Report

```
#=====================================
# Init Env
#=====================================
rm(list=ls())

#=====================================
# Set Env
#=====================================
# Sys.setlocale("LC_ALL", "C")
# options(encoding = "UTF-8")
# Sys.setenv(LANG = "ko_KR.UTF-8")

# Sys.setlocale("LC_ALL", "Korean")
# options(encoding = "UTF-8")
# Sys.setenv(LANG = "ko_KR.UTF-8")

# Sys.setlocale("LC_ALL", "English")
# options(encoding = "UTF-8")
# Sys.setenv(LANG = "en_US.UTF-8")


globalVar = new.env()

globalVar$optDig = 10
globalVar$memLimit = 9999999999999

# config
# globalVar$config = getwd()
globalVar$config = "E:/04. TalentPlatform/Github/TalentPlatform-R"
globalVar$inpConfig = paste(globalVar$config, 'INPUT', 'BDWIDE', sep = '/')
globalVar$figConfig = paste(globalVar$config, 'FIG', 'BDWIDE', sep = '/')
globalVar$outConfig = paste(globalVar$config, 'OUTPUT', 'BDWIDE', sep = '/')
globalVar$logConfig = paste(globalVar$config, 'LOG', 'BDWIDE', sep = '/')
globalVar$mapConfig = paste(globalVar$config, 'CONFIG', 'MAP_INFO', sep = '/')
globalVar$systemConfig = paste(globalVar$config, 'CONFIG', 'system.cfg', sep = '/')
globalVar$seleniumConfig = paste(globalVar$config, 'CONFIG', 'selenium', sep = '/')

fnHeatIndex = function (temp, rh) {

  temp = (temp * 1.8) + 32

  alpha = 61 + ((temp - 68) * 1.2) + (rh * 0.094)
  hi = 0.5*(temp + alpha)

  if (hi > 79) {
    hi = -42.379 + 2.04901523 * temp + 10.14333127 * rh - 0.22475541 * temp * rh - 0.00683783 * (temp^2) - 0.05481717 * (rh^2) + 0.00122874 * (temp^2) * rh + 0.00085282 * temp * (rh^2) - 0.00000199 * (temp^2) * (rh^2)
    if (rh <= 13 && temp >= 80 && temp <= 112) {
      adjustment1 = (13 - rh) / 4
      adjustment2 = sqrt((17 - abs(temp - 95))/17)
      total.adjustment = adjustment1 * adjustment2
      hi = hi - total.adjustment
    } else if (rh > 85 && temp >= 80 && temp <= 87) {
      adjustment1 = (rh - 85) / 10
      adjustment2 = (87 - temp) / 5
      total.adjustment = adjustment1 * adjustment2
      hi = hi + total.adjustment
    }
  }

  heatIndex = (hi - 32) / 1.8
  return(heatIndex)
}


fnHumidIndex = function(temp, rh) {

  vp = rh / 100 * 6.105 * exp(17.27 * temp / (237.7 + temp))
  humidIndex = temp + 0.5555 * (vp - 10)

  return(humidIndex)
}

fnAppTempIndex = function(temp, rh, ws) {

  vp = rh / 100 * 6.105 * exp(17.27 * temp / (237.7 + temp))
  appTempIndex = temp + 0.33 * vp - 0.7 * ws + 4.0

  return(appTempIndex)
}

fnAppTempRadIndex = function(temp, rh, ws, sr) {

  sr = sr * 86400 / (10^6)
  alb = 0.2
  vp = rh / 100 * 6.105 * exp(17.27 * temp / (237.7 + temp))
  appTempRadIndex = temp + 0.33 * vp - 0.7 * ws + 0.7 * (sr * (1 - alb) )

  return(appTempRadIndex)
}

fnWetBulbGolbalTempIndex = function(temp, rh, ws, sr) {

  sr = sr * (10^3) / 86400
  wetBulbGolbalTempIndex = 0.735 * temp + 0.0374 * rh + 0.00292 * temp * rh + 7.619 * sr - 4.557 * (sr^2) - 0.0572 * ws - 4.064

  return(wetBulbGolbalTempIndex)
}

fnStats = function(X, Y) {

  if (length(X) < 1) { return( sprintf("%s", "X 값 없음") ) }
  if (length(Y) < 1) { return( sprintf("%s", "Y 값 없음") ) }

  slope = coef(lm(Y~X))[2]
  interp = coef(lm(Y~X))[1]
  mean_X = mean(X, na.rm=TRUE)
  mean_Y = mean(Y, na.rm=TRUE)
  sd_x = sd(X, na.rm=TRUE)
  sd_y = sd(Y, na.rm=TRUE)
  number = length(X)
  bias = mean(X-Y, na.rm=TRUE)
  rbias = (bias/mean(Y, na.rm=TRUE))*100.0
  # rbias = (bias/mean(X, na.rm=TRUE))*100.0
  rmse = sqrt(mean((X-Y)^2, na.rm=TRUE))
  rrmse = (rmse/mean(Y, na.rm=TRUE))*100.0
  # rrmse = (rmse/mean(X, na.rm=TRUE))*100.0
  r = cor(X, Y)^2
  # r = cor(X, Y)
  diff_mean = mean(X-Y, na.rm=TRUE)
  diff_sd = sd(X-Y, na.rm=TRUE)
  per_diff_mean = mean((X-Y)/Y, na.rm=TRUE)*100.0
  return( c(slope, interp, mean_X, mean_Y, sd_x, sd_y, number, bias, rbias, rmse, rrmse, r, diff_mean, diff_sd, per_diff_mean) )
}



fnCalib = function(X, Y) {
  factor = seq(0, 2, by=0.0001)

  actual = X
  forecast = Y

  # RMSE Fitting
  RMSE = lapply( 1:length(factor), function(i) sqrt(mean( (forecast- (actual*factor[i]) )^2 )))
  RMSE = unlist(RMSE)

  # plot(RMSE)
  ind = which(RMSE == min(RMSE))

  # Best factor
  newFactor = factor[[ind]]

  return(c(newFactor))
}

#=====================================
# Set Data
#=====================================
options(digits = globalVar$optDig)
options(java.parameters = "-Xmx8192m")
memory.limit(size = globalVar$memLimit)
```

```
## [1] 1e+13
```

```
# 패키지 업데이트
# update.packages(ask = FALSE)

# 주석 단계
# ====
# ****
# ++++

# font, colorbar
# font = "New Century Schoolbook"
font = "Palatino Linotype"
# font = "Time Roman"
# font = "Comic Sans MS"
# font = "Helvetica"
# font = "NanumBarunGothic"
# font = "Times New Roman"

# cbSpectral = rev(brewer.pal(11, "Spectral"))
cbViridis =  viridis::viridis(11)
cbMatlab =  colorRamps::matlab.like(11)
cbMatlab2 =  colorRamps::matlab.like2(11)
cbDiverge =  colorspace::diverge_hcl(11)
cbPlasma =  rev(viridis::plasma(11))

#===============================================================================================
# Routine : Main R program
#
# Purpose : 재능상품 오투잡
#
# Author : 해솔
#
# Revisions: V1.0 May 28, 2020 First release (MS. 해솔)
#===============================================================================================

serviceName = "LSH0079"

library(data.table)
library(readr)
library(tidyverse)
library(ggplot2)
library(sf)
library(dplyr)
library(ggmap)
library(readxl)
library(ggrepel)
library(metR)
library(solarPos)
library(Metrics)
library(tidyverse)
library(weathermetrics)
library(data.table)
library(callr)
library(devtools)
library(GGally)
library(factoextra)
library(tidyverse)
library(colorRamps)
library(akima)
library(GGally)
library(tidyverse)
library(factoextra)
library(gridExtra)
library(ggcorrplot)
library(MAT)
library(moments)
library(RColorBrewer)
library(dplyr)
library(zoo)
library(ggplot2)
library(gstat)
library(sp)
library(maptools)
library(RNetCDF)
library(leaflet)
library(colorRamps)
library(gpclib)
library(rgeos)
library(mapdata)
library(MASS)
library(neuralnet)
library(h2o)
library(reshape2)
library(MASS)
library(gstat)
library(pROC)


#===============================================================================================
#  파일 읽기
#===============================================================================================
# 기상 자료 읽기
# 지점, 일시, 기온(°C), 기온 QC플래그, 강수량(mm), 강수량 QC플래그, 풍속(m/s), 풍속 QC플래그
# ,풍향(16방위), 풍향 QC플래그, 습도(%), 습도 QC플래그, 증기압(hPa), 이슬점온도(°C), 현지기압(hPa), 현지기압 QC플래그
# ,해면기압(hPa), 해면기압 QC플래그, 일조(hr), 일조 QC플래그, 일사(MJ/m2), 적설(cm), 3시간신적설(cm)
# ,전운량(10분위), 중하층운량(10분위), 운형(운형약어), 최저운고(100m ), 시정(10m), 지면상태(지면상태코드)
# ,현상번호(국내식), 지면온도(°C), 지면온도 QC플래그, 5cm 지중온도(°C), 10cm 지중온도(°C)
# ,20cm 지중온도(°C), 30cm 지중온도(°C)

fileInfo = Sys.glob(paste(globalVar$inpConfig, "Big_Data_For_Input_ASOS_2011-2019_QC.inp", sep = "/"))
data = data.table::fread(fileInfo, sep=",", header = FALSE, stringsAsFactors = FALSE)
# data = readr::read_csv(file = fileInfo, locale = locale("ko", encoding = "EUC-KR"))

colnames(data) = c("rowNum","stationNum", "dateTimeOri", "temp", "tempQc", "prec", "precQc", "ws", "wsQc", "wd", "wdQc","rh", "rhQc", "waterRh", "dewTemp","localPres", "localPresQc", "seaPres", "seaPresQc","daylight", "daylightQc", "sr", "snowfall","v3hrSnowfall", "allCloudAmount", "middleCloudAmount", "cloudType", "cloudBottomHeight", "vis", "landType", "weatherType", "surfaceTemp", "surfaceTempQc", "surface5mTemp","surface10mTemp", "surface20mTem", "surface30mTemp", "dateTime", "year", "month", "day", "hour", "minute")

dplyr::tbl_df(data)
```

```
 rowNum  stationNum  dateTimeOri           temp  tempQc   prec  precQc    ws  wsQc   wd ▶
  <int>       <int>  <chr>                <dbl>   <int>  <dbl>   <int> <dbl> <int> <int>
      1          90  2011-01-01 00:00      -4.6       0     NA      NA   4.4     0  290
      2          90  2011-01-01 01:00      -5.0       0     NA      NA   3.8     0  290
      3          90  2011-01-01 02:00      -4.5       0     NA      NA   4.2     0  290
      4          90  2011-01-01 03:00      -2.7       0     NA      NA   2.9     0  290
      5          90  2011-01-01 04:00      -1.7       0     NA      NA   3.0     0  290
      6          90  2011-01-01 05:00      -0.4       0     NA      NA   3.6     0  340
```

rowNum <int>	stationNum <int>	dateTimeOri <chr>	temp <dbl>	tempQc <int>	prec <dbl>	precQc <int>	ws <dbl>	wsQc <int>	wd <int>
7	90	2011-01-01 06:00	-0.5	0	0.0	0	3.3	0	340
8	90	2011-01-01 07:00	-0.1	0	NA	9	4.2	0	320
9	90	2011-01-01 08:00	-0.7	0	NA	9	3.9	0	340
10	90	2011-01-01 09:00	-0.2	0	0.0	0	3.9	0	320

