

# R for DS - 2/22/17

## 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:

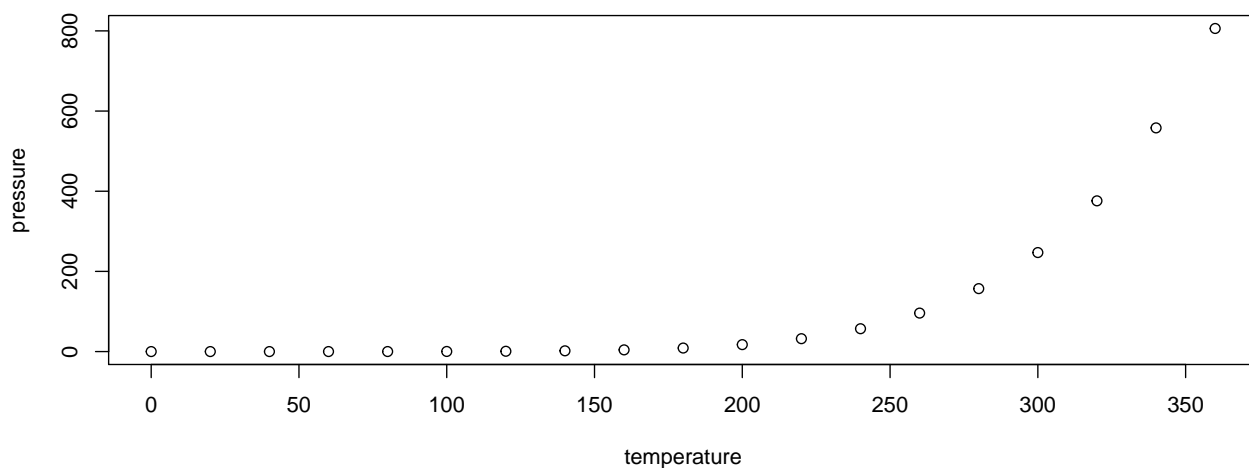
```
summary(cars)
```

```
##      speed      dist
##  Min.   : 4.0    Min.   :  2.00
## 1st Qu.:12.0    1st Qu.: 26.00
##  Median :15.0    Median : 36.00
##   Mean  :15.4    Mean   : 42.98
## 3rd Qu.:19.0    3rd Qu.: 56.00
##   Max.  :25.0    Max.   :120.00
```

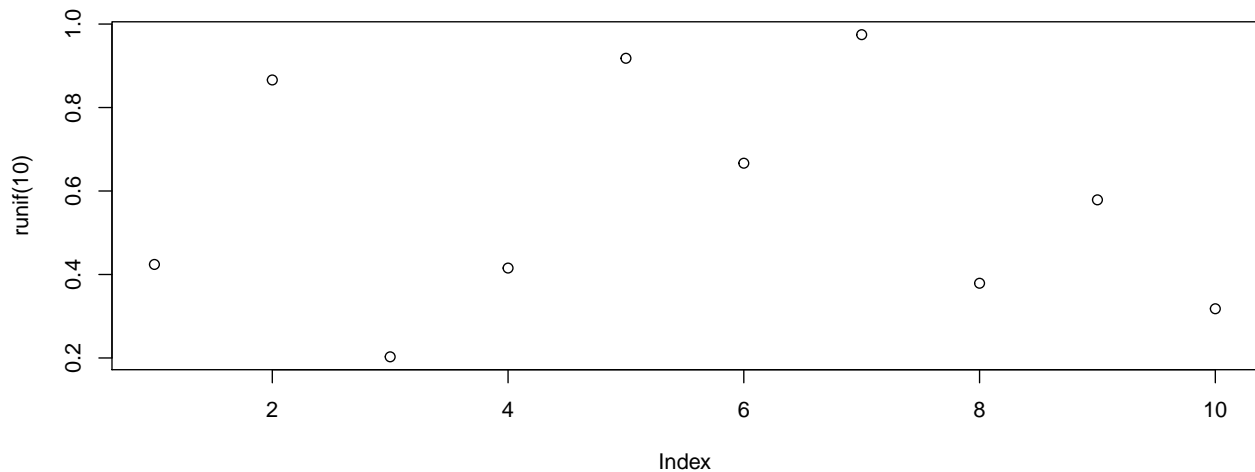
## Including Plots

You can also embed plots, for example:

```
plot(pressure)
```



```
plot(runif(10))
```



Text above

```
fib <- function(n) {
  if (n < 2) return(1)

  fib(n - 1) + fib(n - 2)
}
fib(34)
```

```
## [1] 9227465
```

Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.

