Data Structures, S3, and Subsetting

Colin Rundel

2019-01-17

Tibbles

Modern data frames

Hadley Wickham has a package that modifies data frames to be more modern, or as he calls them surly and lazy.

```
library(tibble)
class(iris)
```

```
## [1] "data.frame"
```

```
tbl_iris = as_tibble(iris)
class(tbl_iris)
```

```
## [1] "tbl_df"     "tbl"        "data.frame"
```

Fancy Printing

```
tbl_iris
```

```
## # A tibble: 150 x 5
##    Sepal.Length Sepal.Width Petal.Length Petal.Width Species
##           <dbl>       <dbl>        <dbl>       <dbl> <fct>
##  1          5.1         3.5          1.4         0.2 setosa
##  2          4.9         3            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           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           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
```

```
data_frame(x = rnorm(10,sd=5), y = rnorm(10))
```

```
## Warning: `data_frame()` is deprecated, use `tibble()`.
## This warning is displayed once per session.
```

Tibbles are lazy

```
tbl_iris[1,]
```

```
## # A tibble: 1 x 5
##   Sepal.Length Sepal.Width Petal.Length Petal.Width Species
##          <dbl>       <dbl>        <dbl>       <dbl> <fct>
## 1          5.1         3.5          1.4         0.2 setosa
```

```
tbl_iris[,"Species"]
```

```
## # A tibble: 150 x 1
##    Species
##    <fct>
##  1 setosa
##  2 setosa
##  3 setosa
##  4 setosa
##  5 setosa
##  6 setosa
##  7 setosa
##  8 setosa
##  9 setosa
```

Tibbles are lazy

```
tbl_iris[1,]
```

```
## # A tibble: 1 x 5
##   Sepal.Length Sepal.Width Petal.Length Petal.Width Species
##          <dbl>       <dbl>        <dbl>       <dbl> <fct>
## 1          5.1         3.5          1.4         0.2 setosa
```

```
tbl_iris[,"Species"]
```

```
## # A tibble: 150 x 1
##    Species
##    <fct>
##  1 setosa
##  2 setosa
##  3 setosa
##  4 setosa
##  5 setosa
##  6 setosa
##  7 setosa
##  8 setosa
##  9 setosa
```

```
data_frame(
  x = 1:3,
  y = c("A","B","C")
)
```

```
## # A tibble: 3 x 2
##       x y
##   <int> <chr>
## 1     1 A
## 2     2 B
## 3     3 C
```

Multiple classes

```
d = data_frame(
x = 1:3,
y = c("A","B","C")
)

class(d)
```

```
## [1] "tbl_df"     "tbl"        "data.frame"
```

Multiple classes

```
d = data_frame(
x = 1:3,
y = c("A","B","C")
)

class(d)
```

```
## [1] "tbl_df"     "tbl"        "data.frame"
```

```
class(d) = rev(class(d))
class(d)
```

```
## [1] "data.frame" "tbl"        "tbl_df"
```

```
d
```

```
##   x y
## 1 1 A
## 2 2 B
## 3 3 C
```

Reverting a tbl

```
d = data_frame(
  x = 1:3,
  y = c("A","B","C")
)

d
```

```
## # A tibble: 3 x 2
##       x y
##   <int> <chr>
## 1     1 A
## 2     2 B
## 3     3 C
```

```
data.frame(d)
```

```
##   x y
## 1 1 A
## 2 2 B
## 3 3 C
```

```
class(d) = "data.frame"
d
```

```
##   x y
## 1 1 A
## 2 2 B
## 3 3 C
```