1-10 of 10,000 rows | 1-10 of 43 columns Previous **1** 2 3 4 5 6 ... 1000 Next

```
#***************************************************************************
# 기상 관측소 정보 읽기
#***************************************************************************
fileInfo = Sys.glob(paste(globalVar$inpConfig, "Station_Information_20190714.info", sep = "/"))
stationData = data.table::fread(fileInfo, sep = "\t", header = FALSE)
colnames(stationData) = c("stationNum", "lon", "lat", "hight", "stationName", "metroStationName")

dplyr::tbl_df(stationData)
```

stationNum <int>	lon <dbl>	lat <dbl>	hight <dbl>	stationName <chr>	metroStationName <chr>
90	128.564720	38.2508500	18.0600	Sokcho	Gwanwon
93	127.754700	37.9475000	95.6100	Bukchunceon	Gwanwon
95	127.304200	38.1478800	155.4800	Cheorwon	Gwanwon
98	127.060690	37.9018600	115.6200	Dongducheon	Gyenggi
99	126.766490	37.8858800	30.5900	Munsan	Gyenggi
100	128.718330	37.6771300	772.5700	Daegwallyeong	Gwanwon
101	127.735700	37.9025600	76.4700	Chuncheon	Gwanwon
102	124.630460	37.9661100	36.0000	Baengnyeongdo	Baengnyeongdo
104	128.855350	37.8045600	78.9000	Bukgangneung	Gwanwon
105	128.890980	37.7514700	26.0400	Gangneung	Gwanwon

1-10 of 97 rows Previous **1** 2 3 4 5 6 ... 10 Next

```
#***************************************************************************
# 온열질환자 자료 읽기
#***************************************************************************
fileInfo = Sys.glob(paste(globalVar$inpConfig, "Big_Data_For_Validation_QC_L2_2015-2019.val", sep = "/"))
valData = data.table::fread(fileInfo, sep = " ", header = FALSE)
colnames(valData) = c("year", "month", "day", "metroStationName", "hwanja", "death")

dplyr::tbl_df(valData)
```

year <int>	month <int>	day <int>	metroStationName <chr>	hwanja <int>	death <int>
2015	5	27	Seoul	0	0
2015	5	27	Busan	0	0
2015	5	27	Daegu	0	0

year	month	day	metroStationName	hwanja	death
<int>	<int>	<int>	<chr>	<int>	<int>
2015	5	27	Incheon	1	0
2015	5	27	Gwangju	0	0
2015	5	27	Daejeon	0	0
2015	5	27	Uasan	0	0
2015	5	27	Gyenggi	2	0
2015	5	27	Gwanwon	1	0
2015	5	27	Chungbuk	0	0

1-10 of 6,839 rows Previous **1** 2 3 4 5 6 ... 684 Next

```
#***********************************************************************
# 일평균 수행
#***********************************************************************
dataL1 = data %>%
  dplyr::na_if(-999.0) %>%
  dplyr::select(dateTime, year, month, day, hour, stationNum, temp, ws, rh, sr) %>%
  dplyr::filter(
    dplyr::between(year, 2015, 2019)
    , dplyr::between(month, 5, 9)
    , dplyr::between(hour, 9, 18)
    ) %>%
  dplyr::group_by(year, month, day, stationNum) %>%
  dplyr::summarise(
    maxTemp = max(temp, na.rm = TRUE)
    , meanTemp = mean(temp, na.rm = TRUE)
    , meanRh = mean(rh, na.rm = TRUE)
    , meanWs = mean(ws, na.rm = TRUE)
    , sumSr = sum(sr, na.rm = TRUE)
  )

dplyr::glimpse(dataL1)
```

```
## Rows: 64,739
## Columns: 9
## Groups: year, month, day [686]
## $ year       <int> 2015, 2015, 2015, 2015, 2015, 2015, 2015, 2015, 2015, 20…
## $ month      <int> 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5,…
## $ day        <int> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,…
## $ stationNum <int> 90, 95, 98, 99, 100, 101, 102, 104, 105, 106, 108, 112, …
## $ maxTemp    <dbl> 18.4, 27.6, 28.3, 26.6, 24.3, 29.9, 20.4, 25.5, 27.9, 23…
## $ meanTemp   <dbl> 16.10, 25.41, 25.62, 24.67, 22.65, 26.35, 18.10, 22.63, …
## $ meanRh     <dbl> 75.0, 39.3, 42.9, 44.3, 43.3, 43.8, 62.1, 50.0, 43.2, 64…
## $ meanWs     <dbl> 1.87, 2.57, 2.76, 2.84, 2.84, 1.83, 3.79, 1.86, 2.08, 1.…
## $ sumSr      <dbl> 0.00, 0.00, 0.00, 0.00, 19.58, 21.05, 0.00, 22.56, 24.44…
```

```
#***********************************************************************
# 4종 폭염지수 계산
#***********************************************************************
dataL2 = dataL1 %>%
  dplyr::filter(sumSr > 0) %>%
  dplyr::mutate(
    heatIndex = fnHeatIndex(meanTemp, meanRh)
    , humidIndex = fnHumidIndex(meanTemp, meanRh)
    , appTempIndex = fnAppTempIndex(meanTemp, meanRh, meanWs)
    , appTempRadIndex = fnAppTempRadIndex(meanTemp, meanRh, meanWs, sumSr)
    , wetBulbGolbalTempIndex = fnWetBulbGolbalTempIndex(meanTemp, meanRh, meanWs, sumSr)
  )

dplyr::glimpse(dataL1)
```

```
## Rows: 64,739
## Columns: 9
## Groups: year, month, day [686]
## $ year       <int> 2015, 2015, 2015, 2015, 2015, 2015, 2015, 2015, 2015, 20…
## $ month      <int> 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5,…
## $ day        <int> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,…
## $ stationNum <int> 90, 95, 98, 99, 100, 101, 102, 104, 105, 106, 108, 112, …
## $ maxTemp    <dbl> 18.4, 27.6, 28.3, 26.6, 24.3, 29.9, 20.4, 25.5, 27.9, 23…
## $ meanTemp   <dbl> 16.10, 25.41, 25.62, 24.67, 22.65, 26.35, 18.10, 22.63, …
## $ meanRh     <dbl> 75.0, 39.3, 42.9, 44.3, 43.3, 43.8, 62.1, 50.0, 43.2, 64…
## $ meanWs     <dbl> 1.87, 2.57, 2.76, 2.84, 2.84, 1.83, 3.79, 1.86, 2.08, 1.…
## $ sumSr      <dbl> 0.00, 0.00, 0.00, 0.00, 19.58, 21.05, 0.00, 22.56, 24.44…
```

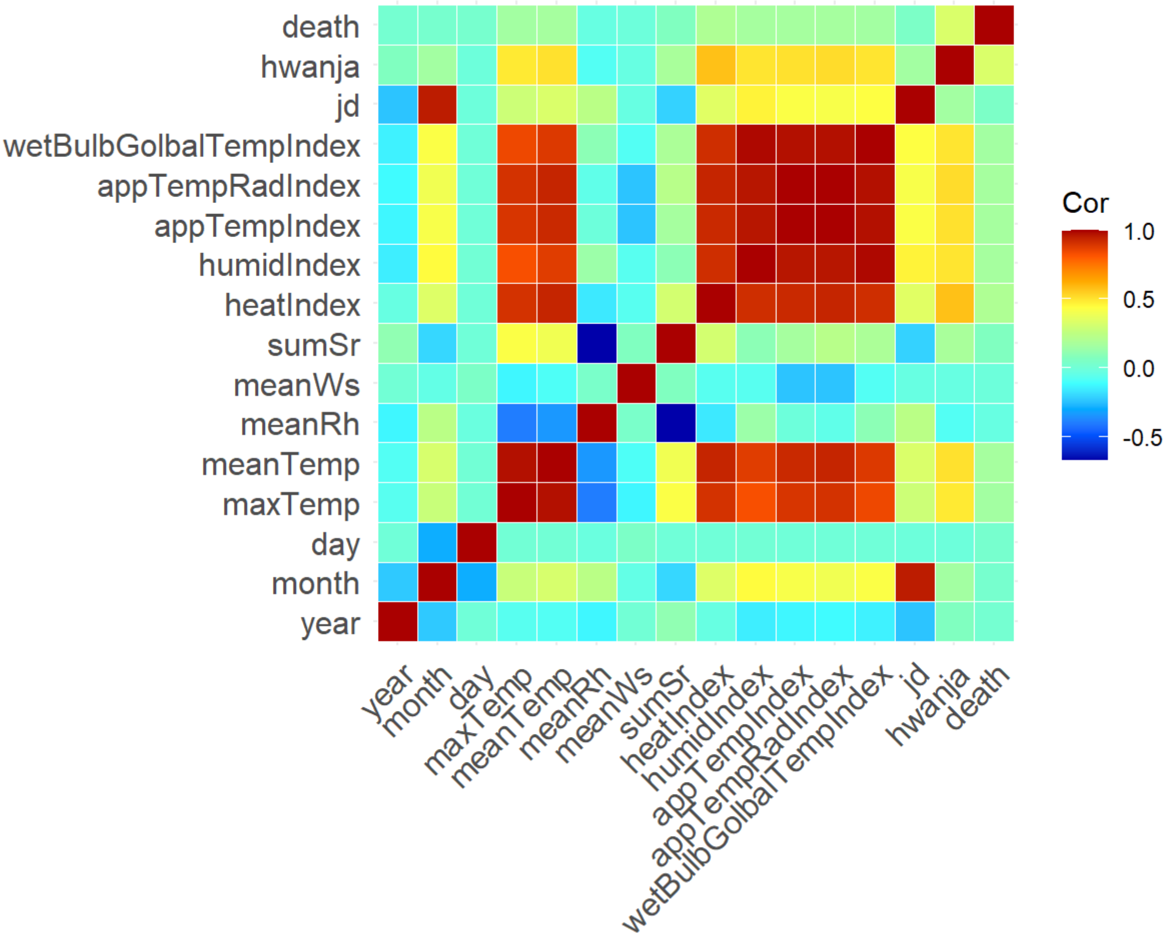
```
#***********************************************************************
# 자료 병합 (기상 자료, 기상 관측소, 온열질환자 자료)
#***********************************************************************
dataL3 = dataL2 %>%
  dplyr::left_join(stationData, by=c("stationNum" = "stationNum")) %>%
  dplyr::left_join(valData, by=c("metroStationName" = "metroStationName", "year" = "year", "month" = "month", "day" = "day")) %>%
  dplyr::mutate(
    sDate = paste(year, month, day, sep = "-")
    , dtDate = readr::parse_date(sDate, "%Y-%m-%d")
    , jd = lubridate::yday(dtDate)
    , xran = lubridate::decimal_date(dtDate)
  ) %>%
  readr::type_convert()

dplyr::glimpse(dataL3)
```

```
## Rows: 27,775
## Columns: 25
## Groups: year, month, day [686]
## $ year                   <int> 2015, 2015, 2015, 2015, 2015, 2015, 2015, 20…
## $ month                  <int> 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5,…
## $ day                    <int> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,…
## $ stationNum             <int> 100, 101, 104, 105, 108, 112, 114, 119, 129,…
## $ maxTemp                <dbl> 24.3, 29.9, 25.5, 27.9, 27.0, 23.0, 29.3, 26…
## $ meanTemp               <dbl> 22.65, 26.35, 22.63, 24.55, 24.91, 21.10, 26…
## $ meanRh                 <dbl> 43.3, 43.8, 50.0, 43.2, 46.2, 74.3, 45.4, 52…
## $ meanWs                 <dbl> 2.84, 1.83, 1.86, 2.08, 3.64, 2.72, 1.93, 2.…
## $ sumSr                  <dbl> 19.58, 21.05, 22.56, 24.44, 20.68, 20.05, 22…
## $ heatIndex              <dbl> 22.10116667, 26.18422222, 22.25411111, 24.18…
## $ humidIndex             <dbl> 23.69232350, 29.11854157, 24.68393380, 26.37…
## $ appTempIndex           <dbl> 28.58120208, 34.01367816, 29.84815869, 31.47…
## $ appTempRadIndex        <dbl> 25.52856080, 31.03216136, 26.93970173, 28.65…
## $ wetBulbGolbalTempIndex <dbl> 18.39708510, 21.79251020, 19.31535180, 20.36…
## $ lon                    <dbl> 128.71833, 127.73570, 128.85535, 128.89098, …
## $ lat                    <dbl> 37.67713, 37.90256, 37.80456, 37.75147, 37.5…
## $ hight                  <dbl> 772.57, 76.47, 78.90, 26.04, 85.67, 68.99, 1…
## $ stationName            <chr> "Daegwallyeong", "Chuncheon", "Bukgangneung"…
## $ metroStationName       <chr> "Gwanwon", "Gwanwon", "Gwanwon", "Gwanwon", …
## $ hwanja                 <int> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, …
## $ death                  <int> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, …
## $ sDate                  <chr> "2015-5-1", "2015-5-1", "2015-5-1", "2015-5-…
## $ dtDate                 <date> 2015-05-01, 2015-05-01, 2015-05-01, 2015-05…
## $ jd                     <dbl> 121, 121, 121, 121, 121, 121, 121, 121, 121,…
## $ xran                   <dbl> 2015.328767, 2015.328767, 2015.328767, 2015.…
```