Notes

```
library(tidyverse)
```

```
## Loading tidyverse: ggplot2
## Loading tidyverse: tibble
## Loading tidyverse: tidyr
## Loading tidyverse: readr
## Loading tidyverse: purrr
## Loading tidyverse: dplyr
```

```
## Conflicts with tidy packages -----
```

```
## filter(): dplyr, stats
## lag():    dplyr, stats
```

```
library(tibble)
as_tibble(iris)
```

```
## # A tibble: 150 × 5
##   Sepal.Length Sepal.Width Petal.Length Petal.Width Species
##   <dbl>         <dbl>         <dbl>         <dbl> <fctr>
## 1         5.1         3.5         1.4         0.2 setosa
## 2         4.9         3.0         1.4         0.2 setosa
## 3         4.7         3.2         1.3         0.2 setosa
## 4         4.6         3.1         1.5         0.2 setosa
## 5         5.0         3.6         1.4         0.2 setosa
## 6         5.4         3.9         1.7         0.4 setosa
## 7         4.6         3.4         1.4         0.3 setosa
```

```
## 8      5.0      3.4      1.5      0.2 setosa
## 9      4.4      2.9      1.4      0.2 setosa
## 10     4.9      3.1      1.5      0.1 setosa
## # ... with 140 more rows
```

```
as_tibble(iris)
```

```
## # A tibble: 150 × 5
##   Sepal.Length Sepal.Width Petal.Length Petal.Width Species
##   <dbl>         <dbl>         <dbl>         <dbl> <fctr>
## 1      5.1      3.5      1.4      0.2 setosa
## 2      4.9      3.0      1.4      0.2 setosa
## 3      4.7      3.2      1.3      0.2 setosa
## 4      4.6      3.1      1.5      0.2 setosa
## 5      5.0      3.6      1.4      0.2 setosa
## 6      5.4      3.9      1.7      0.4 setosa
## 7      4.6      3.4      1.4      0.3 setosa
## 8      5.0      3.4      1.5      0.2 setosa
## 9      4.4      2.9      1.4      0.2 setosa
## 10     4.9      3.1      1.5      0.1 setosa
## # ... with 140 more rows
```

```
as_tibble(diamonds)
```

```
## # A tibble: 53,940 × 10
##   carat      cut color clarity depth table price      x      y      z
##   <dbl>    <ord> <ord>   <ord> <dbl> <dbl> <int> <dbl> <dbl> <dbl>
## 1  0.23    Ideal     E    SI2  61.5   55   326  3.95  3.98  2.43
## 2  0.21    Premium    E    SI1  59.8   61   326  3.89  3.84  2.31
## 3  0.23      Good     E    VS1  56.9   65   327  4.05  4.07  2.31
## 4  0.29    Premium    I    VS2  62.4   58   334  4.20  4.23  2.63
## 5  0.31      Good     J    SI2  63.3   58   335  4.34  4.35  2.75
## 6  0.24 Very Good    J   VVS2  62.8   57   336  3.94  3.96  2.48
## 7  0.24 Very Good    I   VVS1  62.3   57   336  3.95  3.98  2.47
## 8  0.26 Very Good    H    SI1  61.9   55   337  4.07  4.11  2.53
## 9  0.22      Fair     E    VS2  65.1   61   337  3.87  3.78  2.49
## 10 0.23 Very Good    H    VS1  59.4   61   338  4.00  4.05  2.39
## # ... with 53,930 more rows
```

```
library(ggplot2)
```

```
as_tibble(diamonds)
```

```
