Assignment 2

khushboo.satpute

13/11/2019

Q1. Create R Markdown

Q2 Load in the data as an object called DublinAirport

```
DublinAirport<-read.csv("mly532.csv", header=TRUE, sep=",", skip = 19)
str(DublinAirport)
```

```
## 'data.frame':    931 obs. of  12 variables:
##  $ year : int  1941 1941 1942 1942 1942 1942 1942 1942 1942 1942 ...
##  $ month: int  11 12 1 2 3 4 5 6 7 8 ...
##  $ meant: num  6.9 6.5 4.3 2.9 6.3 8.4 10.4 13.1 14.6 14.9 ...
##  $ maxtp: num  14 12.7 11.9 11.6 16.2 16.2 20.9 24.1 22.2 22.3 ...
##  $ mintp: num  -3.1 -3.6 -3.1 -4.3 -6.1 0.8 1.8 1.4 7.2 6.7 ...
##  $ mnmax: num  9.9 9.1 6.9 5.8 9.4 11.9 14.4 18 18.9 18.4 ...
##  $ mnmin: num  3.9 3.9 1.7 0 3.2 4.9 6.3 8.2 10.4 11.4 ...
##  $ rain : num  67.2 41.7 91.9 25.8 76.4 ...
##  $ gmin : num  -5.7 -7.6 -9.5 -10.7 -8.3 -0.4 -0.7 -0.9 2.4 4.6 ...
##  $ wdsp : num  12 12.5 13.1 9 10.7 15.1 12 9.4 13.4 10.8 ...
##  $ maxgt: int  NA NA NA NA NA NA NA NA NA NA ...
##  $ sun  : num  56.1 46.1 72.8 51.4 73.9 ...
```

Q3. Transform the column months.

```
DublinAirport$month<-factor(DublinAirport$month, labels = month.name)
head(DublinAirport)
```

```
##   year    month meant maxtp mintp mnmax mnmin rain  gmin wdsp maxgt   sun
## 1 1941 November   6.9  14.0  -3.1   9.9   3.9 67.2  -5.7 12.0    NA  56.1
## 2 1941 December   6.5  12.7  -3.6   9.1   3.9 41.7  -7.6 12.5    NA  46.1
## 3 1942  January   4.3  11.9  -3.1   6.9   1.7 91.9  -9.5 13.1    NA  72.8
## 4 1942 February   2.9  11.6  -4.3   5.8   0.0 25.8 -10.7  9.0    NA  51.4
## 5 1942    March   6.3  16.2  -6.1   9.4   3.2 76.4  -8.3 10.7    NA  73.9
## 6 1942    April   8.4  16.2   0.8  11.9   4.9 36.9  -0.4 15.1    NA 185.4
```

Q4. Using aggregate function to compute which month has on average the highest and the lowest Precipitation Amount.

```
Agg_dub_rain = aggregate(DublinAirport$rain,by = list(DublinAirport$month) , FUN = mean,na.rm= TRUE)
message("The Minimum Average Precipitation Amount month is")
```

```
## The Minimum Average Precipitation Amount month is
```

```
as.character(Agg_dub_rain$Group.1[Agg_dub_rain$x == max(Agg_dub_rain$x)])
```

```
## [1] "December"
```

```
message("The Minimum average Precipitation amount Month is")
```

```
## The Minimum average Precipitation amount Month is
```

```
as.character(Agg_dub_rain$Group.1[Agg_dub_rain$x == min(Agg_dub_rain$x)])
```

```
## [1] "February"
```

Q5.Adding column which contains a factor indicating the season:

```
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
DublinAirport<-mutate(DublinAirport,season = case_when( month %in% c("December","January","February") ~
                          month %in% c("March","April","May")    ~ "Spring",
                          month %in% c("June","July","August") ~ "Summer",
                          month %in% c("September","October","November") ~ "Autumn",
                         ))
DublinAirport$season<-factor(DublinAirport$season,labels=c("Winter","Spring","Summer","Autumn"))
head(DublinAirport)
```

```
##   year    month meant maxtp mintp mnmax mnmin rain  gmin wdsp maxgt   sun
## 1 1941 November   6.9  14.0  -3.1   9.9   3.9 67.2  -5.7 12.0    NA  56.1
## 2 1941 December   6.5  12.7  -3.6   9.1   3.9 41.7  -7.6 12.5    NA  46.1
## 3 1942  January   4.3  11.9  -3.1   6.9   1.7 91.9  -9.5 13.1    NA  72.8
## 4 1942 February   2.9  11.6  -4.3   5.8   0.0 25.8 -10.7  9.0    NA  51.4
## 5 1942    March   6.3  16.2  -6.1   9.4   3.2 76.4  -8.3 10.7    NA  73.9
## 6 1942    April   8.4  16.2   0.8  11.9   4.9 36.9  -0.4 15.1    NA 185.4
##   season
## 1 Winter
## 2 Autumn
## 3 Autumn
## 4 Autumn
## 5 Spring
## 6 Spring
```