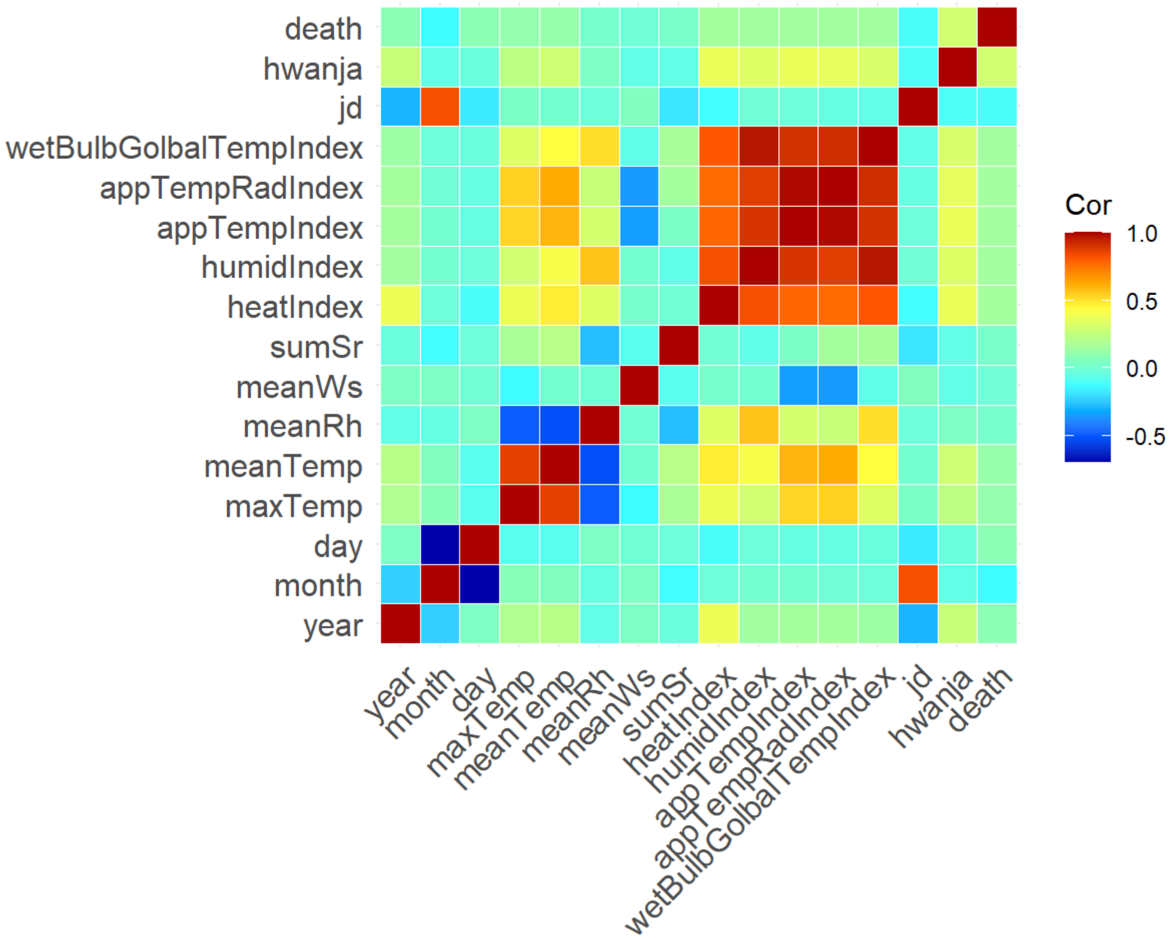
```
#***************************************************************************
# 유의미한 변수 찾기
#***************************************************************************
dataL4 = dataL3 %>%
  na.omit() %>%
  dplyr::select(maxTemp, meanTemp, meanRh, meanWs, sumSr, heatIndex, humidIndex, appTempIndex, appTempRadIndex, wetBulbGolbalTempIndex, jd, hwanja, death)

dplyr::glimpse(dataL4)
```

```
## Rows: 15,958
## Columns: 16
## Groups: year, month, day [404]
## $ year                   <int> 2015, 2015, 2015, 2015, 2015, 2015, 2015, 20…
## $ month                  <int> 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5,…
## $ day                    <int> 27, 27, 27, 27, 27, 27, 27, 27, 27, 27, 27, …
## $ maxTemp                <dbl> 28.1, 31.9, 21.1, 21.6, 31.2, 27.4, 31.7, 31…
## $ meanTemp               <dbl> 24.80, 27.77, 18.29, 18.67, 28.48, 24.28, 28…
## $ meanRh                 <dbl> 25.9, 24.2, 64.4, 62.1, 19.7, 48.1, 21.8, 25…
## $ meanWs                 <dbl> 3.11, 1.73, 1.77, 1.90, 2.76, 2.57, 1.85, 2.…
## $ sumSr                  <dbl> 27.01, 24.74, 22.74, 24.62, 23.40, 22.65, 26…
## $ heatIndex              <dbl> 24.01183333, 27.23444444, 17.85611111, 18.21…
## $ humidIndex             <dbl> 23.73520353, 27.21263827, 20.23627969, 20.52…
## $ appTempIndex           <dbl> 29.29044764, 33.52789402, 25.50720576, 25.74…
## $ appTempRadIndex        <dbl> 26.59729948, 30.72491418, 22.60745792, 22.93…
## $ wetBulbGolbalTempIndex <dbl> 18.76681249, 20.92341958, 16.81547301, 17.05…
## $ jd                     <dbl> 147, 147, 147, 147, 147, 147, 147, 147, 147,…
## $ hwanja                 <int> 1, 1, 1, 1, 0, 1, 1, 2, 0, 0, 0, 0, 0, 0, 0,…
## $ death                  <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,…
```

```
# 상관계수 행렬
corMat = cor(dataL4)

ggcorrplot(corMat, outline.col = "white", lab = FALSE) +
  scale_fill_gradientn(colours = colorRamps::matlab.like(10)) +
  labs(fill = "Cor")
```



```
# 폭염 기준 상관계수 행렬
dataL5 = dataL4 %>%
  dplyr::filter(
    maxTemp >= 33
    , heatIndex >= 32
    )

corMat = cor(dataL5)

ggcorrplot(corMat, outline.col = "white", lab = FALSE) +
  scale_fill_gradientn(colours = colorRamps::matlab.like(10)) +
  labs(fill = "Cor")
```



```
# 각 지점에 따른 다중선형회귀모형/딥러닝 학습
stationNameList = sort(unique(dataL3$metroStationName))
# stationNameInfo = "Busan"

h2o::h2o.init()
```

```
##
## H2O is not running yet, starting it now...
##
## Note:  In case of errors look at the following log files:
##     C:\Users\Public\Documents\ESTsoft\CreatorTemp\RtmpkTUEKb\file33587567633f/h2o_indisystem_started_from_r.out
##     C:\Users\Public\Documents\ESTsoft\CreatorTemp\RtmpkTUEKb\file33581d7e7e25/h2o_indisystem_started_from_r.err
##
##
## Starting H2O JVM and connecting:  Connection successful!
##
## R is connected to the H2O cluster:
##     H2O cluster uptime:         3 seconds 548 milliseconds
##     H2O cluster timezone:       Asia/Seoul
##     H2O data parsing timezone:  UTC
##     H2O cluster version:        3.32.0.1
##     H2O cluster version age:    2 months and 27 days
##     H2O cluster name:           H2O_started_from_R_indisystem_akr880
##     H2O cluster total nodes:    1
##     H2O cluster total memory:   5.95 GB
##     H2O cluster total cores:    8
##     H2O cluster allowed cores:  8
##     H2O cluster healthy:        TRUE
##     H2O Connection ip:          localhost
##     H2O Connection port:        54321
##     H2O Connection proxy:       NA
##     H2O Internal Security:      FALSE
##     H2O API Extensions:         Amazon S3, Algos, AutoML, Core V3, TargetEncoder, Core V4
##     R Version:                  R version 4.0.3 (2020-10-10)
```

```
dataL6 = data.frame()

for (stationNameInfo in stationNameList) {

  dataL4 = dataL3  %>%
    dplyr::filter(
      hwanja >= 0
      , metroStationName == stationNameInfo
      )

  if (nrow(dataL4) < 1) { next }

  dataL5 = na.omit(dataL4)

  #***************************************
  # 데이터 분할
  #***************************************
  # 훈련 및 데이터 셋을 60:40으로 나누기 위한 인덱스 설정
  ind = sample(1:nrow(dataL5), nrow(dataL5) * 0.6)

  # 해당 인덱스에 따라 자료 할당
  trainData = dataL5[ind,]
  testData = dataL5[-ind,]

  # 표준화 수행
  # trainData = dataL5[-ind,]  %>%
    # dplyr::mutate_each_(funs(scale), vars=c("meanTemp", "meanRh", "sumSr", "meanWs", "wetBulbGolbalTempIndex", "jd"))

  # testData = dataL5[ind,]  %>%
    # dplyr::mutate_each_(funs(scale), vars=c("meanTemp", "meanRh", "sumSr", "meanWs", "wetBulbGolbalTempIndex", "jd"))

  # 정규화 수행
  # trainData = dataL5[-ind,]  %>%
  #   dplyr::mutate_each_(funs(scales::rescale), vars=c("meanTemp", "meanRh", "sumSr", "meanWs", "wetBulbGolbalTempIndex", "jd"))

  # testData = dataL5[ind,]  %>%
  #   dplyr::mutate_each_(funs(scales::rescale), vars=c("meanTemp", "meanRh", "sumSr", "meanWs", "wetBulbGolbalTempIndex", "jd"))

  # 훈련 데이터셋 확인
  dplyr::tbl_df(trainData)

  # 테스트 데이터셋 확인
  dplyr::tbl_df(testData)

  # 동적 회귀식 생성
  # allVar = colnames(trainData)
  # predictorVarList = allVar[!allVar %in% "Churn"]
  # predictorVar = paste(predictorVarList, collapse = "+")
  # form = as.formula(paste("Churn ~", predictorVar, collapse = "+"))

  # 수동 회귀식 생성
  # lmForm = hwanja ~ meanTemp + meanRh + heatIndex
  # lmForm = hwanja ~ meanTemp + meanRh + sumSr + meanWs
  form = hwanja ~ meanTemp + meanRh + sumSr + meanWs + wetBulbGolbalTempIndex + jd
  # lmForm = hwanja ~ wetBulbGolbalTempIndex + meanSurfaceTemp + meanAllCloudAmount

  #***************************************
  # 다중선형회귀모형
  #***************************************
  lmFit = lm(form, data = trainData)
  summary(lmFit)
  xAxis = predict(lmFit, new = testData)
  yAxis = testData$hwanja

  # plot(xAxis, yAxis)

  cat(sprintf(
    "MLR [%10s] length : %05s | cor : %05s | Bias : %05s | RMSE : %05s"
    , stationNameInfo
    , length(xAxis)
    , round(cor(xAxis, yAxis), 2)
    , round(Metrics::rmse(xAxis, yAxis), 2)
    , round(Metrics::bias(xAxis, yAxis), 2)
    ), "\n")

  #****************************************
  # 딥러닝
  #****************************************
  # activation : 활성화 함수로서 Rectifier 정규화 선형 함수 (즉 Keras의 ReLU 동일)
  # hidden : 숨겨진 레이어의 수와 뉴런 수 (일반적으로 입력 차원의 1/10 or 1/100 단위)
  # epochs : 반복 횟수 (기본 10-40)
  # nfolds : 훈련 반복 수

  layerNum = as.integer(nrow(trainData) / 10)
  # layerNum = as.integer(nrow(trainData) / 100)

  dlModel = h2o::h2o.deeplearning(
    x = c("meanTemp", "meanRh", "sumSr", "meanWs", "wetBulbGolbalTempIndex", "jd")
    , y = c("hwanja")
    , training_frame = as.h2o(trainData)
    , activation = 'Rectifier'
    , hidden = rep(layerNum, 3)
    , nfolds = 10
    , epochs = 100
  )

  xAxis = as.data.frame(h2o::h2o.predict(object=dlModel, newdata=as.h2o(testData)))$predict
  yAxis = testData$hwanja

  # plot(xAxis, yAxis)

  cat(sprintf(
    "DL [%10s] length : %05s | cor : %05s | Bias : %05s | RMSE : %05s"
    , stationNameInfo
    , length(xAxis)
    , round(cor(xAxis, yAxis), 2)
    , round(Metrics::rmse(xAxis, yAxis), 2)
    , round(Metrics::bias(xAxis, yAxis), 2)
  ), "\n")



  dataL5$stationNameInfo = stationNameInfo
  dataL5$mlr = predict(lmFit, new = dataL5)
  dataL5$dl = as.data.frame(h2o::h2o.predict(object=dlModel, newdata=as.h2o(dataL5)))$predict

  dataL6 = dplyr::bind_rows(dataL6, dataL5)
}
```

```
## MLR [     Busan] length :   139 | cor :  0.59 | Bias :  1.73 | RMSE : -0.14
## 
  |                                                                            
  |                                                                      |   0%
  |                                                                            
  |======================================================================| 100%
## 
  |                                                                            
  |                                                                      |   0%
  |                                                                            
  |======================================================                |  77%
  |                                                                            
  |====================================================================  |  97%
  |                                                                            
  |======================================================================| 100%
## 
  |                                                                            
  |                                                                      |   0%
  |                                                                            
  |======================================================================| 100%
## 
  |                                                                            
  |                                                                      |   0%
  |                                                                            
  |======================================================================| 100%
## DL [     Busan] length :   139 | cor :  0.68 | Bias :  1.63 | RMSE : -0.38
## 
  |                                                                            
  |                                                                      |   0%
  |                                                                            
  |======================================================================| 100%
## 
  |                                                                            
  |                                                                      |   0%
  |                                                                            
  |======================================================================| 100%
## MLR [  Chungbuk] length :   322 | cor :   0.6 | Bias :     2 | RMSE : -0.24
## 
  |                                                                            
  |                                                                      |   0%
  |                                                                            
  |======================================================================| 100%
## 
  |                                                                            
  |                                                                      |   0%
  |                                                                            
  |==========================================================            |  83%
  |                                                                            
  |======================================================================| 100%
## 
  |                                                                            
  |                                                                      |   0%
  |                                                                            
  |======================================================================| 100%
## 
  |                                                                            
  |                                                                      |   0%
  |                                                                            
  |======================================================================| 100%
## DL [  Chungbuk] length :   322 | cor :   0.7 | Bias :  1.77 | RMSE : -0.23
## 
```

```
  |                                                                            
  |                                                                      |   0%
  |                                                                            
  |======================================================================| 100%
## 
  |                                                                            
  |                                                                      |   0%
  |                                                                            
  |======================================================================| 100%
## MLR [  Chungnam] length :   294 | cor :  0.63 | Bias :  2.15 | RMSE :  0.15
## 
  |                                                                            
  |                                                                      |   0%
  |                                                                            
  |======================================================================| 100%
## 
  |                                                                            
  |                                                                      |   0%
  |                                                                            
  |========================================================              |  79%
  |                                                                            
  |======================================================================| 100%
## 
  |                                                                            
  |                                                                      |   0%
  |                                                                            
  |======================================================================| 100%
## 
  |                                                                            
  |                                                                      |   0%
  |                                                                            
  |======================================================================| 100%
## DL [  Chungnam] length :   294 | cor :  0.71 | Bias :  1.93 | RMSE :   0.1
## 
  |                                                                            
  |                                                                      |   0%
  |                                                                            
  |======================================================================| 100%
## 
  |                                                                            
  |                                                                      |   0%
  |                                                                            
  |======================================================================| 100%
## MLR [     Daegu] length :   161 | cor :  0.49 | Bias :  1.22 | RMSE :  0.02
## 
  |                                                                            
  |                                                                      |   0%
  |                                                                            
  |======================================================================| 100%
## 
  |                                                                            
  |                                                                      |   0%
  |                                                                            
  |=======================================================               |  78%
  |                                                                            
  |======================================================================| 100%
## 
  |                                                                            
  |                                                                      |   0%
  |                                                                            
  |======================================================================| 100%
## 
```

```
  |
  |                                                                      |   0%
  |
  |======================================================================| 100%
## DL [     Daegu] length :   161 | cor :  0.62 | Bias :   1.1 | RMSE :  0.12
## 
  |
  |                                                                      |   0%
  |
  |======================================================================| 100%
## 
  |
  |                                                                      |   0%
  |
  |======================================================================| 100%
## MLR [   Daejeon] length :   161 | cor :  0.56 | Bias :  0.71 | RMSE :  0.12
## 
  |
  |                                                                      |   0%
  |
  |======================================================================| 100%
## 
  |
  |                                                                      |   0%
  |
  |======================================================================| 100%
## 
  |
  |                                                                      |   0%
  |
  |======================================================================| 100%
## 
  |
  |                                                                      |   0%
  |
  |======================================================================| 100%
## DL [   Daejeon] length :   161 | cor :  0.49 | Bias :  0.85 | RMSE : -0.06
## 
  |
  |                                                                      |   0%
  |
  |======================================================================| 100%
## 
  |
  |                                                                      |   0%
  |
  |======================================================================| 100%
## MLR [   Gwangju] length :   161 | cor :  0.57 | Bias :  1.19 | RMSE :  0.03
## 
  |
  |                                                                      |   0%
  |
  |======================================================================| 100%
## 
  |
  |                                                                      |   0%
  |
  |=========================================================             |  81%
  |
  |======================================================================| 100%
## 
  |
  |                                                                      |   0%
  |
  |======================================================================| 100%
##
  |
  |                                                                      |   0%
  |
  |======================================================================| 100%
## DL [   Gwangju] length :   161 | cor :  0.69 | Bias :  1.07 | RMSE :  0.18
##
  |
  |                                                                      |   0%
  |
  |======================================================================| 100%
##
  |
  |                                                                      |   0%
  |
  |======================================================================| 100%
## MLR [   Gwanwon] length :   912 | cor :   0.5 | Bias :  2.03 | RMSE : -0.07
##
  |
  |                                                                      |   0%
  |
  |======================================================================| 100%
##
  |
  |                                                                      |   0%
  |
  |===================                                                   |  27%
  |
  |======================                                                |  32%
  |
  |====================================================                  |  75%
  |
  |=========================================================             |  82%
  |
  |===========================================================           |  84%
  |
  |==============================================================        |  88%
  |
  |===================================================================   |  96%
  |
  |======================================================================| 100%
##
  |
  |                                                                      |   0%
  |
  |======================================================================| 100%
##
  |
  |                                                                      |   0%
  |
  |======================================================================| 100%
## DL [   Gwanwon] length :   912 | cor :  0.58 | Bias :   1.9 | RMSE :  -0.1
##
  |
  |                                                                      |   0%
  |
  |======================================================================| 100%
##
  |
  |                                                                      |   0%
  |
  |======================================================================| 100%
## MLR [  Gyengbuk] length :   638 | cor :  0.69 | Bias :  2.28 | RMSE : -0.08
##
  |
  |                                                                      |   0%
  |
  |======================================================================| 100%
##
  |
  |                                                                      |   0%
  |
  |===                                                                   |   5%
  |
  |========================================================              |  80%
  |
  |=========================================================             |  82%
  |
  |============================================================          |  86%
  |
  |===================================================================== |  99%
  |
  |======================================================================| 100%
##
  |
  |                                                                      |   0%
  |
  |======================================================================| 100%
##
  |
  |                                                                      |   0%
  |
  |======================================================================| 100%
## DL [  Gyengbuk] length :   638 | cor :  0.74 | Bias :  2.13 | RMSE : -0.27
##
  |
  |                                                                      |   0%
  |
  |======================================================================| 100%
##
  |
  |                                                                      |   0%
  |
  |======================================================================| 100%
## MLR [   Gyenggi] length :   161 | cor :  0.67 | Bias :  8.39 | RMSE : -1.07
##
  |
  |                                                                      |   0%
  |
  |======================================================================| 100%
##
  |
  |                                                                      |   0%
  |
  |=======================================================               |  79%
  |
  |======================================================================| 100%
##
  |
  |                                                                      |   0%
  |
  |======================================================================| 100%
##
  |
  |                                                                      |   0%
  |
  |======================================================================| 100%
## DL [   Gyenggi] length :   161 | cor :  0.85 | Bias :  5.85 | RMSE : -0.23
##
  |
  |                                                                      |   0%
  |
  |======================================================================| 100%
##
  |
  |                                                                      |   0%
  |
  |======================================================================| 100%
## MLR [  Gyengnam] length :  1097 | cor :  0.67 | Bias :   3.1 | RMSE : -0.06
##
  |
  |                                                                      |   0%
  |
  |======================================================================| 100%
##
  |
  |                                                                      |   0%
  |
  |=                                                                     |   1%
  |
  |==                                                                    |   3%
  |
  |=====                                                                 |   7%
  |
  |=======                                                               |  10%
  |
  |=========                                                             |  12%
  |
  |==========                                                            |  15%
  |
  |==============                                                        |  20%
  |
  |=================                                                     |  24%
  |
  |===================                                                   |  28%
  |
  |====================                                                  |  29%
  |
  |=====================                                                 |  30%
  |
  |=======================                                               |  33%
  |
  |==========================                                            |  37%
  |
  |===========================                                           |  38%
  |
  |=============================                                         |  41%
  |
  |================================                                      |  46%
  |
  |======================================                                |  54%
  |
  |========================================                              |  57%
  |
  |==========================================                            |  60%
  |
  |================================================                      |  69%
  |
  |=====================================================                 |  76%
  |
  |=======================================================               |  79%
  |
  |=========================================================             |  82%
  |
  |============================================================          |  86%
  |
  |=============================================================         |  87%
  |
  |===============================================================       |  90%
  |
  |================================================================      |  91%
  |
  |===================================================================   |  96%
  |
  |===================================================================== |  98%
  |
  |===================================================================== |  99%
  |
  |======================================================================| 100%
##
  |
  |                                                                      |   0%
  |
  |======================================================================| 100%
##
  |
  |                                                                      |   0%
  |
  |======================================================================| 100%
## DL [  Gyengnam] length :  1097 | cor :  0.83 | Bias :  2.32 | RMSE :  0.07
##
  |
  |                                                                      |   0%
  |
  |======================================================================| 100%
##
  |
  |                                                                      |   0%
  |
  |======================================================================| 100%
## MLR [   Incheon] length :   161 | cor :  0.65 | Bias :  1.89 | RMSE :  0.35
##
  |
  |                                                                      |   0%
  |
  |======================================================================| 100%
##
  |
  |                                                                      |   0%
  |
  |==========                                                            |  14%
  |
  |=============================================================         |  87%
  |
  |======================================================================| 100%
```

```
##
  |
  |                                                                      |   0%
  |
  |======================================================================| 100%
##
  |
  |                                                                      |   0%
  |
  |======================================================================| 100%
## DL [   Incheon] length :   161 | cor :  0.76 | Bias :  2.41 | RMSE :  0.58
##
  |
  |                                                                      |   0%
  |
  |======================================================================| 100%
##
  |
  |                                                                      |   0%
  |
  |======================================================================| 100%
## MLR [      Jeju] length :   323 | cor :  0.56 | Bias :  1.13 | RMSE :  0.17
##
  |
  |                                                                      |   0%
  |
  |======================================================================| 100%
##
  |
  |                                                                      |   0%
  |
  |=======================================================               |  79%
  |
  |======================================================================| 100%
##
  |
  |                                                                      |   0%
  |
  |======================================================================| 100%
##
  |
  |                                                                      |   0%
  |
  |======================================================================| 100%
## DL [      Jeju] length :   323 | cor :  0.45 | Bias :   1.4 | RMSE :  0.15
##
  |
  |                                                                      |   0%
  |
  |======================================================================| 100%
##
  |
  |                                                                      |   0%
  |
  |======================================================================| 100%
## MLR [   Jeonbuk] length :   801 | cor :   0.6 | Bias :  2.06 | RMSE : -0.02
##
  |
  |                                                                      |   0%
  |
  |======================================================================| 100%
##
```

```
  |
  |                                                                      |   0%
  |
  |=                                                                     |   1%
  |
  |=                                                                     |   2%
  |
  |=====                                                                 |   7%
  |
  |=======                                                               |  11%
  |
  |==========                                                            |  15%
  |
  |=============                                                         |  19%
  |
  |=============================================                         |  65%
  |
  |=========================================================             |  81%
  |
  |============================================================          |  86%
  |
  |==============================================================        |  89%
  |
  |==================================================================    |  95%
  |
  |======================================================================| 100%
##
  |
  |                                                                      |   0%
  |
  |======================================================================| 100%
##
  |
  |                                                                      |   0%
  |
  |======================================================================| 100%
## DL [   Jeonbuk] length :   801 | cor :  0.63 | Bias :  2.04 | RMSE : -0.07
##
  |
  |                                                                      |   0%
  |
  |======================================================================| 100%
##
  |
  |                                                                      |   0%
  |
  |======================================================================| 100%
## MLR [   Jeonnam] length :   897 | cor :  0.66 | Bias :  2.71 | RMSE : -0.17
##
  |
  |                                                                      |   0%
  |
  |======================================================================| 100%
##
  |
  |                                                                      |   0%
  |
  |=                                                                     |   1%
  |
  |==                                                                    |   3%
  |
  |====                                                                  |   6%
  |
  |=====                                                                 |   7%
  |
  |=======                                                               |  10%
  |
  |==========                                                            |  15%
  |
  |============                                                          |  17%
  |
  |==============                                                        |  20%
  |
  |===============                                                       |  21%
  |
  |=================                                                     |  25%
  |
  |==================                                                    |  26%
  |
  |=====================                                                 |  30%
  |
  |=========================                                             |  35%
  |
  |=============================                                         |  42%
  |
  |===============================                                       |  44%
  |
  |==================================                                    |  48%
  |
  |=====================================================                 |  76%
  |
  |========================================================              |  80%
  |
  |==========================================================            |  83%
  |
  |=============================================================         |  87%
  |
  |===================================================================== |  99%
  |
  |======================================================================| 100%
##
  |
  |                                                                      |   0%
  |
  |======================================================================| 100%
##
  |
  |                                                                      |   0%
  |
  |======================================================================| 100%
## DL [   Jeonnam] length :  897 | cor :  0.72 | Bias :  2.65 | RMSE : -0.04
##
  |
  |                                                                      |   0%
  |
  |======================================================================| 100%
##
  |
  |                                                                      |   0%
  |
  |======================================================================| 100%
## MLR [     Seoul] length :  161 | cor :  0.61 | Bias :   4.3 | RMSE :  0.54
##
  |
  |                                                                      |   0%
  |
  |======================================================================| 100%
##
  |
  |                                                                      |   0%
  |
  |===================================================================   |  95%
  |
  |======================================================================| 100%
##
  |
  |                                                                      |   0%
  |
  |======================================================================| 100%
##
  |
  |                                                                      |   0%
  |
  |======================================================================| 100%
## DL [     Seoul] length :   161 | cor :  0.84 | Bias :  3.62 | RMSE :     0
##
  |
  |                                                                      |   0%
  |
  |======================================================================| 100%
##
  |
  |                                                                      |   0%
  |
  |======================================================================| 100%
```