## # A tibble: 53,940 × 10
##   carat      cut color clarity depth table price      x      y      z
##   <dbl>    <ord> <ord>   <ord> <dbl> <dbl> <int> <dbl> <dbl> <dbl>
## 1  0.23    Ideal     E    SI2  61.5   55   326  3.95  3.98  2.43
## 2  0.21    Premium    E    SI1  59.8   61   326  3.89  3.84  2.31
## 3  0.23      Good     E    VS1  56.9   65   327  4.05  4.07  2.31
## 4  0.29    Premium    I    VS2  62.4   58   334  4.20  4.23  2.63
## 5  0.31      Good     J    SI2  63.3   58   335  4.34  4.35  2.75
## 6  0.24 Very Good    J   VVS2  62.8   57   336  3.94  3.96  2.48
## 7  0.24 Very Good    I   VVS1  62.3   57   336  3.95  3.98  2.47
## 8  0.26 Very Good    H    SI1  61.9   55   337  4.07  4.11  2.53
## 9  0.22      Fair     E    VS2  65.1   61   337  3.87  3.78  2.49
## 10 0.23 Very Good    H    VS1  59.4   61   338  4.00  4.05  2.39
## # ... with 53,930 more rows
```

diamonds

```
## # A tibble: 53,940 × 10
```

```
##   carat      cut color clarity depth table price      x      y      z
##   <dbl>    <ord> <ord>   <ord> <dbl> <dbl> <int> <dbl> <dbl> <dbl>
## 1  0.23    Ideal     E    SI2  61.5   55   326  3.95  3.98  2.43
## 2  0.21  Premium     E    SI1  59.8   61   326  3.89  3.84  2.31
## 3  0.23     Good     E    VS1  56.9   65   327  4.05  4.07  2.31
## 4  0.29  Premium     I    VS2  62.4   58   334  4.20  4.23  2.63
## 5  0.31     Good     J    SI2  63.3   58   335  4.34  4.35  2.75
## 6  0.24 Very Good     J   VVS2  62.8   57   336  3.94  3.96  2.48
## 7  0.24 Very Good     I   VVS1  62.3   57   336  3.95  3.98  2.47
## 8  0.26 Very Good     H    SI1  61.9   55   337  4.07  4.11  2.53
## 9  0.22     Fair     E    VS2  65.1   61   337  3.87  3.78  2.49
## 10 0.23 Very Good     H    VS1  59.4   61   338  4.00  4.05  2.39
## # ... with 53,930 more rows
```

```
as.data.frame(diamonds) %>% head(100) # only show top 100
```

```
##   carat      cut color clarity depth table price      x      y      z
## 1  0.23    Ideal     E    SI2  61.5  55.0   326  3.95  3.98  2.43
## 2  0.21  Premium     E    SI1  59.8  61.0   326  3.89  3.84  2.31
## 3  0.23     Good     E    VS1  56.9  65.0   327  4.05  4.07  2.31
## 4  0.29  Premium     I    VS2  62.4  58.0   334  4.20  4.23  2.63
## 5  0.31     Good     J    SI2  63.3  58.0   335  4.34  4.35  2.75
## 6  0.24 Very Good     J   VVS2  62.8  57.0   336  3.94  3.96  2.48
## 7  0.24 Very Good     I   VVS1  62.3  57.0   336  3.95  3.98  2.47
## 8  0.26 Very Good     H    SI1  61.9  55.0   337  4.07  4.11  2.53
## 9  0.22     Fair     E    VS2  65.1  61.0   337  3.87  3.78  2.49
## 10 0.23 Very Good     H    VS1  59.4  61.0   338  4.00  4.05  2.39
## 11 0.30     Good     J    SI1  64.0  55.0   339  4.25  4.28  2.73
## 12 0.23    Ideal     J    VS1  62.8  56.0   340  3.93  3.90  2.46
## 13 0.22  Premium     F    SI1  60.4  61.0   342  3.88  3.84  2.33
## 14 0.31    Ideal     J    SI2  62.2  54.0   344  4.35  4.37  2.71
## 15 0.20  Premium     E    SI2  60.2  62.0   345  3.79  3.75  2.27
## 16 0.32  Premium     E     I1  60.9  58.0   345  4.38  4.42  2.68
## 17 0.30    Ideal     I    SI2  62.0  54.0   348  4.31  4.34  2.68
## 18 0.30     