Tutorial9

```
library(tidyverse)
```

```
## Warning in file(con, "r"): cannot open file '/var/db/timezone/zoneinfo/
## +VERSION': No such file or directory
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.0 --
```

```
## v ggplot2 3.3.3     v purrr   0.3.4
## v tibble  3.0.6     v dplyr   1.0.7
## v tidyr   1.1.2     v stringr 1.4.0
## v readr   1.4.0     v forcats 0.5.1
```

```
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
library(dplyr)

iris <- read_csv("iris.csv")
```

```
##
## -- Column specification --------------------------------------------------------
## cols(
##   Sepal.Length = col_double(),
##   Sepal.Width = col_double(),
##   Petal.Length = col_double(),
##   Petal.Width = col_double(),
##   Species = col_character()
## )
```

```
##Question 1
readLines('iris.csv', n=2)
```

```
## [1] "\"Sepal.Length\",\"Sepal.Width\",\"Petal.Length\",\"Petal.Width\",\"Species\""
## [2] "5.1,3.5,1.4,0.2,\"setosa\""
```

```
##Question 2
#load data
iris <- read.csv('iris.csv', header = TRUE, stringsAsFactors = FALSE)
#look at data
class(iris)
```

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

```
str(iris)
```

```
## 'data.frame':    150 obs. of  5 variables:
##  $ Sepal.Length: num  5.1 4.9 4.7 4.6 5 5.4 4.6 5 4.4 4.9 ...
##  $ Sepal.Width : num  3.5 3 3.2 3.1 3.6 3.9 3.4 3.4 2.9 3.1 ...
##  $ Petal.Length: num  1.4 1.4 1.3 1.5 1.4 1.7 1.4 1.5 1.4 1.5 ...
##  $ Petal.Width : num  0.2 0.2 0.2 0.2 0.2 0.4 0.3 0.2 0.2 0.1 ...
##  $ Species     : chr  "setosa" "setosa" "setosa" "setosa" ...
```

```
#print last 2 columns and rows
iris[149:150, 4:5]
```

```
##     Petal.Width   Species
## 149         2.3 virginica
## 150         1.8 virginica
```

```
#number observations per species
iris %>%
  count(Species)
```

```
##      Species  n
## 1     setosa 50
## 2 versicolor 50
## 3  virginica 50
```

```
#filter sepal width
iris %>%
  filter(Sepal.Width > 3.5)
```

```
##    Sepal.Length Sepal.Width Petal.Length Petal.Width   Species
## 1           5.0         3.6          1.4         0.2    setosa
## 2           5.4         3.9          1.7         0.4    setosa
## 3           5.4         3.7          1.5         0.2    setosa
## 4           5.8         4.0          1.2         0.2    setosa
## 5           5.7         4.4          1.5         0.4    setosa
## 6           5.4         3.9          1.3         0.4    setosa
## 7           5.7         3.8          1.7         0.3    setosa
## 8           5.1         3.8          1.5         0.3    setosa
## 9           5.1         3.7          1.5         0.4    setosa
## 10          4.6         3.6          1.0         0.2    setosa
## 11          5.2         4.1          1.5         0.1    setosa
## 12          5.5         4.2          1.4         0.2    setosa
## 13          4.9         3.6          1.4         0.1    setosa
## 14          5.1         3.8          1.9         0.4    setosa
## 15          5.1         3.8          1.6         0.2    setosa
## 16          5.3         3.7          1.5         0.2    setosa
## 17          7.2         3.6          6.1         2.5 virginica
## 18          7.7         3.8          6.7         2.2 virginica
## 19          7.9         3.8          6.4         2.0 virginica
```

```r
#create csv with species- setosa
write.csv(iris %>%
            filter(Species=="setosa"),
          "setosa.csv")
#calculate mean, min, and max for petal width of virginica
iris %>%
  filter(Species=="virginica") %>%
  summarise(mean=mean(Petal.Width), minimum=min(Petal.Width),
  maximum=max(Petal.Width))
```

```
##    mean minimum maximum
## 1 2.026     1.4     2.5
```

‘‘‘