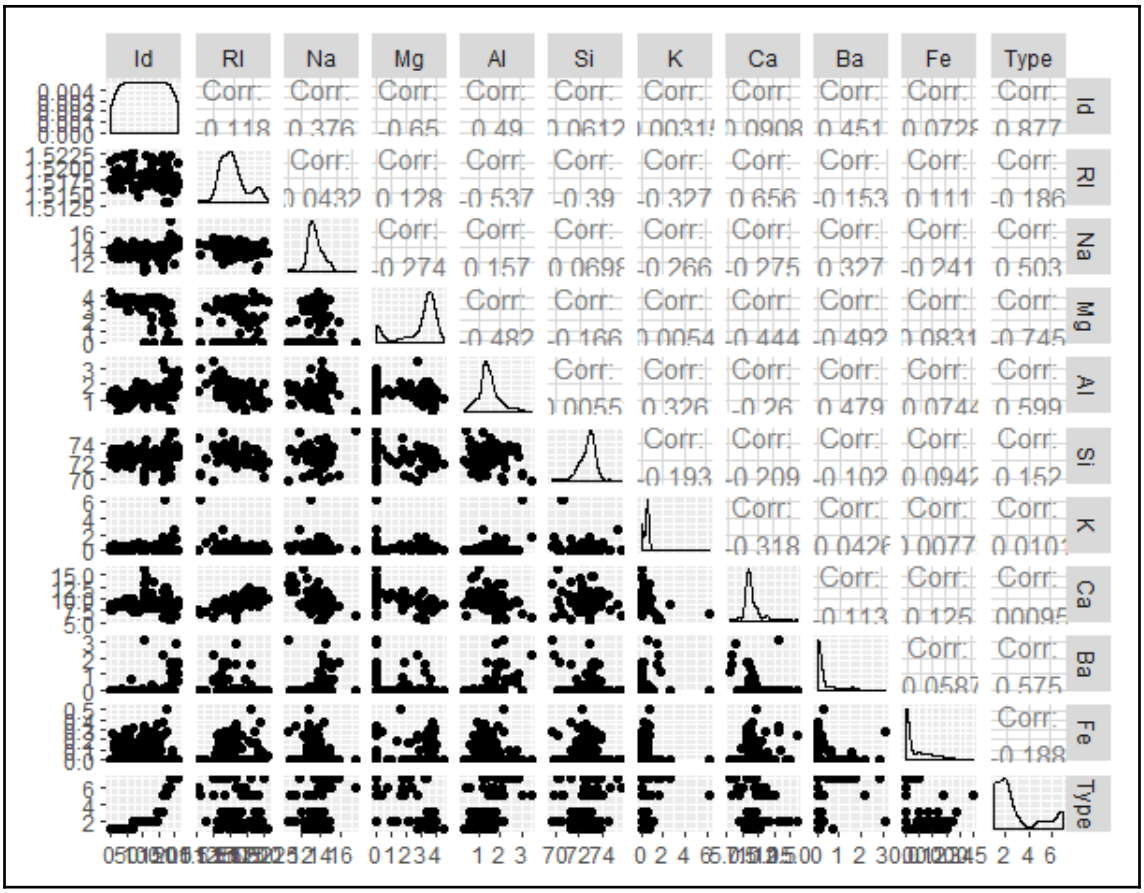

2
4
0
1
2
3
4
Mg
Uncertainty
0.0
0.6
0
1
2
3
4
Mg

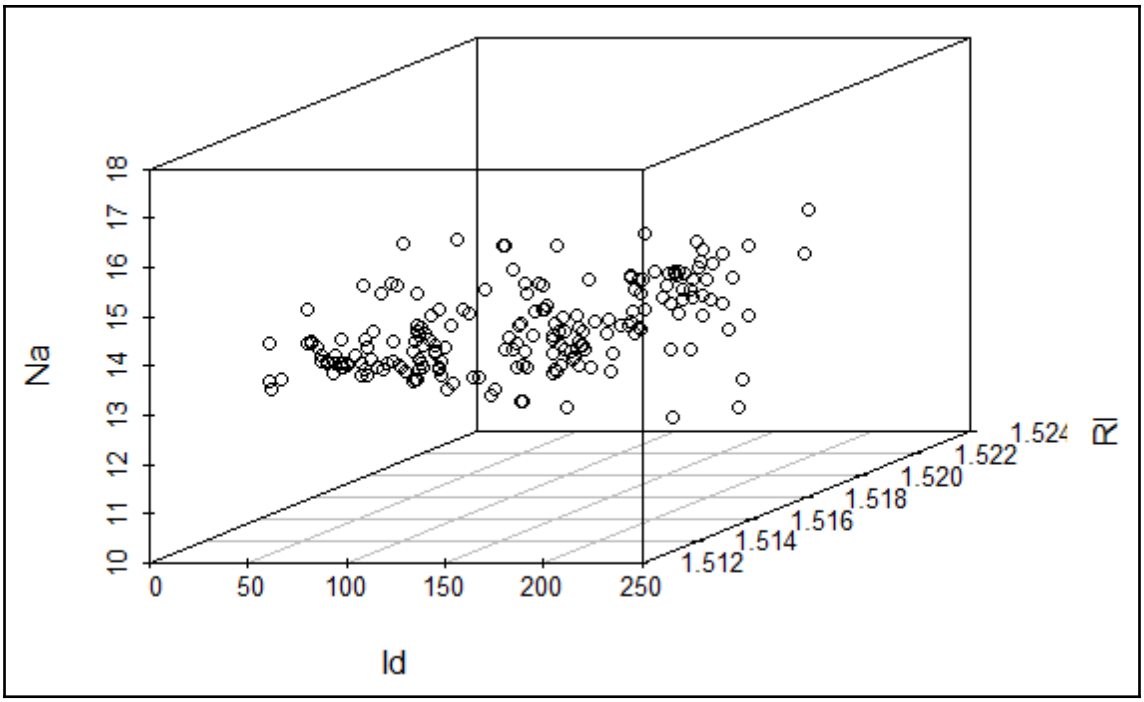

1
2
3
4
MAP
143
44
10