Good     J    SI1  63.4  54.0   351  4.23  4.29  2.70
## 19 0.30     Good     J    SI1  63.8  56.0   351  4.23  4.26  2.71
## 20 0.30 Very Good     J    SI1  62.7  59.0   351  4.21  4.27  2.66
## 21 0.30     Good     I    SI2  63.3  56.0   351  4.26  4.30  2.71
## 22 0.23 Very Good     E    VS2  63.8  55.0   352  3.85  3.92  2.48
## 23 0.23 Very Good     H    VS1  61.0  57.0   353  3.94  3.96  2.41
## 24 0.31 Very Good     J    SI1  59.4  62.0   353  4.39  4.43  2.62
## 25 0.31 Very Good     J    SI1  58.1  62.0   353  4.44  4.47  2.59
## 26 0.23 Very Good     G   VVS2  60.4  58.0   354  3.97  4.01  2.41
## 27 0.24  Premium     I    VS1  62.5  57.0   355  3.97  3.94  2.47
## 28 0.30 Very Good     J    VS2  62.2  57.0   357  4.28  4.30  2.67
## 29 0.23 Very Good     D    VS2  60.5  61.0   357  3.96  3.97  2.40
## 30 0.23 Very Good     F    VS1  60.9  57.0   357  3.96  3.99  2.42
## 31 0.23 Very Good     F    VS1  60.0  57.0   402  4.00  4.03  2.41
## 32 0.23 Very Good     F    VS1  59.8  57.0   402  4.04  4.06  2.42
## 33 0.23 Very Good     E    VS1  60.7  59.0   402  3.97  4.01  2.42
## 34 0.23 Very Good     E    VS1  59.5  58.0   402  4.01  4.06  2.40
## 35 0.23 Very Good     D    VS1  61.9  58.0   402  3.92  3.96  2.44
```

## 36	0.23	Good	F	VS1	58.2	59.0	402	4.06	4.08	2.37
## 37	0.23	Good	E	VS1	64.1	59.0	402	3.83	3.85	2.46
## 38	0.31	Good	H	SI1	64.0	54.0	402	4.29	4.31	2.75
## 39	0.26	Very Good	D	VS2	60.8	59.0	403	4.13	4.16	2.52
## 40	0.33	Ideal	I	SI2	61.8	55.0	403	4.49	4.51	2.78
## 41	0.33	Ideal	I	SI2	61.2	56.0	403	4.49	4.50	2.75
## 42	0.33	Ideal	J	SI1	61.1	56.0	403	4.49	4.55	2.76
## 43	0.26	Good	D	VS2	65.2	56.0	403	3.99	4.02	2.61
## 44	0.26	Good	D	VS1	58.4	63.0	403	4.19	4.24	2.46
## 45	0.32	Good	H	SI2	63.1	56.0	403	4.34	4.37	2.75
## 46	0.29	Premium	F	SI1	62.4	58.0	403	4.24	4.26	2.65
## 47	0.32	Very Good	H	SI2	61.8	55.0	403	4.35	4.42	2.71
## 48	0.32	Good	H	SI2	63.8	56.0	403	4.36	4.38	2.79
## 49	0.25	Very Good	E	VS2	63.3	60.0	404	4.00	4.03	2.54
## 50	0.29	Very Good	H	SI2	60.7	60.0	404	4.33	4.37	2.64
## 51	0.24	Very Good	F	SI1	60.9	61.0	404	4.02	4.03	2.45
## 52	0.23	Ideal	G	VS1	61.9	54.0	404	3.93	3.95	2.44
## 53	0.32	Ideal	I	SI1	60.9	55.0	404	4.45	4.48	2.72
## 54	0.22	Premium	E	VS2	61.6	58.0	404	3.93	3.89	2.41
## 55	0.22	Premium	D	VS2	59.3	62.0	404	3.91	3.88	2.31
## 56	0.30	Ideal	I	SI2	61.0	59.0	405	4.30	4.33	2.63
## 57	0.30	Premium	J	SI2	59.3	61.0	405	4.43	4.38	2.61
## 58	0.30	Very Good	I	SI1	62.6	57.0	405	4.25	4.28	2.67
## 59	0.30	Very Good	I	SI1	63.0	57.0	405	4.28	4.32	2.71
## 60	0.30	Good	I	SI1	63.2	55.0	405	4.25	4.29	2.70
## 61	0.35	Ideal	I	VS1	60.9	57.0	552	4.54	4.59	2.78
## 62	0.30	Premium	D	SI1	62.6	59.0	552	4.23	4.27	2.66
## 63	0.30	Ideal	D	SI1	62.5	57.0	552	4.29	4.32	2.69
## 64	0.30	Ideal	D	SI1	62.1	56.0	552	4.30	4.33	2.68
## 65	0.42	Premium	I	SI2	61.5	59.0	552	4.78	4.84	2.96
## 66	0.28	Ideal	G	VVS2	61.4	56.0	553	4.19	4.22	2.58
## 67	0.32	Ideal	I	VVS1	62.0	55.3	553	4.39	4.42	2.73
## 68	0.31	Very Good	