Homework Cars -Haley Mincin

Haley Mincin

September 28, 2016

Number of rows and columns (in that sequence)

```
dim(iris)
```

```
## [1] 150   5
```

Number of rows

```
nrow(iris)
```

```
## [1] 150
```

Number of columns

```
ncol(iris)
```

```
## [1] 5
```

Names of variables

```
names(iris) # colnames(iris) also gives that information
```

```
## [1] "Sepal.Length" "Sepal.Width"  "Petal.Length" "Petal.Width"
## [5] "Species"
```

First 6 rows

```
head(iris)
```

```
##   Sepal.Length Sepal.Width Petal.Length Petal.Width Species
## 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
```

First 2 rows

```
head(iris,2) # alternately, can use iris[1:2,]
```

```
##   Sepal.Length Sepal.Width Petal.Length Petal.Width Species
## 1          5.1         3.5          1.4         0.2  setosa
## 2          4.9         3.0          1.4         0.2  setosa
```

Last 6 rows

```
tail(iris) # Number of rows can be controlled, see earlier example involving the head command
```

```
##     Sepal.Length Sepal.Width Petal.Length Petal.Width   Species
## 145          6.7         3.3          5.7         2.5 virginica
## 146          6.7         3.0          5.2         2.3 virginica
## 147          6.3         2.5          5.0         1.9 virginica
## 148          6.5         3.0          5.2         2.0 virginica
## 149          6.2         3.4          5.4         2.3 virginica
## 150          5.9         3.0          5.1         1.8 virginica
```

First row

```
iris[1,]
```

```
##   Sepal.Length Sepal.Width Petal.Length Petal.Width Species
## 1          5.1         3.5          1.4         0.2  setosa
```

First row, first column

```
iris[1,1]
```

```
## [1] 5.1
```

Name of third column

```
names(iris)[3]
```

```
## [1] "Petal.Length"
```

3 entries from third column

```
head(iris[3],3) # alternately, can use iris[1:3,3]
```

```
##   Petal.Length
## 1          1.4
## 2          1.4
## 3          1.3
```

Structure of the dataframe (dataset)

```
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     : Factor w/ 3 levels "setosa","versicolor",..: 1 1 1 1 1 1 1 1 1 1 ...
```

Summary of the dataframe

```r
summary(iris) # for factor/categorical variables, this gives a count of all categories
```

```
##   Sepal.Length    Sepal.Width     Petal.Length    Petal.Width
##  Min.   :4.300   Min.   :2.000   Min.   :1.000   Min.   :0.100
##  1st Qu.:5.100   1st Qu.:2.800   1st Qu.:1.600   1st Qu.:0.300
##  Median :5.800   Median :3.000   Median :4.350   Median :1.300
##  Mean   :5.843   Mean   :3.057   Mean   :3.758   Mean   :1.199
##  3rd Qu.:6.400   3rd Qu.:3.300   3rd Qu.:5.100   3rd Qu.:1.800
##  Max.   :7.900   Max.   :4.400   Max.   :6.900   Max.   :2.500
##        Species
##  setosa    :50
##  versicolor:50
##  virginica :50
##
##
##
```

Create a dataframe

```r
Person=c("A","B","C","D","E")
Age=c(15,20,25,30,35)
page=data.frame(Person,Age)
mean(Age) # gives the mean of the variable Age, prior to the creation of the dataset
```

```
## [1] 25
```

```r
Age="" # (resetting that)
mean(Age) # Haha
```

```
## Warning in mean.default(Age): argument is not numeric or logical: returning
## NA
```

```
## [1] NA
```

```r
mean(page$Age)
```

```
## [1] 25
```

```r
mean(page[,2])
```

```
## [1] 25
```

```r
summary(page)
```

```
##  Person      Age
##  A:1    Min.   :15
##  B:1    1st Qu.:20
##  C:1    Median :25
##  D:1    Mean   :25
##  E:1    3rd Qu.:30
##         Max.   :35
```