17
Mg
0
1
2
3
4

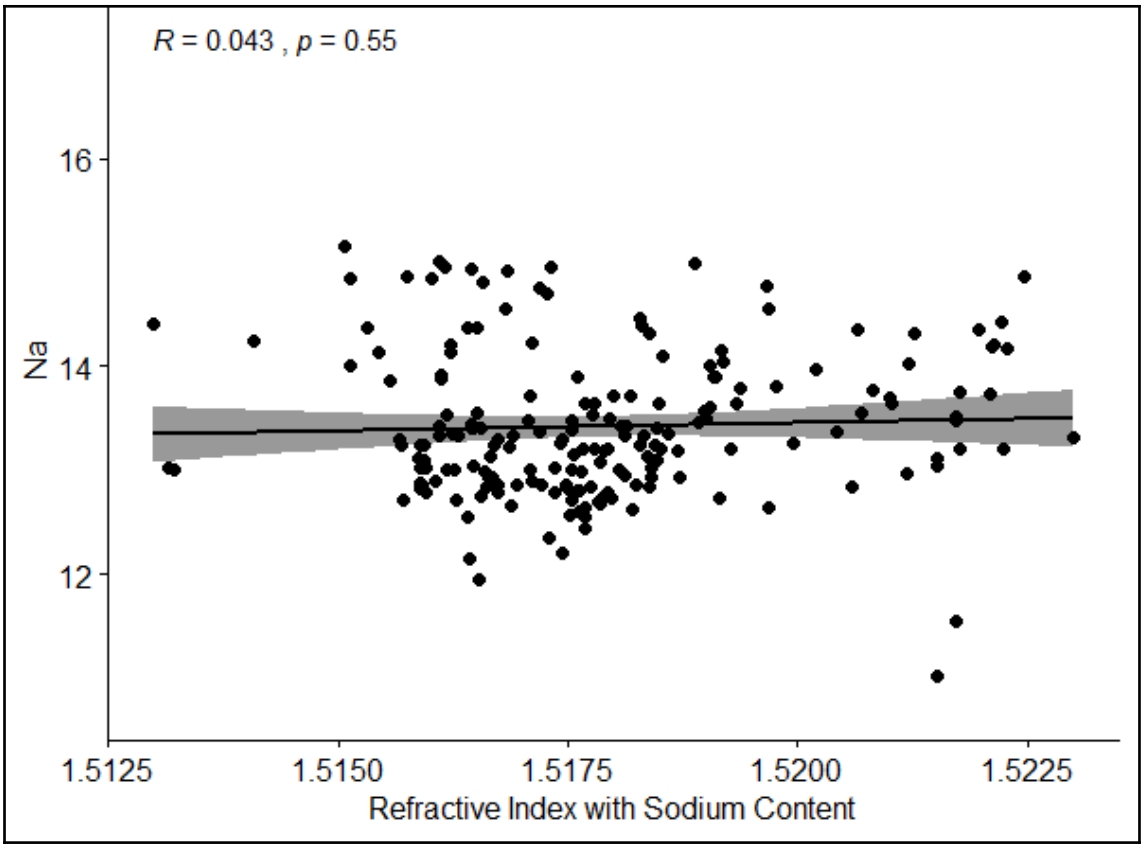

Scatter plot for sodium content
Refractive Index
1.514
1.522
11
12
13
14
15
16
17
Sodium Content