G	SI1	63.3	57.0	553	4.33	4.30	2.73
## 69	0.31	Premium	G	SI1	61.8	58.0	553	4.35	4.32	2.68
## 70	0.24	Premium	E	VVS1	60.7	58.0	553	4.01	4.03	2.44
## 71	0.24	Very Good	D	VVS1	61.5	60.0	553	3.97	4.00	2.45
## 72	0.30	Very Good	H	SI1	63.1	56.0	554	4.29	4.27	2.70
## 73	0.30	Premium	H	SI1	62.9	59.0	554	4.28	4.24	2.68
## 74	0.30	Premium	H	SI1	62.5	57.0	554	4.29	4.25	2.67
## 75	0.30	Good	H	SI1	63.7	57.0	554	4.28	4.26	2.72
## 76	0.26	Very Good	F	VVS2	59.2	60.0	554	4.19	4.22	2.49
## 77	0.26	Very Good	E	VVS2	59.9	58.0	554	4.15	4.23	2.51
## 78	0.26	Very Good	D	VVS2	62.4	54.0	554	4.08	4.13	2.56
## 79	0.26	Very Good	D	VVS2	62.8	60.0	554	4.01	4.05	2.53
## 80	0.26	Very Good	E	VVS1	62.6	59.0	554	4.06	4.09	2.55
## 81	0.26	Very Good	E	VVS1	63.4	59.0	554	4.00	4.04	2.55
## 82	0.26	Very Good	D	VVS1	62.1	60.0	554	4.03	4.12	2.53
## 83	0.26	Ideal	E	VVS2	62.9	58.0	554	4.02	4.06	2.54
## 84	0.38	Ideal	I	SI2	61.6	56.0	554	4.65	4.67	2.87
## 85	0.26	Good	E	VVS1	57.9	60.0	554	4.22	4.25	2.45
## 86	0.24	Premium	G	VVS1	62.3	59.0	554	3.95	3.92	2.45
## 87	0.24	Premium	H	VVS1	61.2	58.0	554	4.01	3.96	2.44
## 88	0.24	Premium	H	VVS1	60.8	59.0	554	4.02	4.00	2.44
## 89	0.24	Premium	H	VVS2	60.7	58.0	554	4.07	4.04	2.46

```
## 90  0.32  Premium    I    SI1  62.9  58.0   554  4.35  4.33  2.73
## 91  0.70    Ideal    E    SI1  62.5  57.0  2757  5.70  5.72  3.57
## 92  0.86     Fair    E    SI2  55.1  69.0  2757  6.45  6.33  3.52
## 93  0.70    Ideal    G    VS2  61.6  56.0  2757  5.70  5.67  3.50
## 94  0.71 Very Good    E    VS2  62.4  57.0  2759  5.68  5.73  3.56
## 95  0.78 Very Good    G    SI2  63.8  56.0  2759  5.81  5.85  3.72
## 96  0.70     Good    E    VS2  57.5  58.0  2759  5.85  5.90  3.38
## 97  0.70     Good    F    VS1  59.4  62.0  2759  5.71  5.76  3.40
## 98  0.96     Fair    F    SI2  66.3  62.0  2759  6.27  5.95  4.07
## 99  0.73 Very Good    E    SI1  61.6  59.0  2760  5.77  5.78  3.56
## 100 0.80  Premium    H    SI1  61.5  58.0  2760  5.97  5.93  3.66
```

```
diamonds
```

```
## # A tibble: 53,940 × 10
```

```
##   carat      cut color clarity depth table price      x      y      z
##   <dbl>    <ord> <ord>   <ord> <dbl> <dbl> <int> <dbl> <dbl> <dbl>
## 1  0.23    Ideal    E     SI2  61.5   55   326   3.95   3.98   2.43
## 2  0.21  Premium    E     SI1  59.8   61   326   3.89   3.84   2.31
## 3  0.23     Good    E     VS1  56.9   65   327   4.05   4.07   2.31
## 4  0.29  Premium    I     VS2  62.4   58   334   4.20   4.23   2.63
## 5  0.31     Good    J     SI2  63.3   58   335   4.34   4.35   2.75
## 6  0.24 Very Good    J    VVS2  62.8   57   336   3.94   3.96   2.48
## 7  0.24 Very Good    I    VVS1  62.3   57   336   3.95   3.98   2.47
## 8  0.26 Very Good    H     SI1  61.9   55   337   4.07   4.11   2.53
## 9  0.22     Fair    E     VS2  65.1   61   337   3.87   3.78   2.49
## 10 0.23 Very Good    H     VS1  59.4   61   338   4.00   4.05   2.39
## # ... with 53,930 more rows
```

```
str(as.data.frame(diamonds) %>% head(100))
```