Q6. Assign to the DublinAiport object the classes WeatherData and data.frame

```
class(DublinAirport)<- c('weatherdata','data.frame')
class(DublinAirport)
```

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

Q7. The S3 summary method for an object of class WeatherData

```
Summary.WeatherData <- function(val){
  val %>%
    group_by(season) %>%
    select(rain, maxtp, mintp, maxgt) %>%
    summarise_all(funs(mean, max, sd), na.rm = TRUE)
}
Summary.WeatherData(DublinAirport)
```

```
## Adding missing grouping variables: `season`

## Warning: funs() is soft deprecated as of dplyr 0.8.0
## Please use a list of either functions or lambdas:
##
##   # Simple named list:
##   list(mean = mean, median = median)
##
##   # Auto named with `tibble::lst()`:
##   tibble::lst(mean, median)
##
##   # Using lambdas
##   list(~ mean(., trim = .2), ~ median(., na.rm = TRUE))
## This warning is displayed once per session.

## # A tibble: 4 x 13
##   season rain_mean maxtp_mean mintp_mean maxgt_mean rain_max maxtp_max
##   <fct>      <dbl>      <dbl>      <dbl>      <dbl>    <dbl>     <dbl>
## 1 Winter      70.2       18.0       1.26       47.2     186.      25.1
## 2 Spring      53.5       17.2     -0.587       45.6     152.      23.5
## 3 Summer      63.9       23.0       5.93       39.5     190.      28.7
## 4 Autumn      63.9       12.7     -3.19        53.0     217       17.1
## # ... with 6 more variables: mintp_max <dbl>, maxgt_max <int>,
## #   rain_sd <dbl>, maxtp_sd <dbl>, mintp_sd <dbl>, maxgt_sd <dbl>
```

Q8. S3 plot method for the class WeatherData that produces the following plots.

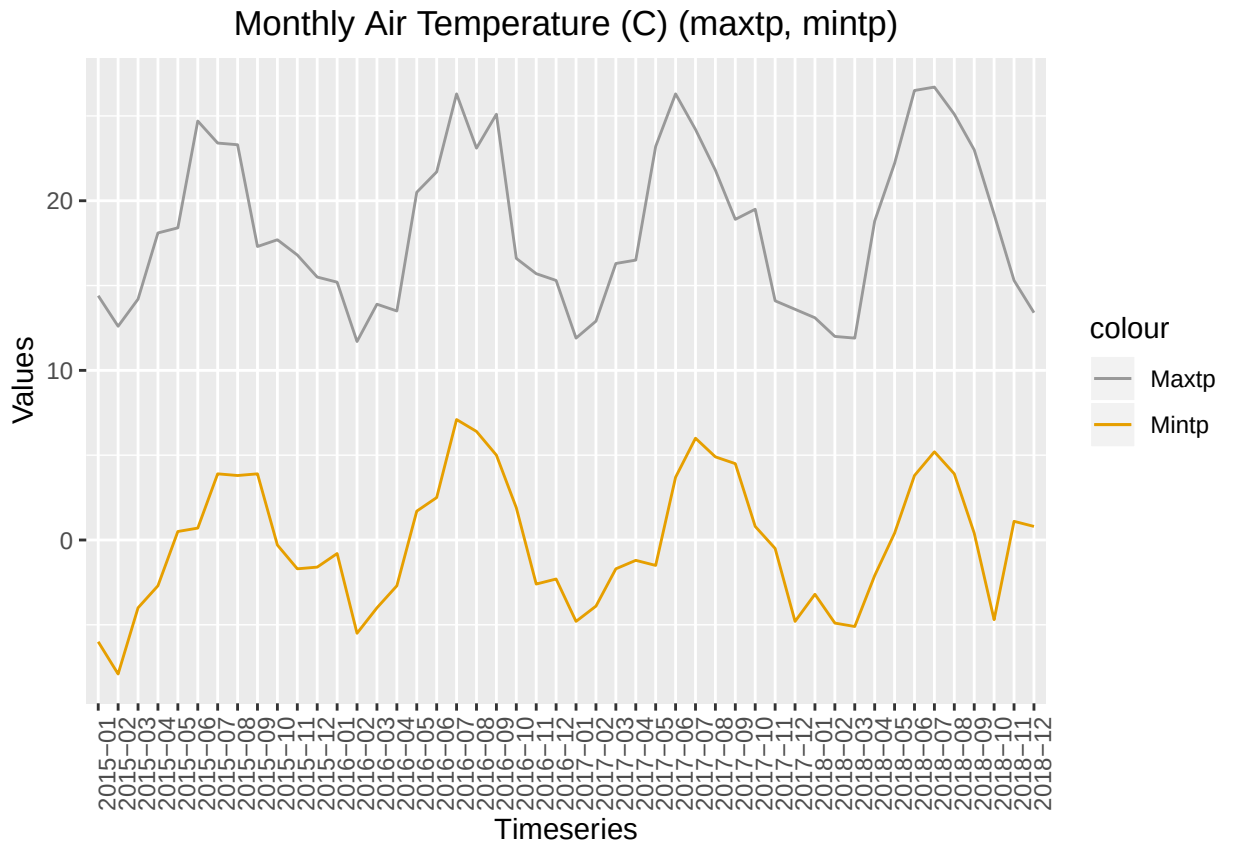
#(1) Plot of the monthly Air Temperature (C) (maxtp, mintp). #(2) Plot of the Precipitation Amount (mm) (rain). #(3) Plot of the Highest Gust (knot) (maxgt).

```r
library(gridExtra)
```