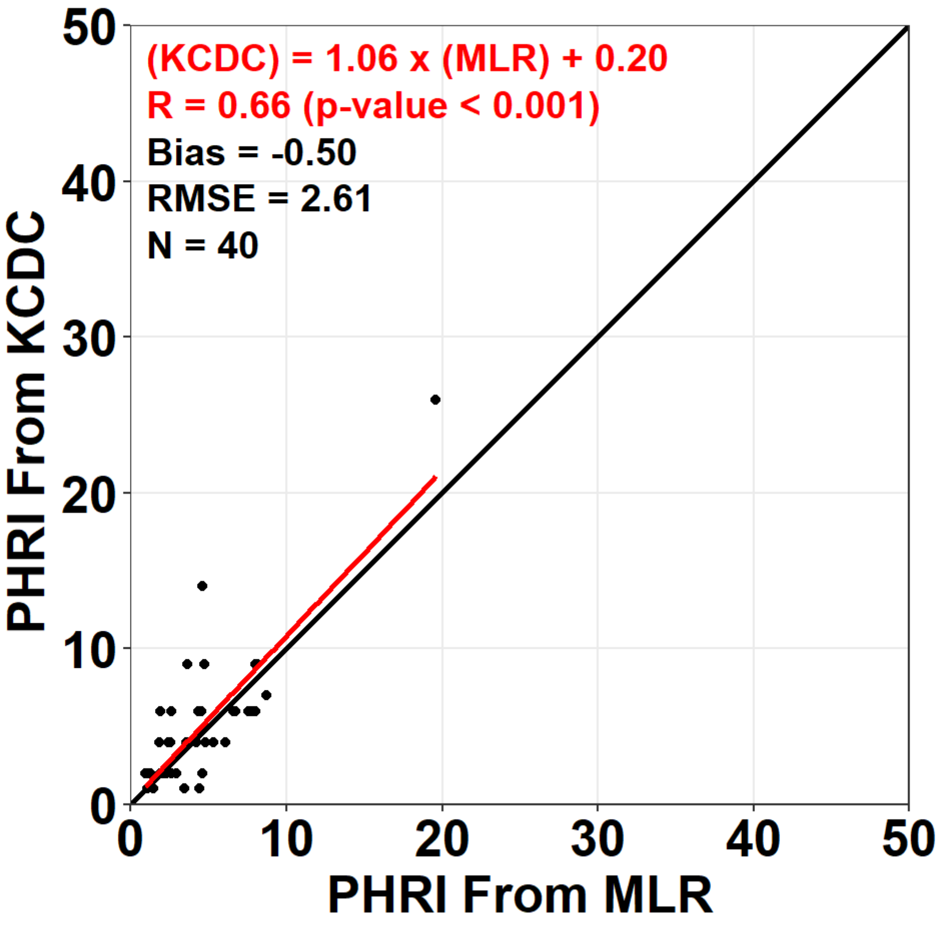
```
saveFile = sprintf("%s/%s_%s", globalVar$outConfig, serviceName, "Big_Data_For_Output_ASOS_2015-2019.out")
readr::write_csv(dataL6, file = saveFile)

dataL7 = dataL6

#=====================================================================================
# 시각화를 위한 파일 읽기
#=====================================================================================
fileInfo = Sys.glob(paste(globalVar$outConfig, "LSH0079_Big_Data_For_Output_ASOS_2015-2019.out", sep = "/"))
dataL7 = readr::read_csv(file = fileInfo, locale = locale("ko", encoding = "EUC-KR")) %>%
  na.omit() %>%
  dplyr::rename(
    "kcdc" = "hwanja"
  ) %>%
  dplyr::filter(
    mlr >= 0
    , dl >= 0
  )

dplyr::glimpse(dataL7)
```

```
## Rows: 11,131
## Columns: 28
## $ year                   <dbl> 2015, 2015, 2015, 2015, 2015, 2015, 2015, 20…
## $ month                  <dbl> 5, 5, 5, 6, 6, 6, 6, 6, 6, 6, 7, 7, 7, 7, 7,…
## $ day                    <dbl> 27, 28, 31, 3, 4, 8, 10, 24, 25, 29, 15, 19,…
## $ stationNum             <dbl> 159, 159, 159, 159, 159, 159, 159, 159, 159,…
## $ maxTemp                <dbl> 26.0, 26.3, 26.3, 27.3, 28.1, 22.8, 27.7, 25…
## $ meanTemp               <dbl> 24.09, 24.16, 24.41, 23.90, 24.86, 21.63, 25…
## $ meanRh                 <dbl> 59.6, 54.3, 56.5, 48.3, 45.7, 85.8, 65.0, 67…
## $ meanWs                 <dbl> 2.74, 2.44, 3.22, 3.53, 3.26, 1.30, 1.78, 2.…
## $ sumSr                  <dbl> 23.55, 19.53, 24.85, 24.35, 24.09, 6.92, 19.…
## $ heatIndex              <dbl> 24.11077778, 24.04938889, 24.38183333, 23.60…
## $ humidIndex             <dbl> 28.43859458, 27.66580712, 28.42494004, 26.28…
## $ appTempIndex           <dbl> 32.05532351, 31.83465770, 31.84111289, 30.14…
## $ appTempRadIndex        <dbl> 29.19476671, 28.77959722, 29.04345529, 27.32…
## $ wetBulbGolbalTempIndex <dbl> 21.64504125, 20.90493725, 21.64780426, 20.26…
## $ lon                    <dbl> 129.03203, 129.03203, 129.03203, 129.03203, …
## $ lat                    <dbl> 35.10468, 35.10468, 35.10468, 35.10468, 35.1…
## $ hight                  <dbl> 69.56, 69.56, 69.56, 69.56, 69.56, 69.56, 69…
## $ stationName            <chr> "Busan", "Busan", "Busan", "Busan", "Busan",…
## $ metroStationName       <chr> "Busan", "Busan", "Busan", "Busan", "Busan",…
## $ kcdc                   <dbl> 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 3,…
## $ death                  <dbl> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,…
## $ sDate                  <chr> "2015-5-27", "2015-5-28", "2015-5-31", "2015…
## $ dtDate                 <date> 2015-05-27, 2015-05-28, 2015-05-31, 2015-06…
## $ jd                     <dbl> 147, 148, 151, 154, 155, 159, 161, 175, 176,…
## $ xran                   <dbl> 2015.400000, 2015.402740, 2015.410959, 2015.…
## $ stationNameInfo        <chr> "Busan", "Busan", "Busan", "Busan", "Busan",…
## $ mlr                    <dbl> 0.5019428106, 0.6669360887, 0.4803191188, 0.…
## $ dl                     <dbl> 0.27353698593, 0.58488525992, 0.14657478766,…
```

```
#*******************************************************
# 일별 사례 : 2018년 07월 23일
#*******************************************************
dayData = dataL7 %>%
  dplyr::filter(
    year == 2018
    , month == 8
    , day == 7
  )


#*******************************************************
# 월별 사례 : 2018년 08월
#*******************************************************
monthData = dataL7 %>%
  dplyr::group_by(stationName, lon, lat, year, month) %>%
  dplyr::summarise(
    cor = cor(mlr, kcdc)
    , bias = bias(mlr, kcdc)
    , rmse = rmse(mlr, kcdc)
    , n = n()
    , meanMlr = mean(mlr, na.rm = TRUE)
    , meanDl = mean(dl, na.rm = TRUE)
    , meanKcdc = mean(kcdc, na.rm = TRUE)
  ) %>%
  dplyr::filter(
    year == 2018
    , month == 8
    )


#*******************************************************
# 연별 사례 : 2018-2019년
#*******************************************************
yearData = dataL7 %>%
  dplyr::filter(dplyr::between(year, 2018, 2019))


#*******************************************************
# 산포도 시각화
#*******************************************************
# 일별 (다중선형회귀 예측 vs 환자 관측값)
X = dayData$mlr
Y = dayData$kcdc

# 일별 (딥러닝 예측 vs 환자 관측값)
X = dayData$dl
Y = dayData$kcdc

# 연별 (다중선형회귀 예측 vs 환자 관측값)
# X = monthData$meanMlr
# Y = monthData$meanKcdc

# 연별 (딥러닝 예측 vs 환자 관측값)
# X = monthData$meanDl
# Y = monthData$meanKcdc

val = fnStats(X, Y)
sprintf("%.3f", val)
```

```
## [1] "1.064"  "0.201"  "4.629"  "5.125"  "3.372"  "4.421"  "40.000" "-0.496"
## [9] "-9.680" "2.609"  "50.906" "0.658"  "-0.496" "2.594"  "12.207"
```

