Weather Data 12.2014 - 12.2015

Eseniia Novokhatskaia

Weather Report

This is an initial 'Weather' data set, which contains an information about different weather observations from December 2014 till December 2015.

```
library(tidyverse)
```

```
## -- Attaching packages ----------------------------------------------------------------- tidyverse 1.2.1

## v ggplot2 3.2.1     v purrr   0.3.3
## v tibble  2.1.3     v dplyr   0.8.3
## v tidyr   1.0.0     v stringr 1.4.0
## v readr   1.3.1     v forcats 0.4.0

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

```
weather <- readRDS("/Users/apple/Desktop/weather.rds")
head(weather)
```

```
##   X year month           measure X1 X2 X3 X4 X5 X6 X7 X8 X9 X10 X11 X12
## 1 1 2014    12  Max.TemperatureF 64 42 51 43 42 45 38 29 49  48  39  39
## 2 2 2014    12 Mean.TemperatureF 52 38 44 37 34 42 30 24 39  43  36  35
## 3 3 2014    12  Min.TemperatureF 39 33 37 30 26 38 21 18 29  38  32  31
## 4 4 2014    12    Max.Dew.PointF 46 40 49 24 37 45 36 28 49  45  37  28
## 5 5 2014    12    MeanDew.PointF 40 27 42 21 25 40 20 16 41  39  31  27
## 6 6 2014    12     Min.DewpointF 26 17 24 13 12 36 -3  3 28  37  27  25
##   X13 X14 X15 X16 X17 X18 X19 X20 X21 X22 X23 X24 X25 X26 X27 X28 X29 X30
## 1  42  45  42  44  49  44  37  36  36  44  47  46  59  50  52  52  41  30
## 2  37  39  37  40  45  40  33  32  33  39  45  44  52  44  45  46  36  26
## 3  32  33  32  35  41  36  29  27  30  33  42  41  44  37  38  40  30  22
## 4  28  29  33  42  46  34  25  30  30  39  45  46  58  31  34  42  26  10
## 5  26  27  29  36  41  30  22  24  27  34  42  44  43  29  31  35  20   4
## 6  24  25  27  30  32  26  20  20  25  25  37  41  29  28  29  27  10  -6
##   X31
## 1  30
## 2  25
## 3  20
## 4   8
## 5   5
## 6   1
```

The problem with initial data set is you cannot really follow the logic of the data gathering and sorting, so it is needed to be fixed.

First of all, we clearly see that 'X-' are days. We do not really need Xs, so we put numbers of days instead.

```
weather$X <- NULL
names(weather)[4:34] <- '1':'31'
head(weather)
```

```
##   year month           measure  1  2  3  4  5  6  7  8  9 10 11 12 13 14
## 1 2014    12  Max.TemperatureF 64 42 51 43 42 45 38 29 49 48 39 39 42 45
## 2 2014    12 Mean.TemperatureF 52 38 44 37 34 42 30 24 39 43 36 35 37 39
## 3 2014    12  Min.TemperatureF 39 33 37 30 26 38 21 18 29 38 32 31 32 33
## 4 2014    12    Max.Dew.PointF 46 40 49 24 37 45 36 28 49 45 37 28 28 29
## 5 2014    12    MeanDew.PointF 40 27 42 21 25 40 20 16 41 39 31 27 26 27
## 6 2014    12     Min.DewpointF 26 17 24 13 12 36 -3  3 28 37 27 25 24 25
##   15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31
## 1 42 44 49 44 37 36 36 44 47 46 59 50 52 52 41 30 30
## 2 37 40 45 40 33 32 33 39 45 44 52 44 45 46 36 26 25
## 3 32 35 41 36 29 27 30 33 42 41 44 37 38 40 30 22 20
## 4 33 42 46 34 25 30 30 39 45 46 58 31 34 42 26 10  8
## 5 29 36 41 30 22 24 27 34 42 44 43 29 31 35 20  4  5
## 6 27 30 32 26 20 20 25 25 37 41 29 28 29 27 10 -6  1
```

Then, we need to sort out the date of the observation. It is easier to unite day, month and year in one column and call it "date".

```
weather <- weather %>%
  gather(day, data, '1': '31') %>%
  unite (Date, year, month, day)
head(weather)
```

```
##        Date           measure data
## 1 2014_12_1  Max.TemperatureF   64
## 2 2014_12_1 Mean.TemperatureF   52
## 3 2014_12_1  Min.TemperatureF   39
## 4 2014_12_1    Max.Dew.PointF   46
## 5 2014_12_1    MeanDew.PointF   40
## 6 2014_12_1     Min.DewpointF   26
```

Now we need to spread all the measures in different columns, so we have different dates as rows in our data set and different measures in columns.

