Summarys

Jennifer Schlindwein

20 12 2021

```r
for (filename in list.files('data')){
  print('----------------------------------------------------------')
  print(filename)
  dt <- fread(paste('data/',filename,sep=''))
  print(summary(dt))
}
```

```
## [1] "----------------------------------------------------------"
## [1] "accidents_2017.csv"
##       Id             District Name      Neighborhood Name    Street
##  Length:10339       Length:10339       Length:10339       Length:10339
##  Class :character   Class :character   Class :character   Class :character
##  Mode  :character   Mode  :character   Mode  :character   Mode  :character
##
##
##
##     Weekday             Month                 Day              Hour
##  Length:10339       Length:10339       Min.   : 1.00     Min.   : 0.00
##  Class :character   Class :character   1st Qu.: 8.00     1st Qu.:10.00
##  Mode  :character   Mode  :character   Median :16.00     Median :14.00
##                                        Mean   :15.78     Mean   :13.81
##                                        3rd Qu.:23.00     3rd Qu.:18.00
##                                        Max.   :31.00     Max.   :23.00
##  Part of the day    Mild injuries     Serious injuries     Victims
##  Length:10339       Min.   : 0.000    Min.   :0.00000    Min.   : 0.000
##  Class :character   1st Qu.: 1.000    1st Qu.:0.00000    1st Qu.: 1.000
##  Mode  :character   Median : 1.000    Median :0.00000    Median : 1.000
##                     Mean   : 1.154    Mean   :0.02331    Mean   : 1.179
##                     3rd Qu.: 1.000    3rd Qu.:0.00000    3rd Qu.: 1.000
##                     Max.   :10.000    Max.   :4.00000    Max.   :10.000
##  Vehicles involved   Longitude         Latitude
##  Min.   : 0.000     Min.   :2.092    Min.   :41.32
##  1st Qu.: 2.000     1st Qu.:2.145    1st Qu.:41.39
##  Median : 2.000     Median :2.163    Median :41.40
##  Mean   : 1.921     Mean   :2.163    Mean   :41.40
##  3rd Qu.: 2.000     3rd Qu.:2.180    3rd Qu.:41.41
##  Max.   :14.000     Max.   :2.223    Max.   :41.47
## [1] "----------------------------------------------------------"
## [1] "air_quality_Nov2017.csv"
##    Station            Air Quality          Longitude         Latitude
##  Length:5744        Length:5744        Min.   :2.115    Min.   :   41.4
##  Class :character   Class :character   1st Qu.:2.131    1st Qu.:   41.4
```

```
##  Mode  :character   Mode  :character   Median :2.151   Median :    41.4
##                                         Mean   :2.152   Mean   : 13877.8
##                                         3rd Qu.:2.162   3rd Qu.:    41.4
##                                         Max.   :2.204   Max.   :414261.0
##
##     O3 Hour           O3 Quality          O3 Value         NO2 Hour
##  Length:5744        Length:5744        Min.   :  1.00   Length:5744
##  Class :character   Class :character   1st Qu.: 14.00   Class :character
##  Mode  :character   Mode  :character   Median : 34.00   Mode  :character
##                                        Mean   : 34.08
##                                        3rd Qu.: 52.00
##                                        Max.   :100.00
##                                        NA's   :1643
##   NO2 Quality         NO2 Value         PM10 Hour          PM10 Quality
##  Length:5744        Min.   :  1.00   Length:5744        Length:5744
##  Class :character   1st Qu.: 17.00   Class :character   Class :character
##  Mode  :character   Median : 33.00   Mode  :character   Mode  :character
##                     Mean   : 35.74
##                     3rd Qu.: 52.00
##                     Max.   :117.00
##                     NA's   :284
##    PM10 Value       Generated           Date Time
##  Min.   : 2.00   Length:5744        Min.   :1.541e+09
##  1st Qu.:10.00   Class :character   1st Qu.:1.542e+09
##  Median :15.00   Mode  :character   Median :1.542e+09
##  Mean   :16.59                      Mean   :1.542e+09
##  3rd Qu.:22.00                      3rd Qu.:1.543e+09
##  Max.   :44.00                      Max.   :1.544e+09
##  NA's   :2097
## [1] "----------------------------------------------------------"
## [1] "air_stations_Nov2017.csv"
##    Station            Longitude         Latitude         Ubication
##  Length:8           Min.   :2.115   Min.   :41.38   Length:8
##  Class :character   1st Qu.:2.130   1st Qu.:41.39   Class :character
##  Mode  :character   Median :2.151   Median :41.39   Mode  :character
##                     Mean   :2.152   Mean   :41.40
##                     3rd Qu.:2.162   3rd Qu.:41.41
##                     Max.   :2.204   Max.   :41.43
##  District Name      Neighborhood Name
##  Length:8           Length:8
##  Class :character   Class :character
##  Mode  :character   Mode  :character
##
##
##
## [1] "----------------------------------------------------------"
## [1] "births.csv"
##       Year       District Code    District Name      Neighborhood Code
##  Min.   :2013   Min.   : 1.000   Length:734         Min.   : 1.00
##  1st Qu.:2014   1st Qu.: 4.000   Class :character   1st Qu.:19.00
##  Median :2015   Median : 7.000   Mode  :character   Median :37.00
##  Mean   :2015   Mean   : 6.752                      Mean   :37.34
##  3rd Qu.:2016   3rd Qu.: 8.000                      3rd Qu.:55.75
##  Max.   :2017   Max.   :99.000                      Max.   :99.00
```

```
##  Neighborhood Name      Gender                 Number
##  Length:734         Length:734         Min.   :  0.00
##  Class :character   Class :character   1st Qu.: 43.25
##  Mode  :character   Mode  :character   Median : 84.50
##                                        Mean   : 91.75
##                                        3rd Qu.:129.25
##                                        Max.   :283.00
## [1] "----------------------------------------------------------"
## [1] "bus_stops.csv"
##     Code              Transport          Longitude         Latitude
##  Length:3162        Length:3162        Min.   :2.056   Min.   :41.32
##  Class :character   Class :character   1st Qu.:2.141   1st Qu.:41.39
##  Mode  :character   Mode  :character   Median :2.161   Median :41.41
##                                        Mean   :2.160   Mean   :41.41
##                                        3rd Qu.:2.180   3rd Qu.:41.42
##                                        Max.   :2.222   Max.   :41.47
##    Bus.Stop            District.Name      Neighborhood.Name
##  Length:3162        Length:3162        Length:3162
##  Class :character   Class :character   Class :character
##  Mode  :character   Mode  :character   Mode  :character
##
##
##
## [1] "----------------------------------------------------------"
## [1] "deaths.csv"
##       Year       District.Code    District.Name       Neighborhood.Code
##  Min.   :2015   Min.   : 1.000   Length:4599        Min.   : 1
##  1st Qu.:2015   1st Qu.: 4.000   Class :character   1st Qu.:19
##  Median :2016   Median : 7.000   Mode  :character   Median :37
##  Mean   :2016   Mean   : 6.247                      Mean   :37
##  3rd Qu.:2017   3rd Qu.: 8.000                      3rd Qu.:55
##  Max.   :2017   Max.   :10.000                      Max.   :73
##  Neighborhood.Name      Age                Number
##  Length:4599        Length:4599        Min.   :  0.00
##  Class :character   Class :character   1st Qu.:  0.00
##  Mode  :character   Mode  :character   Median :  2.00
##                                        Mean   : 10.05
##                                        3rd Qu.: 11.00
##                                        Max.   :139.00
## [1] "----------------------------------------------------------"
## [1] "immigrants_by_nationality.csv"
##       Year       District Code   District Name       Neighborhood Code
##  Min.   :2015   Min.   : 1.0    Length:35224       Min.   : 1.00
##  1st Qu.:2015   1st Qu.: 4.0    Class :character   1st Qu.:19.00
##  Median :2016   Median : 7.0    Mode  :character   Median :37.50
##  Mean   :2016   Mean   : 7.5                       Mean   :37.84
##  3rd Qu.:2017   3rd Qu.: 8.0                       3rd Qu.:56.00
##  Max.   :2017   Max.   :99.0                       Max.   :99.00
##  Neighborhood Name  Nationality            Number
##  Length:35224       Length:35224       Min.   :  0.000
##  Class :character   Class :character   1st Qu.:  0.000
##  Mode  :character   Mode  :character   Median :  0.000
##                                        Mean   :  7.707
##                                        3rd Qu.:  2.000
```