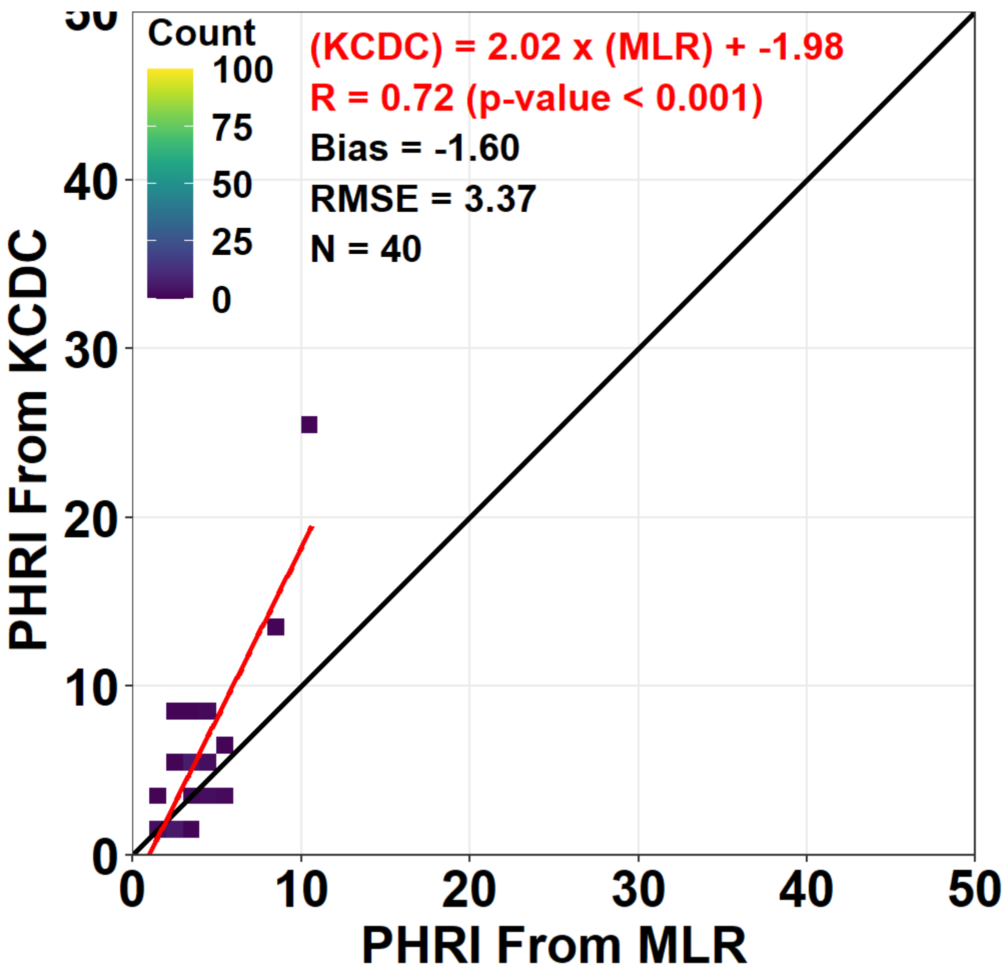
```
xcord = 1
ycord = seq(48, 0, -3)

saveImg = sprintf("%s/Img_%s_%s.png", globalVar$figConfig, serviceName, 1)

ggplot() +
  coord_fixed(ratio=1) +
  theme_bw() +
  geom_point(aes(X, Y)) +
  # stat_bin2d(binwidth = c(1, 1), aes(X, Y)) +
  # stat_bin2d(binwidth = c(5, 5), aes(X, Y)) +
  # scale_fill_gradientn(colours = cbViridis, limits=c(0, 100), na.value=cbViridis[1]) +
  annotate("text", x=xcord, y=ycord[1], label=paste0("(KCDC) = ", sprintf("%.2f",val[1])," x (MLR) + ", sprintf("%.2f",val[2])), size=5, hjust=0, color="red", fontface="bold", family=font) +
  # annotate("text", x=250, y=ycord[2], label=paste0("R = ", sprintf("%.2f",val[12]), " (p < 0.001) | Stdev = ", sprintf("%.2f",val[14])), size=5, hjust=0, color="red", family=font, fontface="bold") +
  annotate("text", x=xcord, y=ycord[2], label=paste0("R = ", sprintf("%.2f",val[12]), " (p-value < 0.001)"), size=5, hjust=0, color="red", family=font, fontface="bold") +
  # annotate("text", x=250, y=ycord[3], label=paste0("Stdev = ", sprintf("%.2f",val[14])), parse=F, size=5, hjust=0, family=font, fontface="bold") +
  # annotate("text", x=250, y=ycord[4], label=paste0("AMI = ", sprintf("%.3f",val[3]), " | CERES = ", sprintf("%.3f",val[4])), parse=F, size=5, hjust=0, family=font, fontface="bold") +
  annotate("text", x=xcord, y=ycord[3], label=paste0("Bias = ", sprintf("%.2f",val[8])), parse=F, size=5, hjust=0, family=font, fontface="bold") +
  # annotate("text", x=220, y=ycord[3], label=paste0("Bias = ", sprintf("%.2f",val[8])), parse=F, size=5, hjust=0, family=font, fontface="bold") +
  annotate("text", x=xcord, y=ycord[4], label=paste0("RMSE = ", sprintf("%.2f",val[10])), parse=F, size=5, hjust=0, family=font, fontface="bold") +
  # annotate("text", x=220, y=ycord[4], label=paste0("RMSE = ", sprintf("%.2f",val[10]), " (MPE = ", sprintf("%.2f",val[15])," %)"), parse=F, size=5, hjust=0, family=font, fontface="bold") +
  # annotate("text", x=2, y=ycord[5], label=paste0("MPE = ", sprintf("%.2f",val[15])," %"), parse=F, size=5, hjust=0, family=font, fontface="bold") +
  annotate("text", x=xcord, y=ycord[5], label=paste0("N = ", sprintf("%.0f",val[7]) ), size=5, hjust=0, color="black", family=font, fontface="bold") +
  geom_abline(intercept=0, slope=1, linetype=1, color="black", size=1.0) +
  stat_smooth(method="lm", color="red", se=F, aes(X, Y)) +
  scale_x_continuous(minor_breaks = seq(0, 50, by=10), breaks=seq(0, 50, by=10),  expand=c(0,0), limits=c(0, 50)) +
  scale_y_continuous(minor_breaks = seq(0, 50, by=10), breaks=seq(0, 50, by=10),  expand=c(0,0), limits=c(0, 50)) +
  labs(title = "" ) +
  labs(x = expression(paste(bold("PHRI From MLR"))),
       y = expression(paste(bold("PHRI From KCDC "))),
       fill = "Count") +
  theme(plot.title=element_text(face="bold", size=20, color="black")) +
  theme(axis.title.x = element_text(face="bold", size=19, colour="black")) +
  theme(axis.title.y = element_text(face="bold", size=19, colour="black", angle=90)) +
  theme(axis.text.x = element_text(face="bold", size=19, colour="black")) +
  theme(axis.text.y = element_text(face="bold", size=19, colour="black")) +
  theme(legend.title=element_text(face="bold", size=14, colour="black")) +
  theme(legend.position=c(0,1), legend.justification=c(0,0.96)) +
  theme(legend.key=element_blank()) +
  theme(legend.text=element_text(size=14, face="bold")) +
  theme(legend.background=element_blank()) +
  theme(text=element_text(family=font)) +
  theme(plot.margin=unit(c(0, 8, 0, 0),"mm")) +
  ggsave(filename = saveImg, width=6, height=6, dpi=600)
```