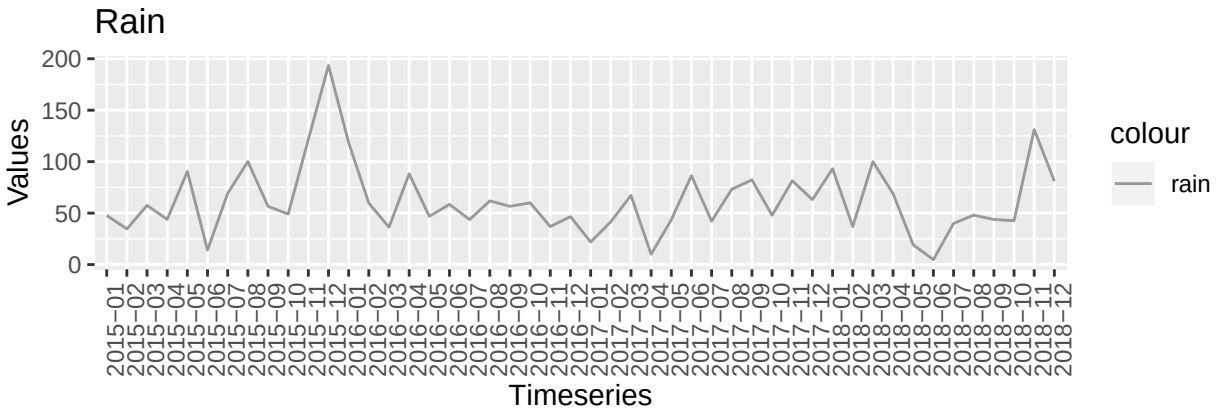
```
##
## Attaching package: 'gridExtra'

## The following object is masked from 'package:dplyr':
##
##     combine
```