```
## 'data.frame':   100 obs. of  10 variables:
## $ carat : num  0.23 0.21 0.23 0.29 0.31 0.24 0.24 0.26 0.22 0.23 ...
## $ cut : Ord.factor w/ 5 levels "Fair"<"Good"<...: 5 4 2 4 2 3 3 3 1 3 ...
## $ color : Ord.factor w/ 7 levels "D"<"E"<"F"<"G"<...: 2 2 2 6 7 7 6 5 2 5 ...
## $ clarity: Ord.factor w/ 8 levels "I1"<"SI2"<"SI1"<...: 2 3 5 4 2 6 7 3 4 5 ...
## $ depth : num  61.5 59.8 56.9 62.4 63.3 62.8 62.3 61.9 65.1 59.4 ...
## $ table : num  55 61 65 58 58 57 57 55 61 61 ...
## $ price : int  326 326 327 334 335 336 336 337 337 338 ...
## $ x : num  3.95 3.89 4.05 4.2 4.34 3.94 3.95 4.07 3.87 4 ...
## $ y : num  3.98 3.84 4.07 4.23 4.35 3.96 3.98 4.11 3.78 4.05 ...
## $ z : num  2.43 2.31 2.31 2.63 2.75 2.48 2.47 2.53 2.49 2.39 ...
```

```
glimpse(diamonds)
```

```
## Observations: 53,940
## Variables: 10
## $ carat <dbl> 0.23, 0.21, 0.23, 0.29, 0.31, 0.24, 0.24, 0.26, 0.22, ...
## $ cut <ord> Ideal, Premium, Good, Premium, Good, Very Good, Very G...
## $ color <ord> E, E, E, I, J, J, I, H, E, H, J, J, F, J, E, E, I, J, ...
## $ clarity <ord> SI2, SI1, VS1, VS2, SI2, VVS2, VVS1, SI1, VS2, VS1, SI...
## $ depth <dbl> 61.5, 59.8, 56.9, 62.4, 63.3, 62.8, 62.3, 61.9, 65.1, ...
## $ table <dbl> 55, 61, 65, 58, 58, 57, 57, 55, 61, 61, 55, 56, 61, 54...
## $ price <int> 326, 326, 327, 334, 335, 336, 336, 337, 337, 338, ...
## $ x <dbl> 3.95, 3.89, 4.05, 4.20, 4.34, 3.94, 3.95, 4.07, 3.87, ...
## $ y <dbl> 3.98, 3.84, 4.07, 4.23, 4.35, 3.96, 3.98, 4.11, 3.78, ...
```

```
## $ z      <dbl> 2.43, 2.31, 2.31, 2.63, 2.75, 2.48, 2.47, 2.53, 2.49, ...
```

```
diamonds[,1]
```

```
## # A tibble: 53,940 × 1
##   carat
##   <dbl>
## 1    0.23
## 2    0.21
## 3    0.23
## 4    0.29
## 5    0.31
## 6    0.24
## 7    0.24
## 8    0.26
## 9    0.22
## 10   0.23
## # ... with 53,930 more rows
```

```
as.data.frame(diamonds)[,1] %>% head(100)
```

```
##   [1] 0.23 0.21 0.23 0.29 0.31 0.24 0.24 0.26 0.22 0.23 0.30 0.23 0.22 0.31
##  [15] 0.20 0.32 0.30 0.30 0.30 0.30 0.30 0.23 0.23 0.31 0.31 0.23 0.24 0.30
##  [29] 0.23 0.23 0.23 0.23 0.23 0.23 0.23 0.23 0.31 0.26 0.33 0.33 0.33
##  [43] 0.26 0.26 0.32 0.29 0.32 0.32 0.25 0.29 0.24 0.23 0.32 0.22 0.22 0.30
##  [57] 0.30 0.30 0.30 0.30 0.35 0.30 0.30 0.30 0.42 0.28 0.32 0.31 0.31 0.24
##  [71] 0.24 0.30 0.30 0.30 0.30 0.26 0.26 0.26 0.26 0.26 0.26 0.26 0.26 0.38
##  [85] 0.26 0.24 0.24 0.24 0.24 0.32 0.70 0.86 0.70 0.71 0.78 0.70 0.70 0.96
##  [99] 0.73 0.80
```

```
diamonds[,1]
```

```
## # A tibble: 53,940 × 1
##   carat
##   <dbl>
## 1    0.23
## 2    0.21
## 3    0.23
## 4    0.29
## 5    0.31
## 6    0.24
## 7    0.24
## 8    0.26
## 9    0.22
## 10   0.23
## # ... with 53,930 more rows
```

```
diamonds[[1]] %>% head(100)
```

```
##   [1] 0.23 0.21 0.23 0.29 0.31 0.24 0.24 0.26 0.22 0.23 0.30 0.23 0.22 0.31
##  [15] 0.20 0.32 0.30 0.30 0.30 0.30 0.30 0.23 0.23 0.31 0.31 0.23 0.24 0.30
##  [29] 0.23 0.23 0.23 0.23 0.23 0.23 0.23 0.23 0.31 0.26 0.33 0.33 0.33
##  [43] 0.26 0.26 0.32 0.29 0.32 0.32 0.25 0.29 0.24 0.23 0.32 0.22 0.22 0.30
##  [57] 0.30 0.30 0.30 0.30 0.35 0.30 0.30 0.30 0.42 0.28 0.32 0.31 0.31 0.24
##  [71] 0.24 0.30 0.30 0.30 0.30 0.26 0.26 0.26 0.26 0.26 0.26 0.26 0.26 0.38
##  [85] 0.26 0.24 0.24 0.24 0.24 0.32 0.70 0.86 0.70 0.71 0.78 0.70 0.70 0.96
##  [99] 0.73 0.80
```