```
##                                             Max.   :1603.000
## [1] "------------------------------------------------------------"
## [1] "immigrants_emigrants_by_age.csv"
##       Year        District Code   District Name      Neighborhood Code
##  Min.   :2015   Min.   : 1.0   Length:4662        Min.   : 1.00
##  1st Qu.:2015   1st Qu.: 4.0   Class :character   1st Qu.:19.00
##  Median :2016   Median : 7.0   Mode  :character   Median :37.50
##  Mean   :2016   Mean   : 7.5                      Mean   :37.84
##  3rd Qu.:2017   3rd Qu.: 8.0                      3rd Qu.:56.00
##  Max.   :2017   Max.   :99.0                      Max.   :99.00
##  Neighborhood Name      Age             Immigrants        Emigrants
##  Length:4662        Length:4662      Min.   :   0.00   Min.   :  0.00
##  Class :character   Class :character 1st Qu.:   5.00   1st Qu.:  6.00
##  Mode  :character   Mode  :character Median :  21.00   Median : 18.00
##                                      Mean   :  58.23   Mean   : 33.68
##                                      3rd Qu.:  64.00   3rd Qu.: 42.00
##                                      Max.   :1393.00   Max.   :451.00
## [1] "------------------------------------------------------------"
## [1] "immigrants_emigrants_by_destination.csv"
##      from              to              weight
##  Length:40          Length:40       Min.   :   12.0
##  Class :character   Class :character 1st Qu.:  203.2
##  Mode  :character   Mode  :character Median :  613.5
##                                      Mean   : 3749.3
##                                      3rd Qu.: 1605.5
##                                      Max.   :53085.0
## [1] "------------------------------------------------------------"
## [1] "immigrants_emigrants_by_destination2.csv"
##      from              to              weight
##  Length:400         Length:400      Min.   :    0.0
##  Class :character   Class :character 1st Qu.:   17.0
##  Mode  :character   Mode  :character Median :   51.0
##                                      Mean   :  374.9
##                                      3rd Qu.:  150.8
##                                      Max.   :11473.0
## [1] "------------------------------------------------------------"
## [1] "immigrants_emigrants_by_sex.csv"
##       Year        District Code    District Name      Neighborhood Code
##  Min.   :2013   Min.   : 1.000   Length:730         Min.   : 1
##  1st Qu.:2014   1st Qu.: 4.000   Class :character   1st Qu.:19
##  Median :2015   Median : 7.000   Mode  :character   Median :37
##  Mean   :2015   Mean   : 6.247                      Mean   :37
##  3rd Qu.:2016   3rd Qu.: 8.000                      3rd Qu.:55
##  Max.   :2017   Max.   :10.000                      Max.   :73
##  Neighborhood Name      Gender           Immigrants        Emigrants
##  Length:730         Length:730       Min.   :   5.0    Min.   :   3.0
##  Class :character   Class :character 1st Qu.: 251.2    1st Qu.: 178.0
##  Mode  :character   Mode  :character Median : 474.5    Median : 329.0
##                                      Mean   : 588.8    Mean   : 362.2
##                                      3rd Qu.: 784.2    3rd Qu.: 498.8
##                                      Max.   :3219.0    Max.   :1541.0
## [1] "------------------------------------------------------------"
## [1] "life_expectancy.csv"
##  Neighborhood        2006-2010        2007-2011        2008-2012
```

```
## Length:146         Min.    :74.20    Min.    :70.70    Min.    :69.50
## Class :character   1st Qu.:81.70    1st Qu.:79.50    1st Qu.:79.90
## Mode  :character   Median :83.90    Median :81.70    Median :82.00
##                    Mean    :84.93   Mean    :82.47   Mean    :82.78
##                    3rd Qu.:88.50    3rd Qu.:86.12    3rd Qu.:86.30
##                    Max.    :90.90   Max.    :88.20   Max.    :88.20
##                    NA's    :8       NA's    :6       NA's    :8
##    2009-2013        2010-2014         Gender
## Min.    :69.00   Min.    :71.30   Length:146
## 1st Qu.:80.10   1st Qu.:80.53   Class :character
## Median :82.70   Median :83.40   Mode  :character
## Mean    :83.10   Mean    :83.45
## 3rd Qu.:86.53   3rd Qu.:86.97
## Max.    :88.50   Max.    :89.80
## NA's    :6        NA's    :4
## [1] "---------------------------------------------------------"
## [1] "most_frequent_baby_names.csv"
##      Order           Name              Gender              Year
## Min.    : 1    Length:1050     Length:1050       Min.    :1996
## 1st Qu.: 7    Class :character   Class :character   1st Qu.:2001
## Median :13    Mode  :character   Mode  :character   Median :2006
## Mean    :13                                          Mean    :2006
## 3rd Qu.:19                                          3rd Qu.:2011
## Max.    :25                                          Max.    :2016
##    Frequency
## Min.    : 44.0
## 1st Qu.: 69.0
## Median : 89.0
## Mean    :101.7
## 3rd Qu.:122.0
## Max.    :348.0
## [1] "---------------------------------------------------------"
## [1] "most_frequent_names.csv"
##      Order           Name              Gender             Decade
## Min.    : 1.0   Length:1100     Length:1100       Length:1100
## 1st Qu.:13.0   Class :character   Class :character   Class :character
## Median :25.5   Mode  :character   Mode  :character   Mode  :character
## Mean    :25.5
## 3rd Qu.:38.0
## Max.    :50.0
##    Frequency
## Min.    :    39.0
## 1st Qu.:  379.0
## Median :  621.5
## Mean    : 1263.6
## 3rd Qu.: 1145.0
## Max.    :19173.0
## [1] "---------------------------------------------------------"
## [1] "population.csv"
##       Year       District.Code    District.Name     Neighborhood.Code
## Min.    :2013   Min.    : 1.000   Length:70080     Min.    : 1
## 1st Qu.:2014   1st Qu.: 4.000   Class :character   1st Qu.:19
## Median :2015   Median : 7.000   Mode  :character   Median :37
## Mean    :2015   Mean    : 6.247                    Mean    :37
```

```
##  3rd Qu.:2016   3rd Qu.: 8.000                      3rd Qu.:55
##  Max.   :2017   Max.   :10.000                      Max.   :73
##  Neighborhood.Name     Gender               Age                Number
##  Length:70080        Length:70080        Length:70080       Min.   :   0.0
##  Class :character    Class :character    Class :character   1st Qu.:  43.0
##  Mode  :character    Mode  :character    Mode  :character   Median :  92.0
##                                                             Mean   : 114.8
##                                                             3rd Qu.: 162.0
##                                                             Max.   : 777.0
## [1] "----------------------------------------------------------"
## [1] "transports.csv"
##      Code             Transport            Longitude          Latitude
##  Length:651         Length:651         Min.   :2.005    Min.   :41.30
##  Class :character   Class :character   1st Qu.:2.135    1st Qu.:41.38
##  Mode  :character   Mode  :character   Median :2.164    Median :41.40
##                                        Mean   :2.157    Mean   :41.40
##                                        3rd Qu.:2.182    3rd Qu.:41.42
##                                        Max.   :2.249    Max.   :41.50
##    Station           District.Name      Neighborhood.Name
##  Length:651         Length:651         Length:651
##  Class :character   Class :character   Class :character
##  Mode  :character   Mode  :character   Mode  :character
##
##
##
## [1] "----------------------------------------------------------"
## [1] "unemployment.csv"
##       Year          Month          District Code   District Name
##  Min.   :2013   Length:14208     Min.   : 1.0    Length:14208
##  1st Qu.:2015   Class :character   1st Qu.: 4.0    Class :character
##  Median :2016   Mode  :character   Median : 7.0    Mode  :character
##  Mean   :2015                     Mean   : 7.5
##  3rd Qu.:2016                     3rd Qu.: 8.0
##  Max.   :2017                     Max.   :99.0
##  Neighborhood Code Neighborhood Name    Gender           Demand_occupation
##  Min.   : 1.00    Length:14208       Length:14208       Length:14208
##  1st Qu.:19.00    Class :character   Class :character   Class :character
##  Median :37.50    Mode  :character   Mode  :character   Mode  :character
##  Mean   :37.84
##  3rd Qu.:56.00
##  Max.   :99.00
##     Number
##  Min.   :   0.0
##  1st Qu.: 121.0
##  Median : 296.0
##  Mean   : 446.7
##  3rd Qu.: 685.0
##  Max.   :3094.0
```