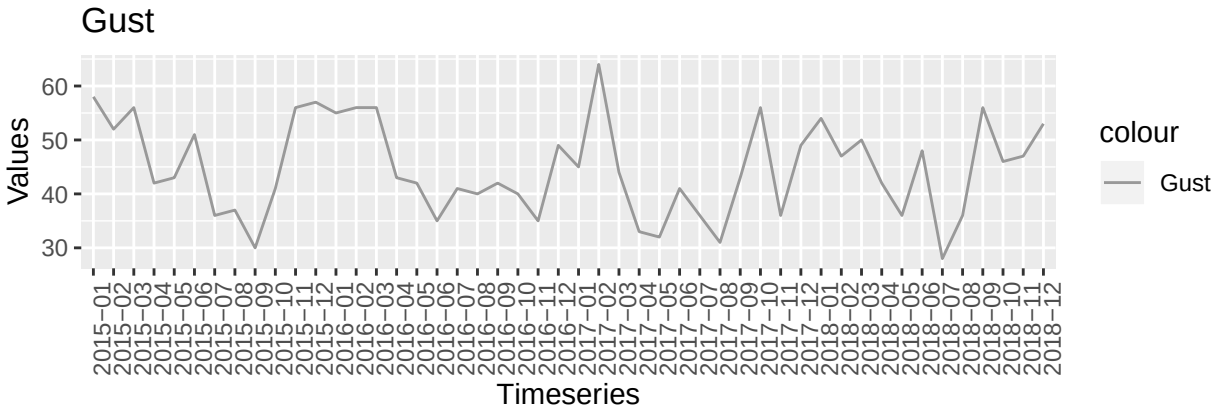
```r
library(ggplot2)
plot.WeatherData <- function(plotdata, start_year=2015, end_year=2018,
                             plot1 = TRUE, plot2 = TRUE, plot3 = TRUE )
  {

  df_sub <- subset(DublinAirport,DublinAirport$year >= 2015 & DublinAirport$year <= 2018)
  data <- data.frame(Timeseries = with(df_sub,
                                       sprintf("%d-%02d", df_sub$year, df_sub$month)),
                     mintp = df_sub$mintp,maxtp = df_sub$maxtp,rain = df_sub$rain,
                     maxgt = df_sub$maxgt)

  plt <- ggplot(data, aes(x= Timeseries)) +
    theme(axis.text.x = element_text(angle = 90, hjust = 1), legend.position="right") +
    labs(y = "Values") + scale_color_manual(values=c('#999999','#E69F00'))

  if(plot1 == TRUE & plot2 == FALSE & plot3 == FALSE ){
    plt <- plt +
      geom_line(aes(y = mintp, group = 1, color = "Mintp")) +
      geom_line(aes(y = maxtp, group = 1, color = "Maxtp")) +
        ggtitle("Monthly Air Temperature (C) (maxtp, mintp)")+theme(plot.title = element_text(hjust = 0
    plt
  }
  else if(plot1 == FALSE & plot2 == TRUE & plot3 == FALSE ){
    plt <- plt +
      geom_line(aes(y = rain, group = 1, color = "Rain")) +
      ggtitle("Monthly Rain in mm")+theme(plot.title = element_text(hjust = 0.5))
    plt
  }
  else if(plot1 == FALSE & plot2 == FALSE & plot3 == TRUE ){
    plt <- plt +
      geom_line(aes(y = maxgt, group = 1, color = "Gust")) +
      ggtitle("Monthly Gust")+theme(plot.title = element_text(hjust = 0.5))
    plt
  }
  else if(plot1 == TRUE & plot2 == TRUE & plot3 == FALSE ){
    plt_1 <- plt +
      geom_line(aes(y = mintp, group = 1, color = "Mintp")) +
      geom_line(aes(y = maxtp, group = 1, color = "Maxtp"))  +
      labs( title = 'Air Temperature')
    plt_2 <- plt +
      geom_line(aes(y = maxgt, group = 1, color = "Rain")) +
      labs( title = 'Rain')
    grid.arrange(plt_1,plt_2)
  }
```

```
  else if(plot1 == TRUE & plot2 == FALSE & plot3 == TRUE ){
    plt_1 <- plt +
      geom_line(aes(y = mintp, group = 1, color = "Mintp")) +
      geom_line(aes(y = maxtp, group = 1, color = "Maxtp")) +
      labs( title = 'Air Temperature')
    plt_2 <- plt +
      geom_line(aes(y = maxgt, group = 1, color = "Gust")) +
      labs( title = 'Gust')
    grid.arrange(plt_1,plt_2)
  }
  else if(plot1 == FALSE & plot2 == TRUE & plot3 == TRUE ){
    plt_1 <- plt +
      geom_line(aes(y = rain, group = 1, color = "rain")) +
      labs( title = 'Rain')
    plt_2 <- plt +
      geom_line(aes(y = maxgt, group = 1, color = "Gust")) +
      labs( title = 'Gust')
    grid.arrange(plt_1,plt_2)
  }else{
    plt_1 <- plt +
      geom_line(aes(y = mintp, group = 1, color = "Mintp")) +
      geom_line(aes(y = maxtp, group = 1, color = "Maxtp")) +
      labs( title = 'Air Temperature')
    plt_2 <- plt +
      geom_line(aes(y = rain, group = 1, color = "Rain")) +
      labs( title = 'Rain')
    plt_3 <- plt +
      geom_line(aes(y = maxgt, group = 1, color = "Gust")) +
      labs( title = 'Gust')
    grid.arrange(plt_1,plt_2,plt_3)
  }
}
# Plot 1
plot.WeatherData(DublinAirport, 2015, 2017, TRUE, FALSE, FALSE)
```

