sb_rmd_ex_05

Sander Bossier

2023-11-11

1.

```r
speed_read_baseR <- microbenchmark(
  read.delim(
    file = "https://raw.githubusercontent.com/kplevoet/texts/main/doyle/baskervilles_TAG.txt",
    sep = "\t",
    encoding = "UTF-8",
    quote = ""
    ),
  times = 20,
  unit = "ms"
)

speed_read_readr <- microbenchmark(
  read_delim(
    file = "https://raw.githubusercontent.com/kplevoet/texts/main/doyle/baskervilles_TAG.txt",
    show_col_types = FALSE,
    quote = ""
  ),
  times = 20,
  unit = "ms"
)

speed_read_dt <- microbenchmark(
  fread(
    input = "https://raw.githubusercontent.com/kplevoet/texts/main/doyle/baskervilles_TAG.txt",
    sep = "\t",
    encoding = "UTF-8",
    quote = ""
  ),
  times = 20,
  unit = "ms"
)
```

The average speed of the base R function `read.delim` is 355.666005 ms, the average speed of the tidyverse readr `read_delim` function is 321.39842 ms, and the average speed of the data.table `fread` function is 122.315025 ms.

```r
data_baseR <- read.delim(
  file = "https://raw.githubusercontent.com/kplevoet/texts/main/doyle/baskervilles_TAG.txt",
  sep = "\t",
  encoding = "UTF-8",
  quote = ""
```

```
  )

data_readr <- read_delim(
  file = "https://raw.githubusercontent.com/kplevoet/texts/main/doyle/baskervilles_TAG.txt",
  show_col_types = FALSE,
  quote = ""
  )

data_dt <- fread(
  input = "https://raw.githubusercontent.com/kplevoet/texts/main/doyle/baskervilles_TAG.txt",
  sep = "\t",
  encoding = "UTF-8",
  quote = ""
)
```

The object size for the base R data.frame is 2820.3 Kb, the object size of of tidyverse tibble is 2823.8 Kb, and the size of the data.frame is 2820.9 Kb.