```
weather <- weather %>%
  spread(key = measure, value = data)
head(weather)
```

```
##         Date CloudCover    Events Max.Dew.PointF Max.Gust.SpeedMPH
## 1  2014_12_1          6      Rain             46                29
## 2 2014_12_10          8      Rain             45                29
## 3 2014_12_11          8 Rain-Snow             37                28
## 4 2014_12_12          7      Snow             28                21
## 5 2014_12_13          5                       28                23
## 6 2014_12_14          4                       29                20
##   Max.Humidity Max.Sea.Level.PressureIn Max.TemperatureF
## 1           74                    30.45               64
## 2          100                    29.58               48
```

```
## 3           92                   29.81               39
## 4           85                   29.88               39
## 5           75                   29.86               42
## 6           82                   29.91               45
##   Max.VisibilityMiles Max.Wind.SpeedMPH Mean.Humidity
## 1                  10                22            63
## 2                  10                23            95
## 3                  10                21            87
## 4                  10                16            75
## 5                  10                17            65
## 6                  10                15            68
##   Mean.Sea.Level.PressureIn Mean.TemperatureF Mean.VisibilityMiles
## 1                     30.13                52                   10
## 2                      29.5                43                    3
## 3                     29.61                36                    7
## 4                     29.85                35                   10
## 5                     29.82                37                   10
## 6                     29.83                39                   10
##   Mean.Wind.SpeedMPH MeanDew.PointF Min.DewpointF Min.Humidity
## 1                 13             40            26           52
## 2                 13             39            37           89
## 3                 13             31            27           82
## 4                 11             27            25           64
## 5                 12             26            24           55
## 6                 10             27            25           53
##   Min.Sea.Level.PressureIn Min.TemperatureF Min.VisibilityMiles
## 1                    30.01               39                   10
## 2                    29.43               38                    1
## 3                    29.44               32                    1
## 4                    29.81               31                    7
## 5                    29.78               32                   10
## 6                    29.78               33                   10
##   PrecipitationIn WindDirDegrees
## 1            0.01            268
## 2            0.28            357
## 3            0.02            230
## 4               T            286
## 5               T            298
## 6            0.00            306
```

Weather data set already looks understandable, however, we can make it even better by sorting the data set by date and bringing all measure names to one style.

```
weather <- weather %>%
  mutate(Date = as.Date(Date, format = "%Y_%m_%d")) %>%
  arrange(Date) %>%
  rename(
    Cloud.CoverOktas = CloudCover,
    Max.HumidityPercent = Max.Humidity,
    Mean.HumidityPercent = Mean.Humidity,
    Mean.Dew.PointF = MeanDew.PointF,
    Min.Dew.PointF = Min.DewpointF,
    Min.HumidityPercent = Min.Humidity,
    Wind.DirDegrees = WindDirDegrees,
```

```
  )
head(weather)
```

```
##         Date Cloud.CoverOktas    Events Max.Dew.PointF Max.Gust.SpeedMPH
## 1 2014-12-01                6      Rain            46                29
## 2 2014-12-02                7 Rain-Snow            40                29
## 3 2014-12-03                8      Rain            49                38
## 4 2014-12-04                3                      24                33
## 5 2014-12-05                5      Rain            37                26
## 6 2014-12-06                8      Rain            45                25
##   Max.HumidityPercent Max.Sea.Level.PressureIn Max.TemperatureF
## 1                  74                    30.45               64
## 2                  92                    30.71               42
## 3                 100                     30.4               51
## 4                  69                    30.56               43
## 5                  85                    30.68               42
## 6                 100                    30.42               45
##   Max.VisibilityMiles Max.Wind.SpeedMPH Mean.HumidityPercent
## 1                  10                22                   63
## 2                  10                24                   72
## 3                  10                29                   79
## 4                  10                25                   54
## 5                  10                22                   66
## 6                  10                22                   93
##   Mean.Sea.Level.PressureIn Mean.TemperatureF Mean.VisibilityMiles
## 1                     30.13                52                   10
## 2                     30.59                38                    8
## 3                     30.07                44                    5
## 4                     30.33                37                   10
## 5                     30.59                34                   10
## 6                     30.24                42                    4
##   Mean.Wind.SpeedMPH Mean.Dew.PointF Min.Dew.PointF Min.HumidityPercent
## 1                 13              40             26                 52
## 2                 15              27             17                 51
## 3                 12              42             24                 57
## 4                 12              21             13                 39
## 5                 10              25             12                 47
## 6                  8              40             36                 85
##   Min.Sea.Level.PressureIn Min.TemperatureF Min.VisibilityMiles
## 1                    30.01               39                  10
## 2                     30.4               33                   2
## 3                    29.87               37                   1
## 4                    30.09               30                  10
## 5                    30.45               26                   5
## 6                    30.16               38                   0
##   PrecipitationIn Wind.DirDegrees
## 1            0.01             268
## 2            0.10              62
## 3            0.44             254
## 4            0.00             292
## 5            0.11              61
## 6            1.09             313
```