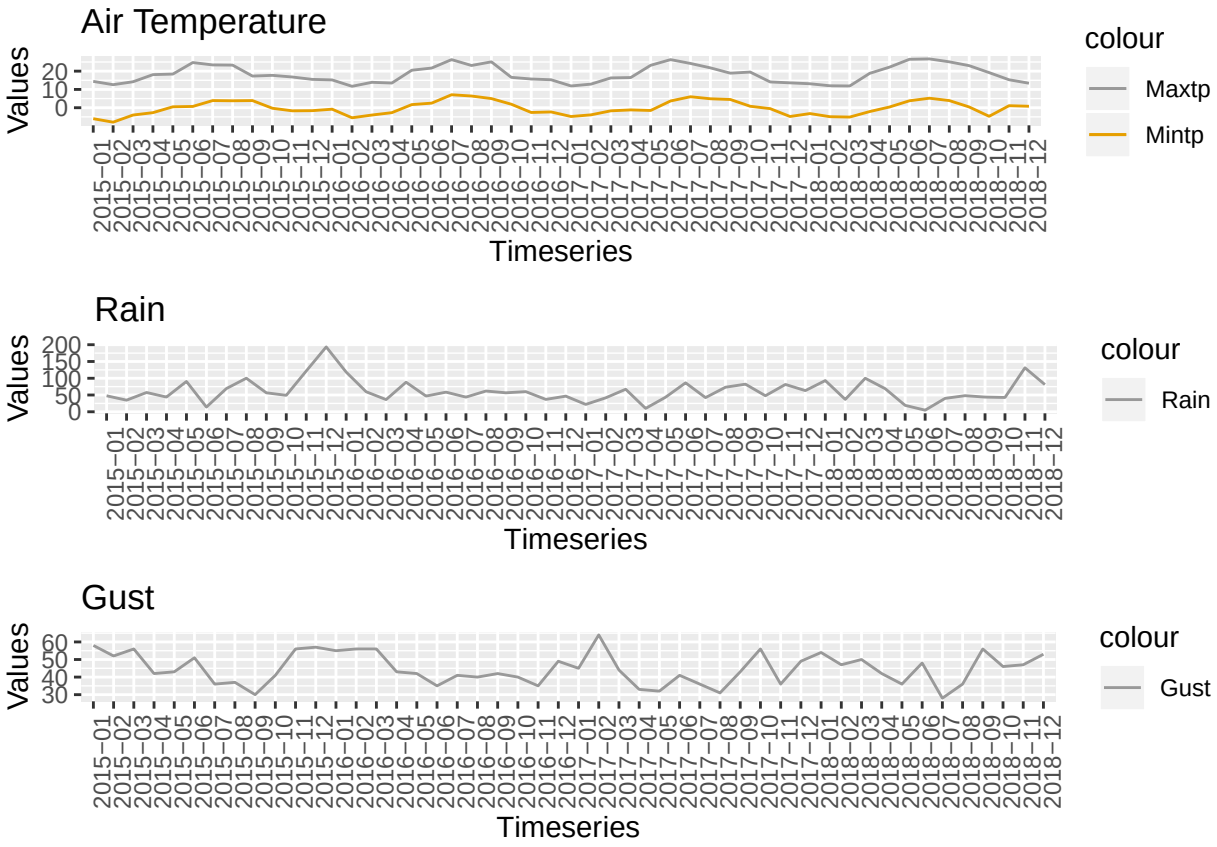



```
# Plot 2
plot.WeatherData(DublinAirport, 1990, 2000, FALSE, TRUE, TRUE)
```






```
# Plot 3
plot.WeatherData(DublinAirport )
```


Air Temperature
Values
20
10
0
Timeseries
colour
Maxtp
Mintp
Rain
Values
200
150
100
50
0
Timeseries
colour
Rain
Gust
Values
60
50
40
30
Timeseries
colour
Gust