Gergely Daróczi

Looong report

Fri Jul 20 12:17:28 2012

Contents

I have written the below report in 10 mins :)

Dataset

Here I will do a pretty fast report on `mtcars` which is:

	mpg	cyl	disp	hp	drat
Mazda RX4	21.0	6	160.0	110	3.90
Mazda RX4 Wag	21.0	6	160.0	110	3.90
Datsun 710	22.8	4	108.0	93	3.85
Hornet 4 Drive	21.4	6	258.0	110	3.08
Hornet Sportabout	18.7	8	360.0	175	3.15
Valiant	18.1	6	225.0	105	2.76
Duster 360	14.3	8	360.0	245	3.21
Merc 240D	24.4	4	146.7	62	3.69
Merc 230	22.8	4	140.8	95	3.92
Merc 280	19.2	6	167.6	123	3.92
Merc 280C	17.8	6	167.6	123	3.92
Merc 450SE	16.4	8	275.8	180	3.07
Merc 450SL	17.3	8	275.8	180	3.07
Merc 450SLC	15.2	8	275.8	180	3.07
Cadillac Fleetwood	10.4	8	472.0	205	2.93
Lincoln Continental	10.4	8	460.0	215	3.00
Chrysler Imperial	14.7	8	440.0	230	3.23
Fiat 128	32.4	4	78.7	66	4.08
Honda Civic	30.4	4	75.7	52	4.93
Toyota Corolla	33.9	4	71.1	65	4.22
Toyota Corona	21.5	4	120.1	97	3.70
Dodge Challenger	15.5	8	318.0	150	2.76
AMC Javelin	15.2	8	304.0	150	3.15
Camaro Z28	13.3	8	350.0	245	3.73
Pontiac Firebird	19.2	8	400.0	175	3.08
Fiat X1-9	27.3	4	79.0	66	4.08
Porsche 914-2	26.0	4	120.3	91	4.43
Lotus Europa	30.4	4	95.1	113	3.77
Ford Pantera L	15.8	8	351.0	264	4.22
Ferrari Dino	19.7	6	145.0	175	3.62
Maserati Bora	15.0	8	301.0	335	3.54
Volvo 142E	21.4	4	121.0	109	4.11

	wt	qsec	vs	am
Mazda RX4	2.620	16.46	0	1
Mazda RX4 Wag	2.875	17.02	0	1
Datsun 710	2.320	18.61	1	1
Hornet 4 Drive	3.215	19.44	1	0
Hornet Sportabout	3.440	17.02	0	0
Valiant	3.460	20.22	1	0
Duster 360	3.570	15.84	0	0
Merc 240D	3.190	20.00	1	0
Merc 230	3.150	22.90	1	0
Merc 280	3.440	18.30	1	0
Merc 280C	3.440	18.90	1	0
Merc 450SE	4.070	17.40	0	0
Merc 450SL	3.730	17.60	0	0
Merc 450SLC	3.780	18.00	0	0
Cadillac Fleetwood	5.250	17.98	0	0
Lincoln Continental	5.424	17.82	0	0
Chrysler Imperial	5.345	17.42	0	0
Fiat 128	2.200	19.47	1	1
Honda Civic	1.615	18.52	1	1
Toyota Corolla	1.835	19.90	1	1
Toyota Corona	2.465	20.01	1	0
Dodge Challenger	3.520	16.87	0	0
AMC Javelin	3.435	17.30	0	0
Camaro Z28	3.840	15.41	0	0
Pontiac Firebird	3.845	17.05	0	0
Fiat X1-9	1.935	18.90	1	1
Porsche 914-2	2.140	16.70	0	1
Lotus Europa	1.513	16.90	1	1
Ford Pantera L	3.170	14.50	0	1
Ferrari Dino	2.770	15.50	0	1
Maserati Bora	3.570	14.60	0	1
Volvo 142E	2.780	18.60	1	1

	gear	carb
Mazda RX4	4	4
Mazda RX4 Wag	4	4
Datsun 710	4	1
Hornet 4 Drive	3	1
Hornet Sportabout	3	2
Valiant	3	1
Duster 360	3	4
Merc 240D	4	2
Merc 230	4	2
Merc 280	4	4
Merc 280C	4	4
Merc 450SE	3	3
Merc 450SL	3	3
Merc 450SLC	3	3
Cadillac Fleetwood	3	4
Lincoln Continental	3	4
Chrysler Imperial	3	4
Fiat 128	4	1
Honda Civic	4	2
Toyota Corolla	4	1
Toyota Corona	3	1
Dodge Challenger	3	2
AMC Javelin	3	2
Camaro Z28	3	4
Pontiac Firebird	3	2
Fiat X1-9	4	1
Porsche 914-2	5	2
Lotus Europa	5	2
Ford Pantera L	5	4
Ferrari Dino	5	6
Maserati Bora	5	8
Volvo 142E	4	2

Descriptives

	Average	Median	Standard.deviation	Variance
mpg	20.0906	19.200	6.0269	3.632e+01
cyl	6.1875	6.000	1.7859	3.190e+00
disp	230.7219	196.300	123.9387	1.536e+04
hp	146.6875	123.000	68.5629	4.701e+03
drat	3.5966	3.695	0.5347	2.859e-01
wt	3.2172	3.325	0.9785	9.574e-01
qsec	17.8487	17.710	1.7869	3.193e+00
vs	0.4375	0.000	0.5040	2.540e-01
am	0.4062	0.000	0.4990	2.490e-01
gear	3.6875	4.000	0.7378	5.444e-01
carb	2.8125	2.000	1.6152	2.609e+00

In details

mpg

We found the folloing values here:

21.0, *21.0*, *22.8*, *21.4*, *18.7*, *18.1*, *14.3*, *24.4*, *22.8*, *19.2*, *17.8*, *16.4*, *17.3*, *15.2*, *10.4*, *10.4*, *14.7*, *32.4*, *30.4*, *33.9*, *21.5*, *15.5*, *15.2*, *13.3*, *19.2*, *27.3*, *26.0*, *30.4*, *15.8*, *19.7*, *15.0* and *21.4*

The mean of mpg is *20.09* while the standard deviation is: *6.027*. The most frequent value in mpg is 10.4, but let us check out the frequency table too:

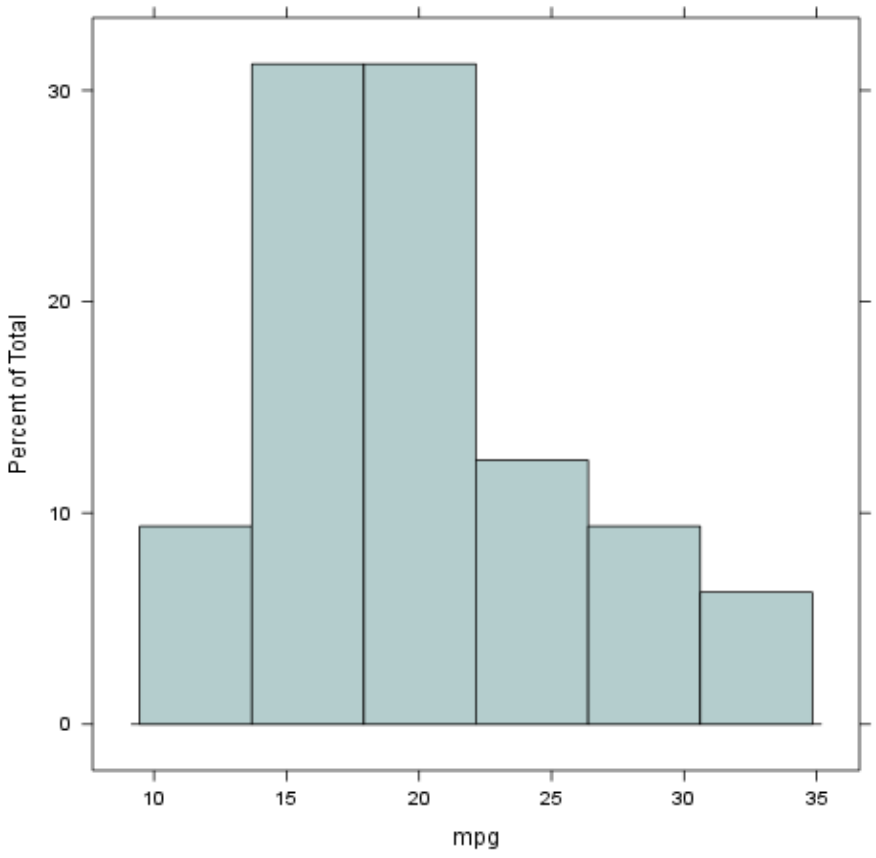
10.4	13.3	14.3	14.7	15	15.2	15.5	15.8
2	1	1	1	1	2	1	1

16.4	17.3	17.8	18.1	18.7	19.2	19.7	21
1	1	1	1	1	2	1	2

21.4	21.5	22.8	24.4	26	27.3	30.4
2	1	2	1	1	1	2

32.4	33.9
1	1

Tables are boring, let us show the same with a `histogram`:



cyl

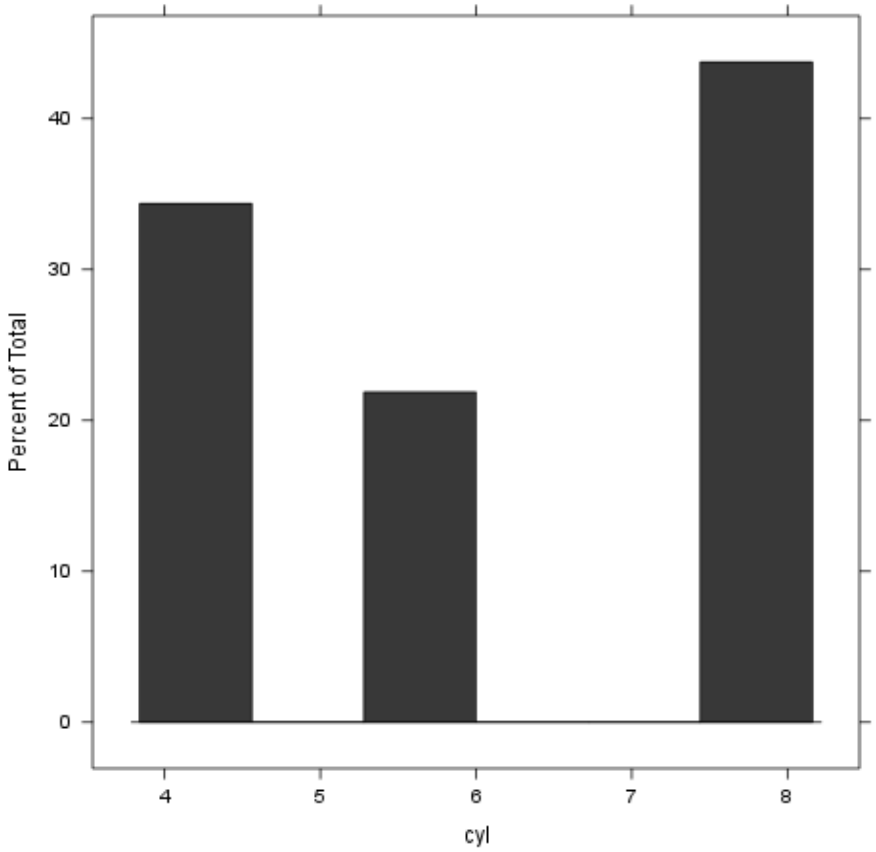
We found the folloing values here:

6, *6*, *4*, *6*, *8*, *6*, *8*, *4*, *4*, *6*, *6*, *8*, *8*, *8*, *8*, *8*, *8*, *4*, *4*, *4*, *4*, *8*, *8*, *8*, *8*, *4*, *4*, *4*, *8*, *6*, *8* and *4*

The mean of cyl is *6.188* while the standard deviation is: *1.786*. The most frequent value in cyl is 8, but let us check out the frequency table too:

4	6	8
11	7	14

Tables are boring, let us show the same with a `histogram`:



disp

We found the folloing values here:

160.0, *160.0*, *108.0*, *258.0*, *360.0*, *225.0*, *360.0*, *146.7*, *140.8*, *167.6*, *167.6*, *275.8*, *275.8*, *275.8*, *472.0*, *460.0*, *440.0*, *78.7*, *75.7*, *71.1*, *120.1*, *318.0*, *304.0*, *350.0*, *400.0*, *79.0*, *120.3*, *95.1*, *351.0*, *145.0*, *301.0* and *121.0*

The mean of disp is *230.7* while the standard deviation is: *123.9*. The most frequent value in disp is 275.8, but let us check out the frequency table too:

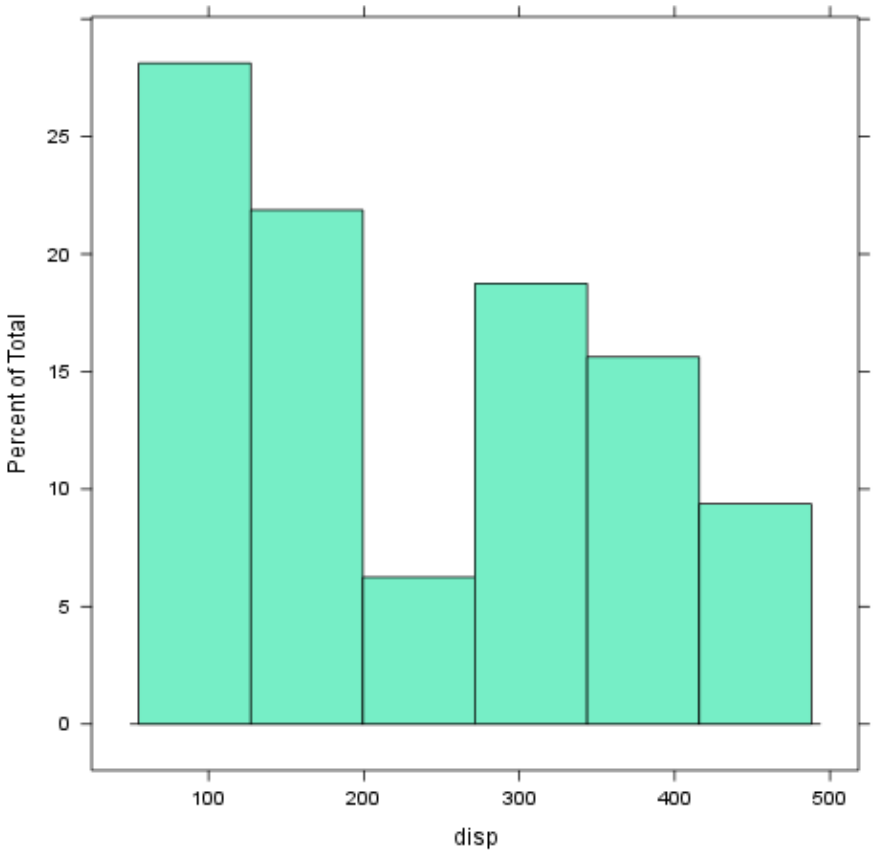
71.1	75.7	78.7	79	95.1	108	120.1	120.3
1	1	1	1	1	1	1	1

121	140.8	145	146.7	160	167.6	225	258
1	1	1	1	2	2	1	1

275.8	301	304	318	350	351	360	400
3	1	1	1	1	1	2	1

440	460	472
1	1	1

Tables are boring, let us show the same with a `histogram`:



hp

We found the folloing values here:

110, *110*, *93*, *110*, *175*, *105*, *245*, *62*, *95*, *123*, *123*, *180*, *180*, *180*, *205*, *215*, *230*, *66*, *52*, *65*, *97*, *150*, *150*, *245*, *175*, *66*, *91*, *113*, *264*, *175*, *335* and *109*

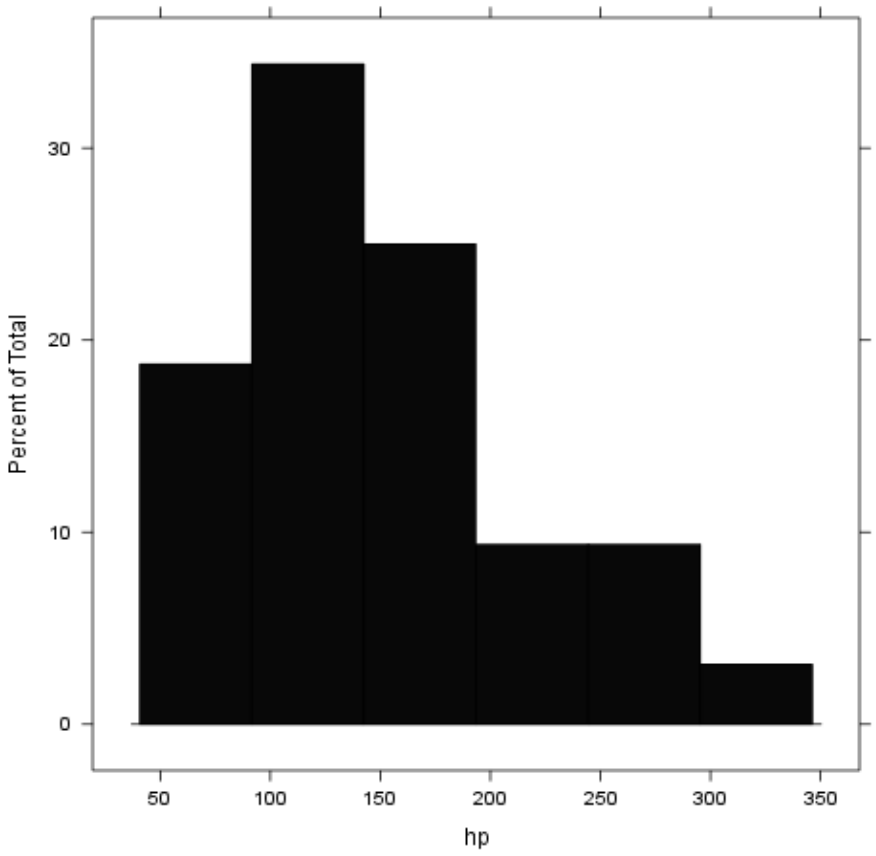
The mean of hp is *146.7* while the standard deviation is: *68.56*. The most frequent value in hp is 110, but let us check out the frequency table too:

52	62	65	66	91	93	95	97	105
1	1	1	2	1	1	1	1	1

109	110	113	123	150	175	180	205
1	3	1	2	2	3	3	1

215	230	245	264	335
1	1	2	1	1

Tables are boring, let us show the same with a `histogram`:



drat

We found the folloing values here:

3.90, *3.90*, *3.85*, *3.08*, *3.15*, *2.76*, *3.21*, *3.69*, *3.92*, *3.92*, *3.92*, *3.07*, *3.07*, *3.07*, *2.93*, *3.00*, *3.23*, *4.08*, *4.93*, *4.22*, *3.70*, *2.76*, *3.15*, *3.73*, *3.08*, *4.08*, *4.43*, *3.77*, *4.22*, *3.62*, *3.54* and *4.11*

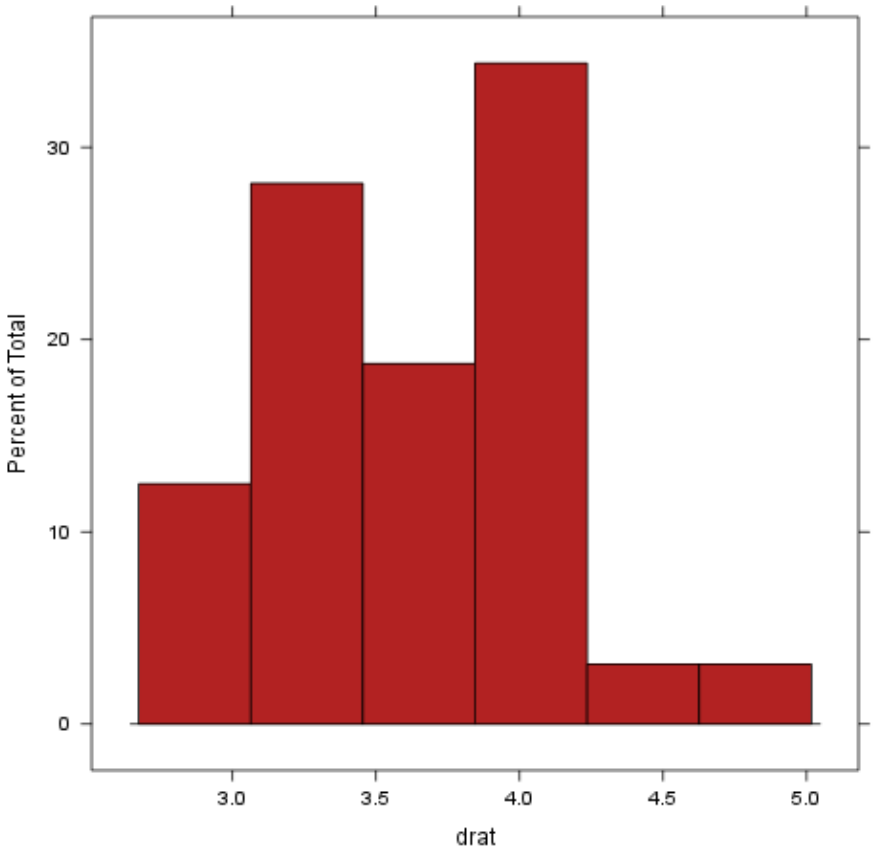
The mean of drat is *3.597* while the standard deviation is: *0.5347*. The most frequent value in drat is 3.07, but let us check out the frequency table too:

2.76	2.93	3	3.07	3.08	3.15	3.21	3.23
2	1	1	3	2	2	1	1

3.54	3.62	3.69	3.7	3.73	3.77	3.85	3.9
1	1	1	1	1	1	1	2

3.92	4.08	4.11	4.22	4.43	4.93
3	2	1	2	1	1

Tables are boring, let us show the same with a `histogram`:



wt

We found the folloing values here:

2.620, *2.875*, *2.320*, *3.215*, *3.440*, *3.460*, *3.570*, *3.190*, *3.150*, *3.440*, *3.440*, *4.070*, *3.730*, *3.780*, *5.250*, *5.424*, *5.345*, *2.200*, *1.615*, *1.835*, *2.465*, *3.520*, *3.435*, *3.840*, *3.845*, *1.935*, *2.140*, *1.513*, *3.170*, *2.770*, *3.570* and *2.780*

The mean of wt is *3.217* while the standard deviation is: *0.9785*. The most frequent value in wt is 3.44, but let us check out the frequency table too:

1.513	1.615	1.835	1.935	2.14	2.2	2.32
1	1	1	1	1	1	1

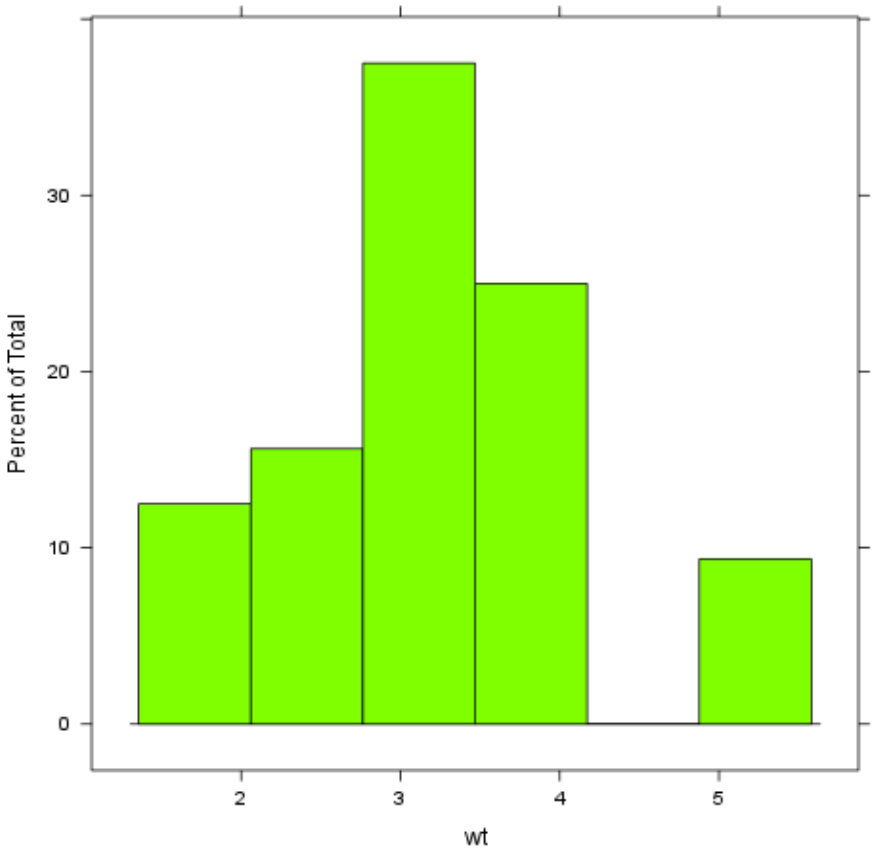
2.465	2.62	2.77	2.78	2.875	3.15	3.17
1	1	1	1	1	1	1