```
diamonds
```

```
## # A tibble: 53,940 × 10
##   carat      cut color clarity depth table price      x      y      z
##   <dbl>    <ord> <ord>   <ord> <dbl> <dbl> <int> <dbl> <dbl> <dbl>
## 1  0.23    Ideal     E     SI2  61.5   55   326  3.95  3.98  2.43
## 2  0.21  Premium     E     SI1  59.8   61   326  3.89  3.84  2.31
## 3  0.23     Good     E     VS1  56.9   65   327  4.05  4.07  2.31
## 4  0.29  Premium     I     VS2  62.4   58   334  4.20  4.23  2.63
## 5  0.31     Good     J     SI2  63.3   58   335  4.34  4.35  2.75
## 6  0.24 Very Good     J    VVS2  62.8   57   336  3.94  3.96  2.48
## 7  0.24 Very Good     I    VVS1  62.3   57   336  3.95  3.98  2.47
## 8  0.26 Very Good     H     SI1  61.9   55   337  4.07  4.11  2.53
## 9  0.22     Fair     E     VS2  65.1   61   337  3.87  3.78  2.49
## 10 0.23 Very Good     H     VS1  59.4   61   338  4.00  4.05  2.39
## # ... with 53,930 more rows
```

```
diamonds$carat %>% head(100)
```

```
##   [1] 0.23 0.21 0.23 0.29 0.31 0.24 0.24 0.26 0.22 0.23 0.30 0.23 0.22 0.31
##  [15] 0.20 0.32 0.30 0.30 0.30 0.30 0.30 0.23 0.23 0.31 0.31 0.23 0.24 0.30
##  [29] 0.23 0.23 0.23 0.23 0.23 0.23 0.23 0.23 0.23 0.31 0.26 0.33 0.33 0.33
##  [43] 0.26 0.26 0.32 0.29 0.32 0.32 0.25 0.29 0.24 0.23 0.32 0.22 0.22 0.30
##  [57] 0.30 0.30 0.30 0.30 0.35 0.30 0.30 0.30 0.42 0.28 0.32 0.31 0.31 0.24
##  [71] 0.24 0.30 0.30 0.30 0.30 0.26 0.26 0.26 0.26 0.26 0.26 0.26 0.26 0.38
##  [85] 0.26 0.24 0.24 0.24 0.24 0.32 0.70 0.86 0.70 0.71 0.78 0.70 0.70 0.96
##  [99] 0.73 0.80
```

```
diamonds
```

```
## # A tibble: 53,940 × 10
##   carat      cut color clarity depth table price      x      y      z
##   <dbl>    <ord> <ord>   <ord> <dbl> <dbl> <int> <dbl> <dbl> <dbl>
## 1  0.23    Ideal     E     SI2  61.5   55   326  3.95  3.98  2.43
## 2  0.21  Premium     E     SI1  59.8   61   326  3.89  3.84  2.31
## 3  0.23     Good     E     VS1  56.9   65   327  4.05  4.07  2.31
## 4  0.29  Premium     I     VS2  62.4   58   334  4.20  4.23  2.63
## 5  0.31     Good     J     SI2  63.3   58   335  4.34  4.35  2.75
## 6  0.24 Very Good     J    VVS2  62.8   57   336  3.94  3.96  2.48
## 7  0.24 Very Good     I    VVS1  62.3   57   336  3.95  3.98  2.47
## 8  0.26 Very Good     H     SI1  61.9   55   337  4.07  4.11  2.53
## 9  0.22     Fair     E     VS2  65.1   61   337  3.87  3.78  2.49
## 10 0.23 Very Good     H     VS1  59.4   61   338  4.00  4.05  2.39
## # ... with 53,930 more rows
```

```
diamonds[, 1]
```

```
## # A tibble: 53,940 × 1
##   carat
##   <dbl>
## 1  0.23
## 2  0.21
## 3  0.23
## 4  0.29
## 5  0.31
## 6  0.24
## 7  0.24
```



```
## 8 0.26
## 9 0.22
## 10 0.23
## # ... with 53,930 more rows
```

```
diamonds[1, ] %>% head(100)
```

```