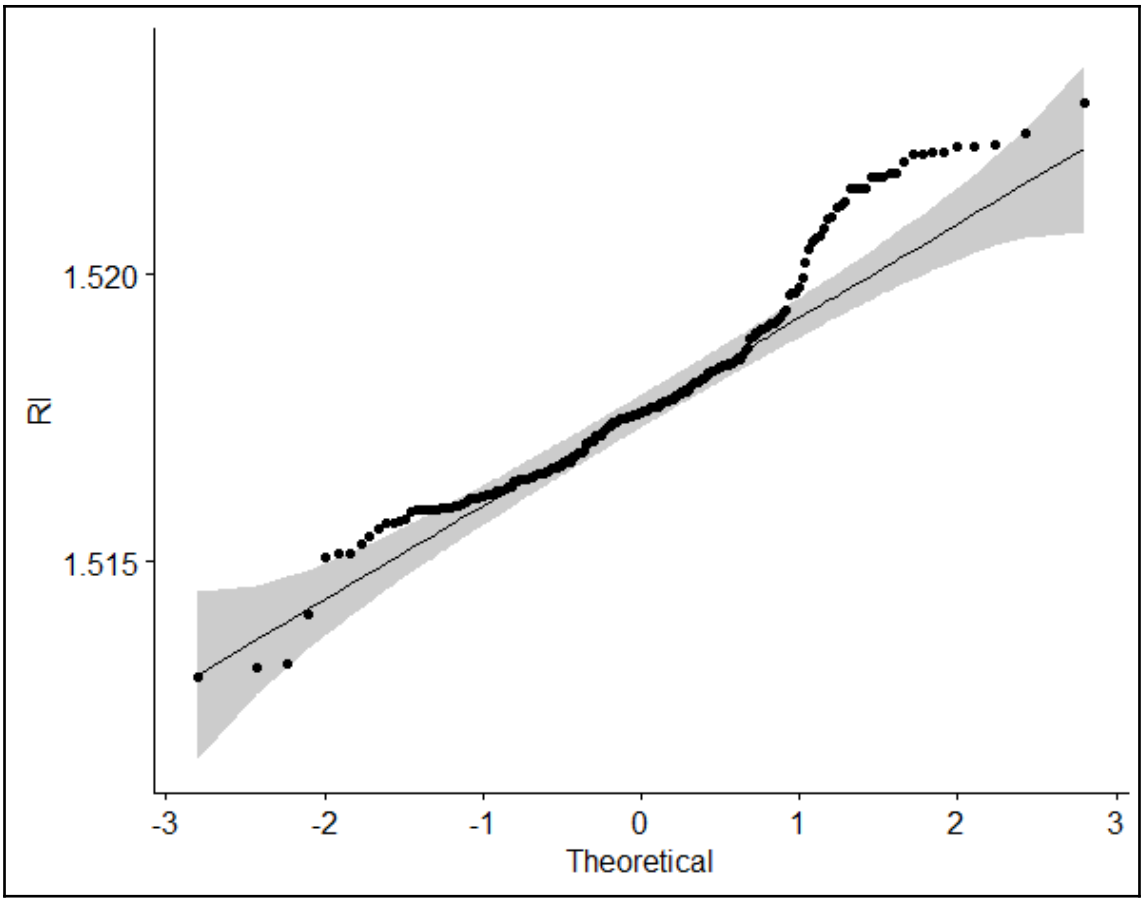

15
10
5
0
Na
0
50
100
150
200
Id

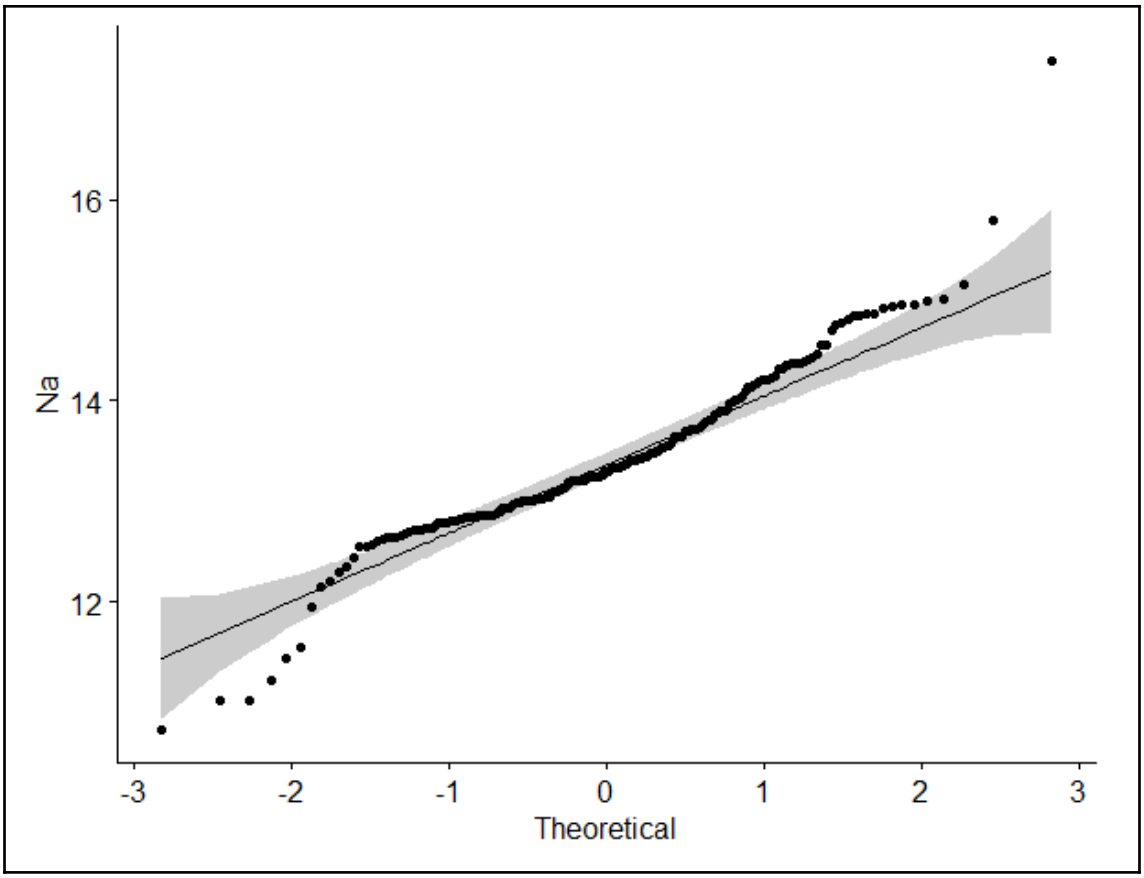

16
14
12
Na
0
50
100
150
200
Id

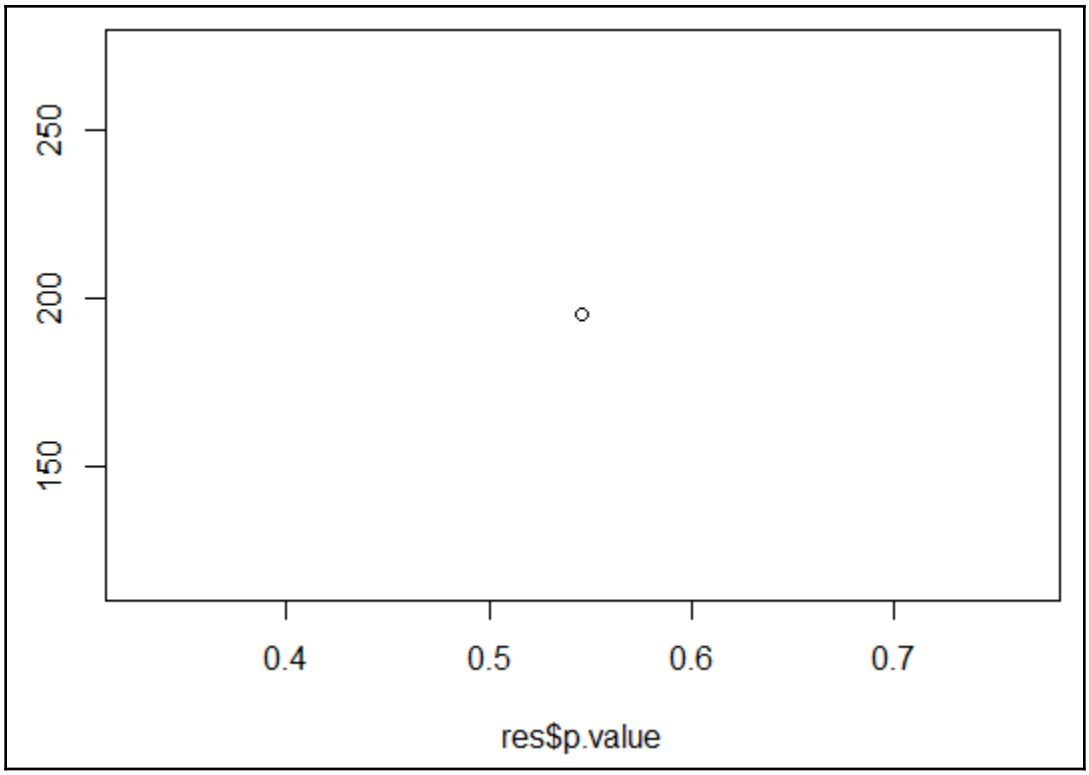

Id
RI
Na
Mg
Al
Si
K
Ca
Ba
Fe
Type
Corr:
-0.118
0.376
-0.65
0.49
0.0612
0.0908
0.451
0.877
0.0432
0.128
-0.537
-0.39
-0.327
0.656
-0.153
0.111
-0.186
-0.274
0.157
0.0698
-0.266
-0.275
0.327
-0.241
0.503
-0.482
-0.166
0.0054
-0.444
-0.492
0.0831
-0.745
0.0055
0.326
-0.26
0.479
0.0744
0.599
-0.193
-0.209
-0.102
0.0942
0.152
-0.318
0.0426
0.0077
0.0101
-0.113
0.125
0.0095
0.0587