3.19	3.215	3.435	3.44	3.46	3.52	3.57
1	1	1	3	1	1	2

3.73	3.78	3.84	3.845	4.07	5.25
1	1	1	1	1	1

5.345	5.424
1	1

Tables are boring, let us show the same with a `histogram`:



qsec

We found the folloing values here:

16.46, *17.02*, *18.61*, *19.44*, *17.02*, *20.22*, *15.84*, *20.00*, *22.90*, *18.30*, *18.90*, *17.40*, *17.60*, *18.00*, *17.98*, *17.82*, *17.42*, *19.47*, *18.52*, *19.90*, *20.01*, *16.87*, *17.30*, *15.41*, *17.05*, *18.90*, *16.70*, *16.90*, *14.50*, *15.50*, *14.60* and *18.60*

The mean of qsec is *17.85* while the standard deviation is: *1.787*. The most frequent value in qsec is 17.02, but let us check out the frequency table too:

14.5	14.6	15.41	15.5	15.84	16.46	16.7
1	1	1	1	1	1	1

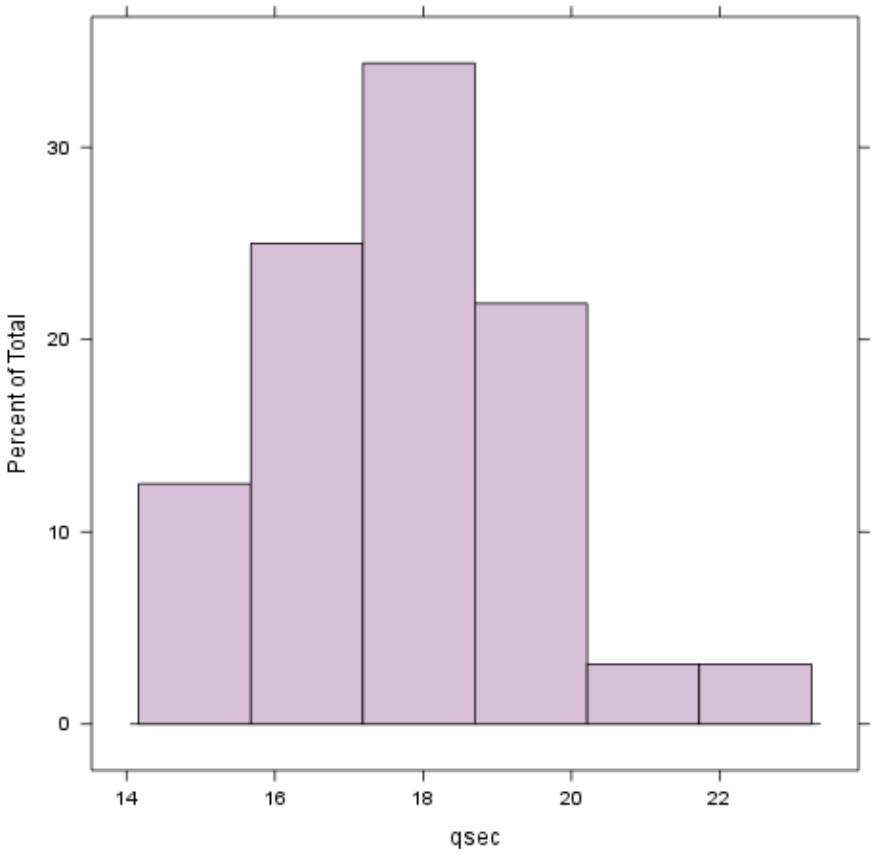
16.87	16.9	17.02	17.05	17.3	17.4	17.42
1	1	2	1	1	1	1

17.6	17.82	17.98	18	18.3	18.52	18.6
1	1	1	1	1	1	1

18.61	18.9	19.44	19.47	19.9	20	20.01
1	2	1	1	1	1	1

20.22	22.9
1	1

Tables are boring, let us show the same with a `histogram`:



vs

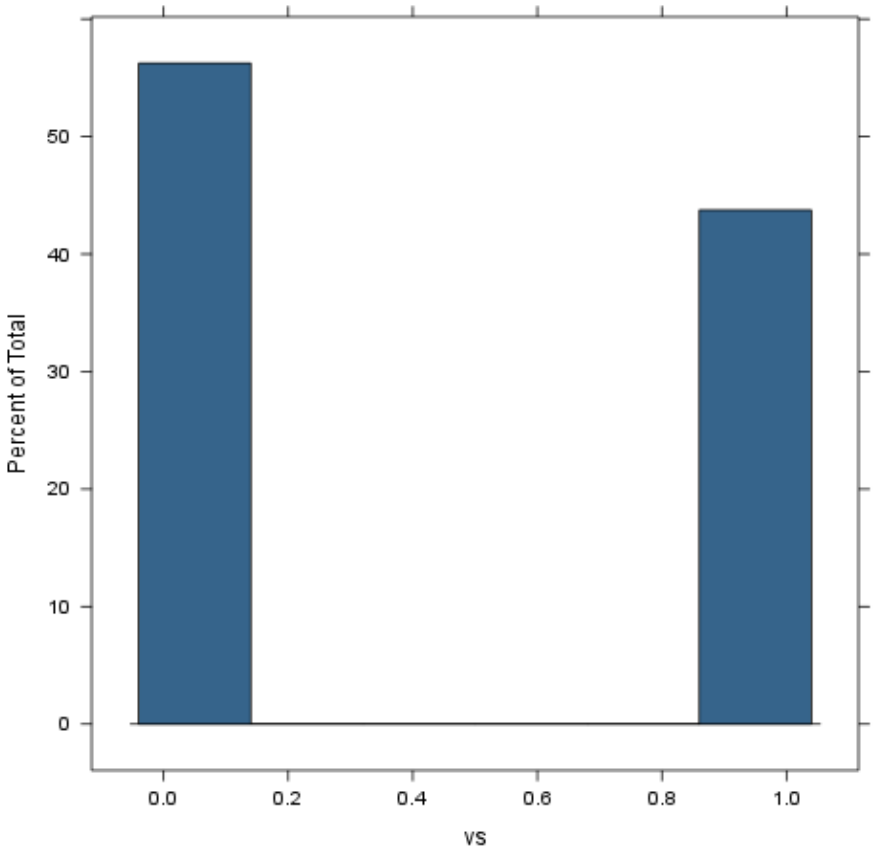
We found the folloing values here:

0, *0*, *1*, *1*, *0*, *1*, *0*, *1*, *1*, *1*, *1*, *0*, *0*, *0*, *0*, *0*, *0*, *1*, *1*, *1*, *1*, *0*, *0*, *0*, *0*, *1*, *0*, *1*, *0*, *0*, *0* and *1*

The mean of vs is *0.4375* while the standard deviation is: *0.504*. The most frequent value in vs is 0, but let us check out the frequency table too:

0	1
18	14

Tables are boring, let us show the same with a `histogram`:



am

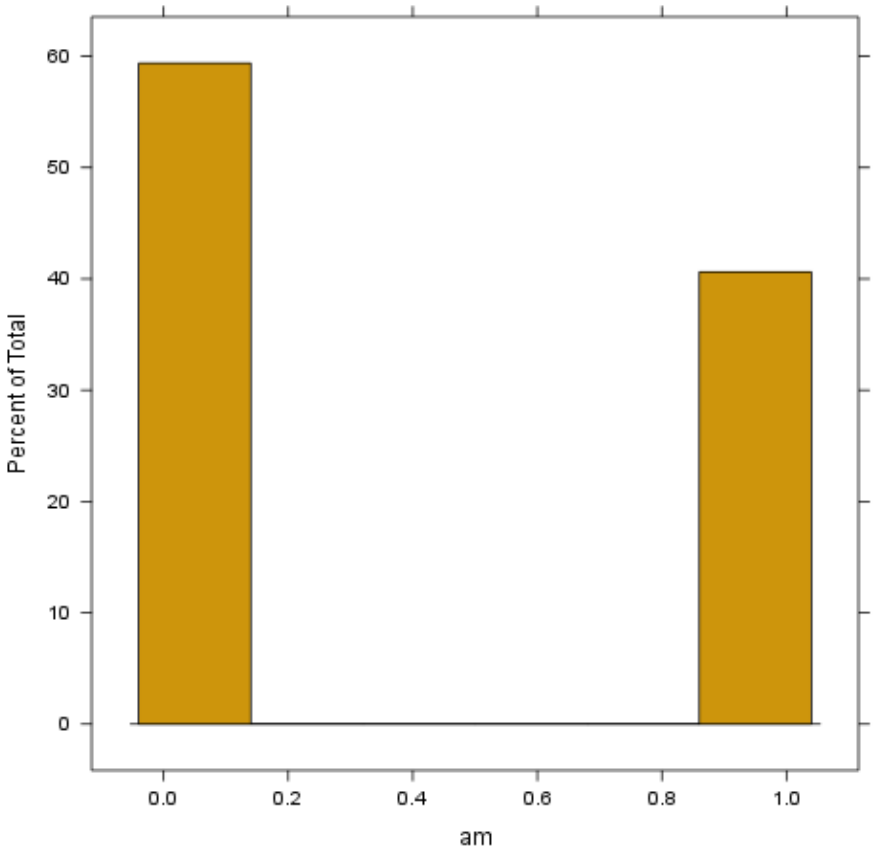
We found the folloing values here:

1, *1*, *1*, *0*, *0*, *0*, *0*, *0*, *0*, *0*, *0*, *0*, *0*, *0*, *0*, *0*, *0*, *1*, *1*, *1*, *0*, *0*, *0*, *0*, *0*, *1*, *1*, *1*, *1*, *1*, *1* and *1*

The mean of am is *0.4062* while the standard deviation is: *0.499*. The most frequent value in am is 0, but let us check out the frequency table too:

0	1
19	13

Tables are boring, let us show the same with a `histogram`:



gear

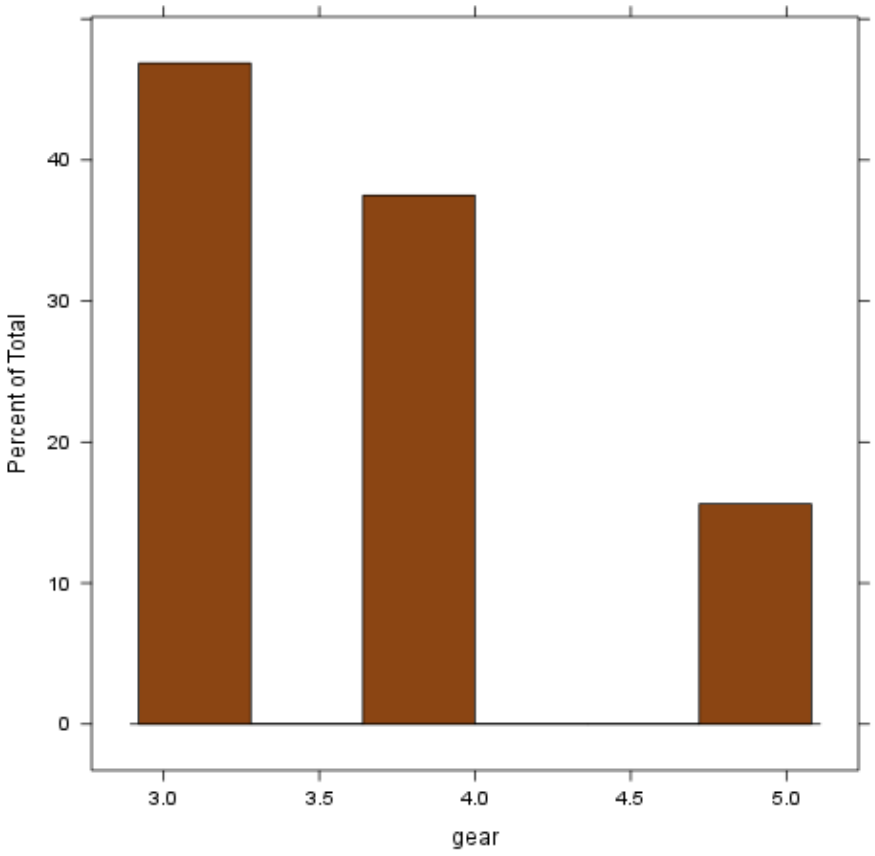
We found the folloing values here:

4, *4*, *4*, *3*, *3*, *3*, *3*, *4*, *4*, *4*, *4*, *3*, *3*, *3*, *3*, *3*, *3*, *4*, *4*, *4*, *3*, *3*, *3*, *3*, *3*, *4*, *5*, *5*, *5*, *5*, *5* and *4*

The mean of gear is *3.688* while the standard deviation is: *0.7378*. The most frequent value in gear is 3, but let us check out the frequency table too:

3	4	5
15	12	5

Tables are boring, let us show the same with a `histogram`:



carb

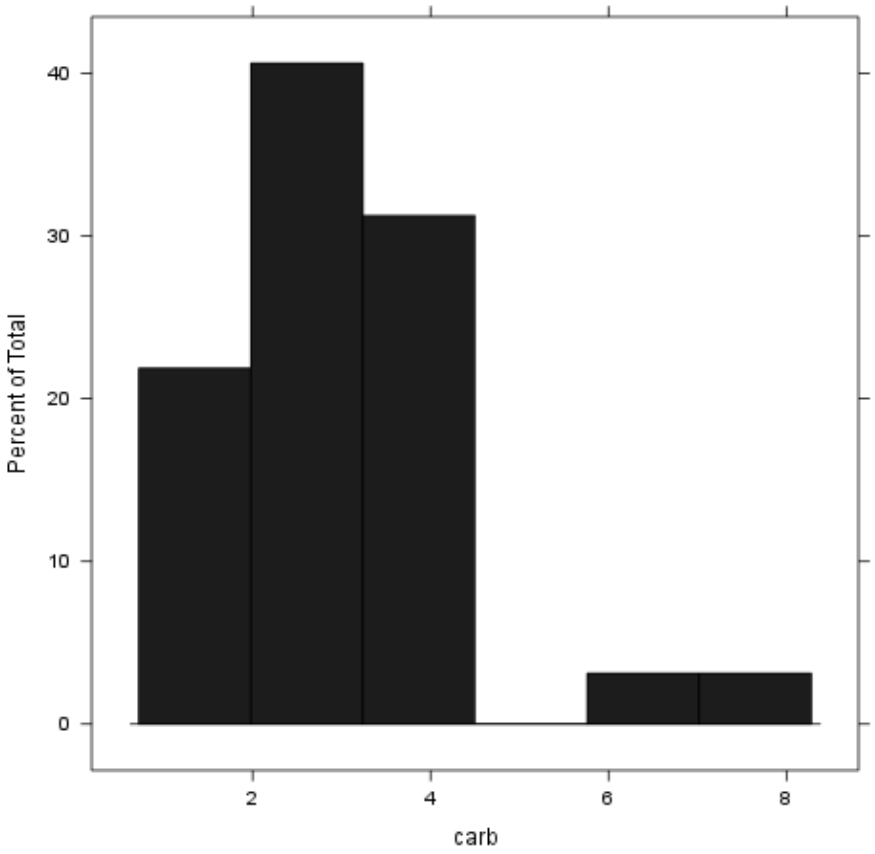
We found the folloing values here:

4, *4*, *1*, *1*, *2*, *1*, *4*, *2*, *2*, *4*, *4*, *3*, *3*, *3*, *4*, *4*, *4*, *1*, *2*, *1*, *1*, *2*, *2*, *4*, *2*, *1*, *2*, *2*, *4*, *6*, *8* and *2*

The mean of carb is *2.812* while the standard deviation is: *1.615*. The most frequent value in carb is 2, but let us check out the frequency table too:

1	2	3	4	6	8
7	10	3	10	1	1

Tables are boring, let us show the same with a `histogram`:



Correlation

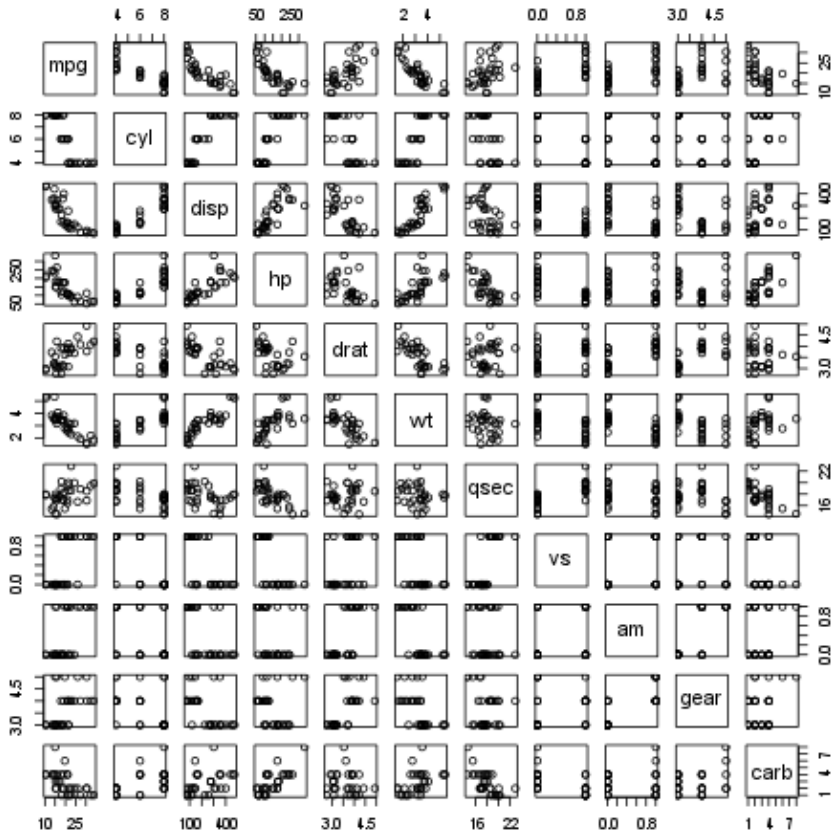
And here goes a correlation table:

	mpg	cyl	disp	hp	drat
mpg	1.00000	-0.85216	-0.84755	-0.77617	0.68117
cyl	-0.85216	1.00000	0.90203	0.83245	-0.69994
disp	-0.84755	0.90203	1.00000	0.79095	-0.71021
hp	-0.77617	0.83245	0.79095	1.00000	-0.44876
drat	0.68117	-0.69994	-0.71021	-0.44876	1.00000
wt	-0.86766	0.78250	0.88798	0.65875	-0.71244
qsec	0.41868	-0.59124	-0.43370	-0.70822	0.09120
vs	0.66404	-0.81081	-0.71042	-0.72310	0.44028
am	0.59983	-0.52261	-0.59123	-0.24320	0.71271
gear	0.48028	-0.49269	-0.55557	-0.12570	0.69961
carb	-0.55093	0.52699	0.39498	0.74981	-0.09079

	wt	qsec	vs	am
mpg	-0.86766	0.41868	0.66404	0.59983
cyl	0.78250	-0.59124	-0.81081	-0.52261
disp	0.88798	-0.43370	-0.71042	-0.59123
hp	0.65875	-0.70822	-0.72310	-0.24320
drat	-0.71244	0.09120	0.44028	0.71271
wt	1.00000	-0.17472	-0.55492	-0.69250
qsec	-0.17472	1.00000	0.74454	-0.22986
vs	-0.55492	0.74454	1.00000	0.16835
am	-0.69250	-0.22986	0.16835	1.00000
gear	-0.58329	-0.21268	0.20602	0.79406
carb	0.42761	-0.65625	-0.56961	0.05753

	gear	carb
mpg	0.48028	-0.55093
cyl	-0.49269	0.52699
disp	-0.55557	0.39498
hp	-0.12570	0.74981
drat	0.69961	-0.09079
wt	-0.58329	0.42761
qsec	-0.21268	-0.65625
vs	0.20602	-0.56961
am	0.79406	0.05753
gear	1.00000	0.27407
carb	0.27407	1.00000

And the same on a graph:



Yeah, that latter took a while to render in an image file :)

That's not a `pander` issue.

Some models

Okay, let us find out how `weight` affects other variables:

mpg

A simple linear model: `mtcars$wt ~ mtcars$mpg`

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	6.047e+00	3.087e-01	1.959e+01	1.204e-18
Independent	-1.409e-01	1.474e-02	-9.559e+00	1.294e-10

Table 1: Fitting linear model: mtcars$wt ~ Independent

cyl

A simple linear model: `mtcars$wt ~ mtcars$cyl`

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	5.646e-01	4.006e-01	1.409e+00	1.690e-01
Independent	4.287e-01	6.228e-02	6.883e+00	1.218e-07

Table 2: Fitting linear model: mtcars$wt ~ Independent

disp

A simple linear model: `mtcars$wt ~ mtcars$disp`

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	1.600e+00	1.730e-01	9.248e+00	2.738e-10
Independent	7.010e-03	6.629e-04	1.058e+01	1.222e-11

Table 3: Fitting linear model: mtcars$wt ~ Independent

hp

A simple linear model: `mtcars$wt ~ mtcars$hp`

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	1.838e+00	3.165e-01	5.808e+00	2.389e-06
Independent	9.401e-03	1.960e-03	4.796e+00	4.146e-05

Table 4: Fitting linear model: mtcars$wt ~ Independent

drat

A simple linear model: `mtcars$wt ~ mtcars$drat`

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	7.906e+00	8.522e-01	9.277e+00	2.547e-10
Independent	-1.304e+00	2.345e-01	-5.561e+00	4.784e-06

Table 5: Fitting linear model: mtcars$wt ~ Independent

qsec

A simple linear model: `mtcars$wt ~ mtcars$qsec`

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	4.924792	1.765407	2.789607	0.009081
Independent	-0.095667	0.098433	-0.971907	0.338868

Table 6: Fitting linear model: mtcars$wt ~ Independent

vs

A simple linear model: `mtcars$wt ~ mtcars$vs`

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	3.689e+00	1.950e-01	1.891e+01	3.203e-18
Independent	-1.077e+00	2.949e-01	-3.654e+00	9.798e-04

Table 7: Fitting linear model: mtcars$wt ~ Independent

am

A simple linear model: `mtcars$wt ~ mtcars$am`

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	3.769e+00	1.646e-01	2.289e+01	1.490e-20
Independent	-1.358e+00	2.583e-01	-5.258e+00	1.125e-05

Table 8: Fitting linear model: mtcars$wt ~ Independent

gear

A simple linear model: `mtcars$wt ~ mtcars$gear`

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	6.070e+00	7.392e-01	8.212e+00	3.632e-09
Independent	-7.735e-01	1.967e-01	-3.933e+00	4.587e-04

Table 9: Fitting linear model: mtcars$wt ~ Independent

carb

A simple linear model: `mtcars$wt ~ mtcars$carb`

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	2.489e+00	3.230e-01	7.705e+00	1.353e-08
Independent	2.590e-01	9.998e-02	2.591e+00	1.464e-02

Table 10: Fitting linear model: mtcars$wt ~ Independent

This report was generated with R (2.15.1) and pander (0.1) on x86_64-unknown-linux-gnu platform.