## # A tibble: 1 × 10
##   carat cut color clarity depth table price x y z
##   <dbl> <ord> <ord> <ord> <dbl> <dbl> <int> <dbl> <dbl> <dbl>
## 1 0.23 Ideal E SI2 61.5 55 326 3.95 3.98 2.43
```

```
as.data.frame(diamonds)[1, ] %>% head(100)
```

```
##   carat cut color clarity depth table price x y z
## 1 0.23 Ideal E SI2 61.5 55 326 3.95 3.98 2.43
```

```
library(dplyr)
diamonds %>% filter(carat <= 0.25)
```

```
## # A tibble: 785 × 10
##   carat cut color clarity depth table price x y z
##   <dbl> <ord> <ord> <ord> <dbl> <dbl> <int> <dbl> <dbl> <dbl>
## 1 0.23 Ideal E SI2 61.5 55 326 3.95 3.98 2.43
## 2 0.21 Premium E SI1 59.8 61 326 3.89 3.84 2.31
## 3 0.23 Good E VS1 56.9 65 327 4.05 4.07 2.31
## 4 0.24 Very Good J VVS2 62.8 57 336 3.94 3.96 2.48
## 5 0.24 Very Good I VVS1 62.3 57 336 3.95 3.98 2.47
## 6 0.22 Fair E VS2 65.1 61 337 3.87 3.78 2.49
## 7 0.23 Very Good H VS1 59.4 61 338 4.00 4.05 2.39
## 8 0.23 Ideal J VS1 62.8 56 340 3.93 3.90 2.46
## 9 0.22 Premium F SI1 60.4 61 342 3.88 3.84 2.33
## 10 0.20 Premium E SI2 60.2 62 345 3.79 3.75 2.27
## # ... with 775 more rows
```

```
my_string <- "
multi
line
string"
```

```
library(stringr)
diamonds %>%
  mutate(
    cut_awesome = str_replace(cut, "Good", "Awesome")
  ) ->
barret_diamonds

barret_diamonds
```

```
## # A tibble: 53,940 × 11
##   carat cut color clarity depth table price x y z
##   <dbl> <ord> <ord> <ord> <dbl> <dbl> <int> <dbl> <dbl> <dbl>
## 1 0.23 Ideal E SI2 61.5 55 326 3.95 3.98 2.43
## 2 0.21 Premium E SI1 59.8 61 326 3.89 3.84 2.31
## 3 0.23 Good E VS1 56.9 65 327 4.05 4.07 2.31
## 4 0.29 Premium I VS2 62.4 58 334 4.20 4.23 2.63
## 5 0.31 Good J SI2 63.3 58 335 4.34 4.35 2.75
```

```
## 6 0.24 Very Good J VVS2 62.8 57 336 3.94 3.96 2.48
## 7 0.24 Very Good I VVS1 62.3 57 336 3.95 3.98 2.47
## 8 0.26 Very Good H SI1 61.9 55 337 4.07 4.11 2.53
## 9 0.22 Fair E VS2 65.1 61 337 3.87 3.78 2.49
## 10 0.23 Very Good H VS1 59.4 61 338 4.00 4.05 2.39
## # ... with 53,930 more rows, and 1 more variables: cut_awsome <chr>
```

```
barret_diamonds$cut %>% head(100)
```

```
## [1] Ideal Premium Good Premium Good Very Good Very Good
## [8] Very Good Fair Very Good Good Ideal Premium Ideal
## [15] Premium Premium Ideal Good Good Very Good Good
## [22] Very Good Very Good Very Good Very Good Very Good Premium Very Good
## [29] Very Good Very Good Very Good Very Good Very Good Very Good Very Good
## [36] Good Good Good Very Good Ideal Ideal Ideal
## [43] Good Good Good Premium Very Good Good Very Good
## [50] Very Good Very Good Ideal Ideal Premium Premium Ideal
## [57] Premium Very Good Very Good Good Ideal Premium Ideal
## [64] Ideal Premium Ideal Ideal Very Good Premium Premium
## [71] Very Good Very Good Premium Premium Good Very Good Very Good
## [78] Very Good Very Good Very Good Very Good Very Good Ideal Ideal
## [85] Good Premium Premium Premium Premium Premium Ideal
## [92] Fair Ideal Very Good Very Good Good Good Fair
## [99] Very Good Premium
## Levels: Fair < Good < Very Good < Premium < Ideal
```