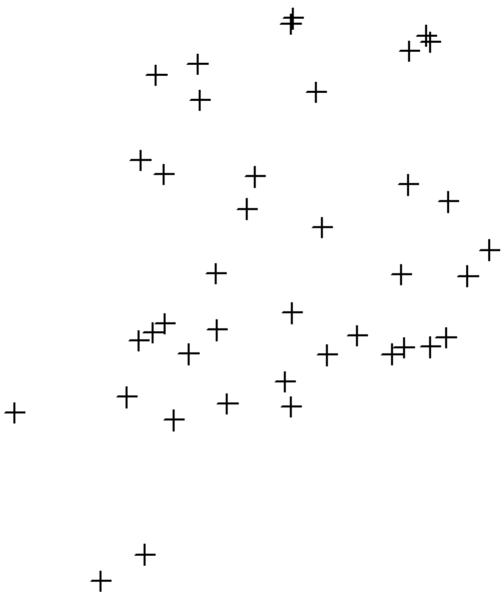
```
#**********************************************************
# 2차원 빈도분포 산포도 시각화
#**********************************************************
# 일별 (다중선형회귀 예측 vs 환자 관측값)
X = dayData$mlr
Y = dayData$kcdc

# 일별 (딥러닝 예측 vs 환자 관측값)
# X = dayData$dl
# Y = dayData$kcdc

# 연별 (다중선형회귀 예측 vs 환자 관측값)
# X = monthData$meanMlr
# Y = monthData$meanKcdc

# 연별 (딥러닝 예측 vs 환자 관측값)
# X = monthData$meanDl
# Y = monthData$meanKcdc


xcord = 10.5
ycord = seq(48, 0, -3)

val = fnStats(X, Y)
sprintf("%.3f", val)
```

```
##  [1] "2.017"   "-1.984"  "3.524"   "5.125"   "1.864"   "4.421"   "40.000"
##  [8] "-1.601"  "-31.245" "3.368"   "65.724"  "0.723"   "-1.601"  "3.001"
## [15] "-11.837"
```

```
saveImg = sprintf("%s/Img_%s_%s.png", globalVar$figConfig, serviceName, 2)

ggplot() +
  coord_fixed(ratio=1) +
  theme_bw() +
  # geom_point(aes(X, Y)) +
  stat_bin2d(binwidth = c(1, 1), aes(X, Y)) +
  # stat_bin2d(binwidth = c(5, 5), aes(X, Y)) +
  scale_fill_gradientn(colours = cbViridis, limits=c(0, 100), na.value=cbViridis[1]) +
  annotate("text", x=xcord, y=ycord[1], label=paste0("(KCDC) = ", sprintf("%.2f",val[1])," x (MLR) + ", sprintf("%.2f",val[2])), size=5, hjust=0, color="red", fontface="bold", family=font) +
  # annotate("text", x=xcord, y=ycord[1], label=paste0("(Val) = ", sprintf("%.2f",val[1])," x (Pred) + ", sprintf("%.2f",val[2])), size=5, hjust=0, color="red", fontface="bold", family=font) +
  # annotate("text", x=250, y=ycord[2], label=paste0("R = ", sprintf("%.2f",val[12]), " (p < 0.001) | Stdev = ", sprintf("%.2f",val[14])), size=5, hjust=0, color="red", family=font, fontface="bold") +
  annotate("text", x=xcord, y=ycord[2], label=paste0("R = ", sprintf("%.2f",val[12]), " (p-value < 0.001)"), size=5, hjust=0, color="red", family=font, fontface="bold") +
  # annotate("text", x=250, y=ycord[3], label=paste0("Stdev = ", sprintf("%.2f",val[14])), parse=F, size=5, hjust=0, family=font, fontface="bold") +
  # annotate("text", x=250, y=ycord[4], label=paste0("AMI = ", sprintf("%.3f",val[3]), " | CERES = ", sprintf("%.3f",val[4])), parse=F, size=5, hjust=0, family=font, fontface="bold") +
  # annotate("text", x=xcord, y=ycord[3], label=paste0("Bias = ", sprintf("%.2f",val[8]), " (", sprintf("%.2f",val[9])," %)"), parse=F, size=5, hjust=0, family=font, fontface="bold") +
  annotate("text", x=xcord, y=ycord[3], label=paste0("Bias = ", sprintf("%.2f",val[8])), parse=F, size=5, hjust=0, family=font, fontface="bold") +
  # annotate("text", x=220, y=ycord[3], label=paste0("Bias = ", sprintf("%.2f",val[8])), parse=F, size=5, hjust=0, family=font, fontface="bold") +
  # annotate("text", x=xcord, y=ycord[4], label=paste0("RMSE = ", sprintf("%.2f",val[10]), " (", sprintf("%.2f",val[11])," %)"), parse=F, size=5, hjust=0, family=font, fontface="bold") +
  annotate("text", x=xcord, y=ycord[4], label=paste0("RMSE = ", sprintf("%.2f",val[10])), parse=F, size=5, hjust=0, family=font, fontface="bold") +
  # annotate("text", x=220, y=ycord[4], label=paste0("RMSE = ", sprintf("%.2f",val[10]), " (MPE = ", sprintf("%.2f",val[15])," %)"), parse=F, size=5, hjust=0, family=font, fontface="bold") +
  # annotate("text", x=2, y=ycord[5], label=paste0("MPE = ", sprintf("%.2f",val[15])," %"), parse=F, size=5, hjust=0, family=font, fontface="bold") +
  annotate("text", x=xcord, y=ycord[5], label=paste0("N = ", sprintf("%.0f", val[7])), size=5, hjust=0, color="black", family=font, fontface="bold") +
  geom_abline(intercept=0, slope=1, linetype=1, color="black", size=1.0) +
  stat_smooth(method="lm", color="red", se=F, aes(X, Y)) +
  scale_x_continuous(minor_breaks = seq(0, 50, by=10), breaks=seq(0, 50, by=10),  expand=c(0,0), limits=c(0, 50)) +
  scale_y_continuous(minor_breaks = seq(0, 50, by=10), breaks=seq(0, 50, by=10),  expand=c(0,0), limits=c(0, 50)) +
  labs(
    x = expression(paste(bold("PHRI From MLR")))
    , y = expression(paste(bold("PHRI From KCDC ")))
    , fill = "Count"
    , title = NULL
    ) +
  theme(
    plot.title=element_text(face="bold", size=20, color="black")
    , axis.title.x = element_text(face="bold", size=19, colour="black")
    , axis.title.y = element_text(face="bold", size=19, colour="black", angle=90)
    , axis.text.x = element_text(face="bold", size=19, colour="black")
    , axis.text.y = element_text(face="bold", size=19, colour="black")
    , legend.title=element_text(face="bold", size=14, colour="black")
    , legend.position=c(0,1), legend.justification=c(0, 0.96)
    , legend.key=element_blank()
    , legend.text=element_text(size=14, face="bold")
    , legend.background=element_blank()
    # , text=element_text(family=font)
    , plot.margin=unit(c(0, 8, 0, 0),"mm")
    ) +
  ggsave(filename = saveImg, width=6, height=6, dpi=600 )
```



```
#*******************************************************
# 지도 매핑
#*******************************************************
# 일별
mapData = dayData

# 월별
# mapData = monthData

saveImg = sprintf("%s/Img_%s_%s.png", globalVar$figConfig, serviceName, 2)

coordinates(mapData)  = ~lon + lat
plot(mapData)
```