0.575
-0.188


Na
10
11
12
13
14
15
16
17
18
Id
0
50
100
150
200
250
RI
1.512
1.514
1.516
1.518
1.520
1.522
1.524


R = 0.043 , p = 0.55
16
14
12
Na
1.5125
1.5150
1.5175
1.5200
1.5225
Refractive Index with Sodium Content


RI
1.520
1.515
-3
-2
-1
0
1
2
3
Theoretical


16
14
12
Na
-3
-2
-1
0
1
2
3
Theoretical


250
200
150
0.4
0.5
0.6
0.7
res$p.value

Chapter 11: Next Steps

plot {graphics} R Documentation

Generic X-Y Plotting

Description

Generic function for plotting of R objects. For more details about the graphical parameter arguments, see `par`.

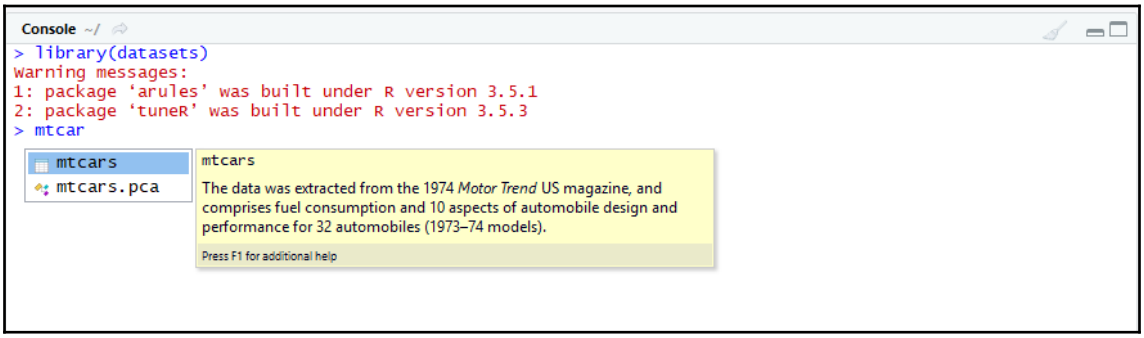
For simple scatter plots, `plot.default` will be used. However, there are `plot` methods for many R objects, including `function`s, `data.frame`s, `density` objects, etc. Use `methods(plot)` and the documentation for these.

Usage

```
plot(x, y, ...)
```

Arguments

`x`	the coordinates of points in the plot. Alternatively, a single plotting structure, function or *any R object with a `plot` method* can be provided
`y`	the y coordinates of points in the plot, *optional* if `x` is an appropriate structure.



Graphics Bundle Ends Here

Index