A Notebook for Data Methods Class Fall '22 (importing a file)

Contents

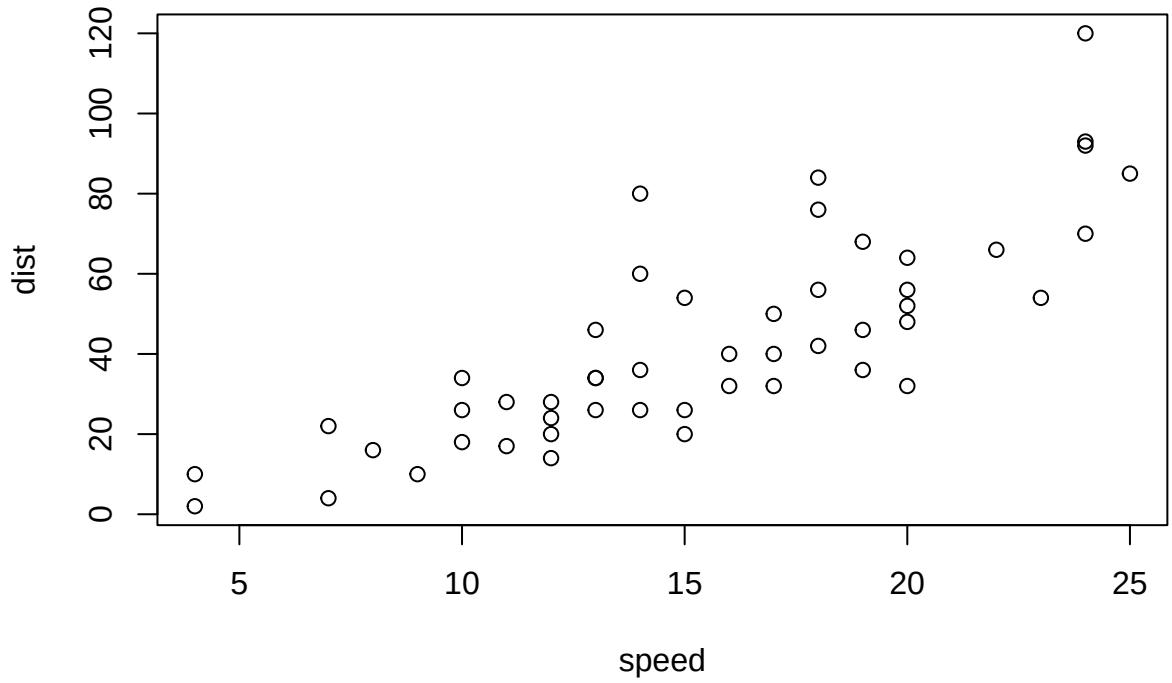
This is an R Markdown Notebook. When you execute code within the notebook, the results appear beneath the code.

```
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.2 --
## v ggplot2 3.3.6     v purrr   0.3.4
## v tibble  3.1.8     v dplyr   1.0.9
## v tidyr   1.2.0     v stringr 1.4.1
## v readr   2.1.2     v forcats 0.5.2
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

Try executing this chunk by clicking the *Run* button within the chunk or by placing your cursor inside it and pressing *Ctrl+Shift+Enter*.

```
plot(cars)
```



Add a new chunk by clicking the *Insert Chunk* button on the toolbar or by pressing *Ctrl+Alt+I*.

When you save the notebook, an HTML file containing the code and output will be saved alongside it (click the *Preview* button or press *Ctrl+Shift+K* to preview the HTML file).

The preview shows you a rendered HTML copy of the contents of the editor. Consequently, unlike *Knit*, *Preview* does not run any R code chunks. Instead, the output of the chunk when it was last run in the editor is displayed.

```
gss <- read_csv("gss2014.csv")
```

```
## Rows: 2538 Columns: 4
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr (1): zodiac
## dbl (3): age, year, childs
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
head(gss)
```

```
## # A tibble: 6 x 4
##     age  year childs zodiac
##   <dbl> <dbl>  <dbl> <chr>
## 1    53  2014      0 virgo
## 2    26  2014      0 virgo
## 3    59  2014      1 gemini
## 4    56  2014      2 scorpio
```

```
## 5    74  2014      3 sagittarius
## 6    56  2014      1 scorpio
```

```
class(gss$age)
```

```
## [1] "numeric"
```

```
class(gss$year)
```

```
## [1] "numeric"
```

```
class(gss$childs)
```

```
## [1] "numeric"
```

```
class(gss$zodiac)
```

```
## [1] "character"
```

```
gss$can_drink <- if_else(gss$age >= 21, 1, 0)
gss
```

```
## # A tibble: 2,538 x 5
##      age  year childs zodiac      can_drink
##    <dbl> <dbl>  <dbl> <chr>           <dbl>
##  1    53  2014      0 virgo               1
##  2    26  2014      0 virgo               1
##  3    59  2014      1 gemini              1
##  4    56  2014      2 scorpio             1
##  5    74  2014      3 sagittarius         1
##  6    56  2014      1 scorpio             1
##  7    63  2014      2 <NA>                1
##  8    34  2014      2 virgo               1
##  9    37  2014      4 virgo               1
## 10    30  2014      3 sagittarius         1
## # ... with 2,528 more rows
```

```
?select
```

```
## starting httpd help server ... done
```

```
?filter
```

```
gss2 <- gss %>% filter(age < 21)
```

```
gss2 <- gss2 %>% select(age, zodiac) %>% arrange(age)
```