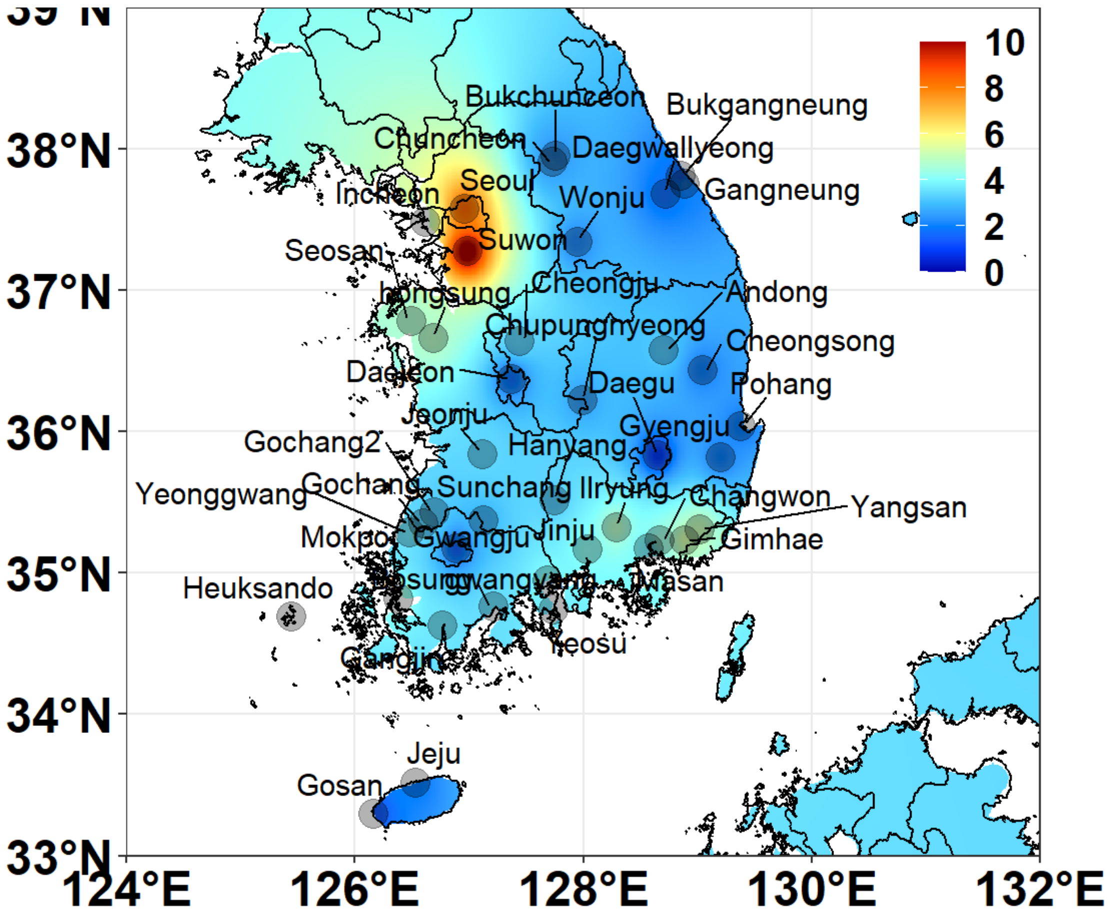
```
mapDataL1 = mapData %>%
  as.tibble()

globalVar$mapConfig
```

```
## [1] "E:/04. TalentPlatform/Github/TalentPlatform-R/CONFIG/MAP_INFO"
```

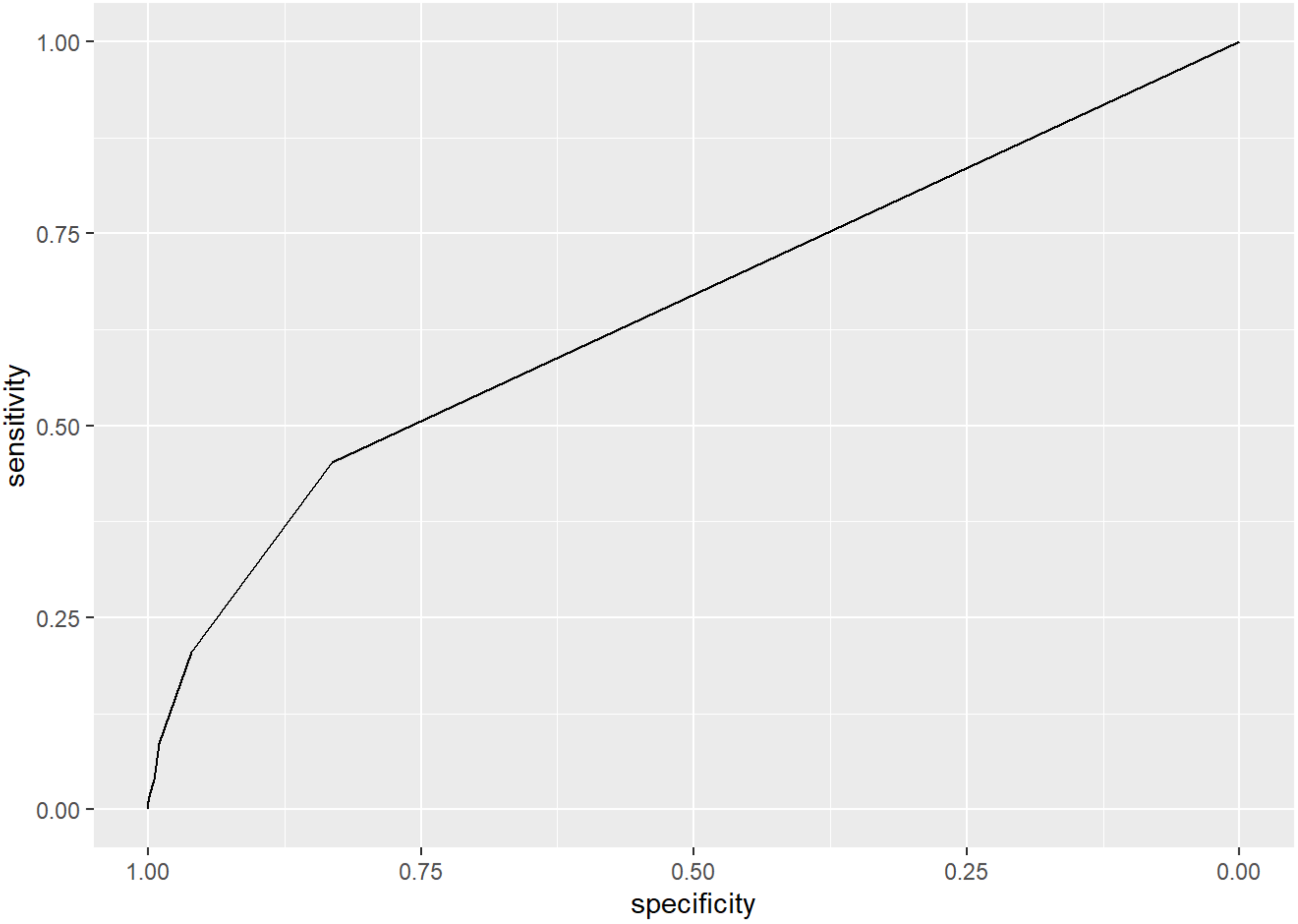
```
mapKor = read_sf(paste(globalVar$mapConfig, "gadm36_KOR_shp/gadm36_KOR_1.shp", sep = "/"))
mapPrk = read_sf(paste(globalVar$mapConfig, "gadm36_PRK_shp/gadm36_PRK_1.shp", sep = "/"))
mapJpn = read_sf(paste(globalVar$mapConfig, "gadm36_JPN_shp/gadm36_JPN_1.shp", sep = "/"))

yRange = as.numeric(c(33, 39))    # min/max latitude  of the interpolation area
xRange = as.numeric(c(124, 132))  # min/max longitude of the interpolation area

# expand points to grid
gridData = expand.grid(
  x = seq(from = xRange[1], to = xRange[2], by = 0.01)
  , y = seq(from = yRange[1], to = yRange[2], by = 0.01)
  )
coordinates(gridData) = ~ x + y
gridded(gridData) = TRUE

#****************************************
# 공간 내삽
#****************************************

#++++++++++++++++++++++
# 일별
#++++++++++++++++++++++
# 일별 다중선형회귀 예측
spData = gstat::idw(formula = mlr ~ 1, locations = mapData, newdata = gridData)
```

```
## [inverse distance weighted interpolation]
```

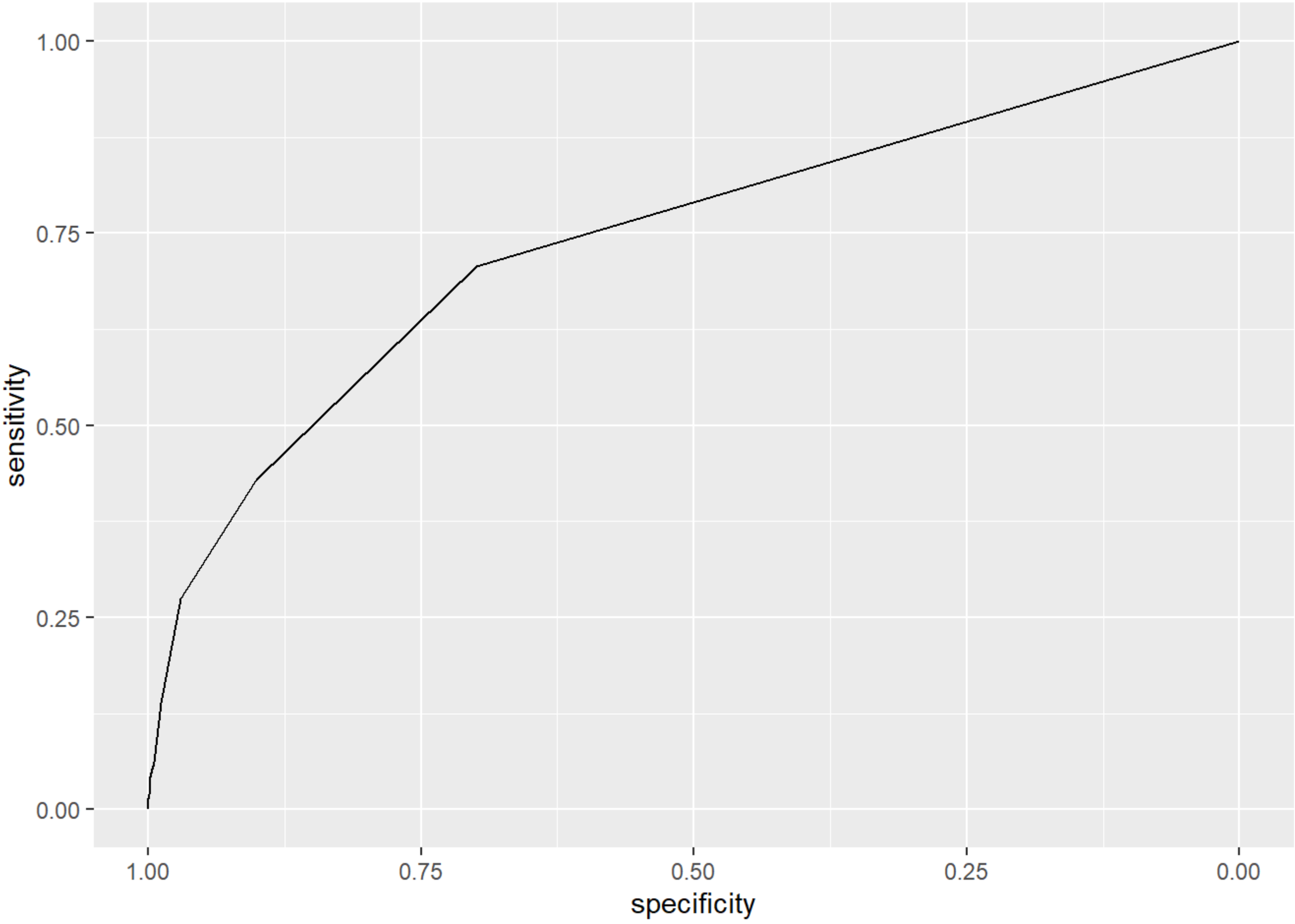
```
# 일별 딥러닝 예측
# spData = gstat::idw(formula = dl ~ 1, locations = mapData, newdata = gridData)

# 일별 환자 관측
# spData = gstat::idw(formula = kcdc ~ 1, locations = mapData, newdata = gridData)

#+++++++++++++++++++++
# 월별
#+++++++++++++++++++++
# 월별 다중선형회귀 예측
# spData = gstat::idw(formula = meanMlr ~ 1, locations = mapData, newdata = gridData)

# 월별 딥러닝 예측
# spData = gstat::idw(formula = meanKcdc ~ 1, locations = mapData, newdata = gridData)

# 월별 딥러닝 예측
# spData = gstat::idw(formula = meanKcdc ~ 1, locations = mapData, newdata = gridData)

# 월별 상관계수 분포
# spData = gstat::idw(formula = cor ~ 1, locations = mapData, newdata = gridData)

# 월별 평균제곱근오차 분포
# spData = gstat::idw(formula = rmse ~ 1, locations = mapData, newdata = gridData)

spDataL1 = spData %>%
  as.data.frame() %>%
  dplyr::rename(
    "lon" = "x"
    , "lat"  = "y"
    , "pred" = "var1.pred"
  ) %>%
  dplyr::mutate(
    isMaskLand = metR::MaskLand(lon, lat, mask = "world")
  ) %>%
  dplyr::filter(isMaskLand == TRUE)

saveImg = sprintf("%s/Img_%s_%s.png", globalVar$figConfig, serviceName, 3)

ggplot() +
  coord_fixed(ratio = 1.1) +
  # coord_fixed(1.3) +
  theme_bw() +
  geom_tile(data = spDataL1, aes(x = lon, y = lat, fill = pred)) +
  scale_fill_gradientn(colours = cbMatlab, limits=c(0, 10), breaks = seq(0, 10, 2), na.value = cbMatlab[length(cbMatlab)]) + # Pred, Val
  # scale_fill_gradientn(colours = cbSpectral, limits=c(0, 10), na.value = cbSpectral[length(cbSpectral)], breaks = seq(0, 10, 2)) + # RMSE
  # scale_fill_gradientn(colours = cbPlasma, limits=c(0.4, 1), na.value = cbPlasma[length(cbPlasma)], breaks = seq(0.4, 1, 0.2)) + # R
  # geom_point(aes(x = long, y = lat, colour=obs), data = data_L6, size=5, alpha=0.3, show.legend = FALSE) +
  geom_point(aes(x = lon, y = lat), colour="black", data = mapDataL1, size=5, alpha=0.3, show.legend = FALSE) +
  # geom_text(aes(x=lon, y=lat, label=obs, colour=obs), hjust=-0.25, vjust=0.5, nudge_x=0, nudge_y=0, size=5, colour='black', data=rsReultL1, family = font) +
  ggrepel::geom_text_repel(aes(x=lon, y=lat, label=stationName, colour=stationName), point.padding = 0.25, box.padding = 0.25,  nudge_y = 0.1, data=mapDataL1, size = 4, colour="black") +

  # geom_label(aes(x=lon, y=lat, label=obs, colour=obs), hjust=-0.25, vjust=0.5, nudge_x=0, nudge_y=0, size=5,
  # data=data_L4, show.legend = FALSE) +
  # coord_sf() +
  # ggplot(data = world) +
  # geom_sf() +
  # coord_sf(xlim = c(124, 132), ylim = c(33, 39), expand = FALSE)
  # geom_path(aes(x=long, y=lat, group=group), data = world, colour="black") +
  # geom_polygon(data = shore, aes(x=long, y = lat, group = group), color = "black", fill = NA) +
 # geom_polygon(data = usa, aes(x = long, y = lat, group = group), fill = NA, color = "red") +
 # scale_x_continuous(breaks = longitude, labels = x_longitude, expand=c(0,0), limits=c(125, 130)) +
# scale_y_continuous(breaks = latitude,  labels = y_latitude,  expand=c(0,0), limits=c(33, 39)) +
# scale_x_continuous(breaks = longitude, labels = x_longitude, expand=c(0,0), limits=c(124, 132)) +
# scale_y_continuous(breaks = latitude,  labels = y_latitude,  expand=c(0,0), limits=c(33, 39)) +
# ggplot2::scale_x_continuous(limits=c(124, 132)) +
# ggplot2::scale_y_continuous(limits=c(33, 39)) +
# geom_polygon(data = shore, aes(x=long, y = lat, group = group), color = "black", fill = NA) +
# geom_sf(data=map,  color = "black", fill=NA) +
 geom_sf(data = mapKor, color = "black", fill = NA) +
 geom_sf(data = mapPrk, color = "black", fill = NA) +
 geom_sf(data = mapJpn, color = "black", fill = NA) +
 metR::scale_x_longitude(expand = c(0, 0), breaks=seq(124, 132, 2), limits=c(124, 132)) +
 metR::scale_y_latitude(expand = c(0, 0), breaks=seq(32, 40, 1), limits=c(33, 39)) +
 theme(
  plot.title = element_text(face="bold", size=18, color="black")
  , axis.title.x = element_text(face="bold", size=18, colour="black")
  , axis.title.y = element_text(face="bold", size=18, colour="black", angle=90)
  , axis.text.x  = element_text(face="bold", size=18, colour="black")
  , axis.text.y  = element_text(face="bold", size=18, colour="black")
  , legend.position=c(1, 1), legend.justification=c(1, 1)
  , legend.key=element_blank()
  , legend.text=element_text(size=14, face="bold")
  , legend.title=element_text(face="bold", size=14, colour="black")
  , legend.background=element_blank()
  # , text=element_text(family = font)
  , plot.margin=unit(c(0, 8, 0, 0), "mm")
 ) +
 labs(x = NULL, y = NULL, fill = NULL, colour = NULL, title  = NULL) +
 ggsave(filename = saveImg, width=8, height=10, dpi=600)
```



```
#******************************************************************
# 검증
#******************************************************************
# 일별 (다중선형회귀 예측 vs 환자 관측)
# rocDayData = roc(as.integer(dayData$mlr), dayData$kcdc)
# pROC::ggroc(rocDayData)

# 월별 (다중선형회귀 예측 vs 환자 관측)
# rocMontyData = roc(as.integer(monthData$meanMlr), monthData$meanKcdc)
# pROC::ggroc(rocMontyData)

# 연별 (다중선형회귀 예측 vs 환자 관측)
rocYearData = roc(as.integer(yearData$mlr), yearData$kcdc)
pROC::ggroc(rocYearData)
```



```
# 일별 (딥러닝 예측 vs 환자 관측)
# rocDayData = roc(as.integer(dayData$dl), dayData$kcdc)
# pROC::ggroc(rocDayData)

# 월별 (딥러닝 예측 vs 환자 관측)
# rocMontyData = roc(as.integer(monthData$meanDl), monthData$meanKcdc)
# pROC::ggroc(rocMontyData)

# 연별 (딥러닝 예측 vs 환자 관측)
rocYearData = roc(as.integer(yearData$dl), yearData$kcdc)
pROC::ggroc(rocYearData)
```


sensitivity
1.00
0.75
0.50
0.25
0.00
specificity
1.00
0.75
0.50
0.25
0.00