```
# install.packages("elasticnet")
# install.packages("tseries")
# install.packages("zoo")
# install.packages("forecast")
# install.packages("chron")
# install.packages("RColorBrewer")
# install.packages("lattice")
# install.packages("ncdf4")

library("readxl")
library("lubridate")
library("proto")
library("gsubfn")
library("RSQLite")
library("sqldf")
library("tidyverse")
library("ggplot2")
library("mice")
library("caret")
library("dplyr")
library("reshape2")
library("janitor")
library("tidyr")
library("MASS")
library("car")
library("Metrics")
library("glmnet")
library("xgboost")
library("gbm")
library("mboost")
library("elasticnet")
```

```
library("tseries")
library("zoo")
library("forecast")
library("chron")
library("RColorBrewer")
library("lattice")
library("ncdf4")
#Load the data
met <- read.csv("./Assignment//Team/Assignment 1//Dataset//MET_HadCRUT4_Data.csv")
nasa  <- read.csv("./Assignment//Team/Assignment 1//Dataset//NASA.csv")

head(met)
head(nasa)

#NASA
#removing the last 5 columns
nasa = subset(nasa, select = -c(14,15,16,17,18,19) )
head(nasa)
df2 <- melt(nasa, id=c("Year"))
df2
final <- separate(data = df2, col = variable, into = c("Month"), sep = "\\.")
final

#convert month to numeric
final$Month <- match(final$Month,month.abb)

final$date <- paste(final$Year,final$Month,"1",sep="-")
head(final)
final$date <- as.Date(final$date, format = "%Y-%m-%d")
head(final)
```

```
final = final[,!(names(final) %in% c("Month"))]
final = final[,!(names(final) %in% c("Year"))]

# reorder by column name
final <- final[c("date", "value")]

#write to csv
colnames(final) <- c("Date", "Temp_Diff")
write.csv(final, "nasa_clean.csv", row.names=FALSE)



#met
#make into date
head(met)

#removing the first columns
met = subset(met, select = -c(1) )
head(met)

met$date <- paste(met$Year,met$Month,"1",sep="-")
head(met)
met$date <- as.Date(met$date, format = "%Y-%m-%d")
head(met)

met = met[,!(names(met) %in% c("Month"))]
met = met[,!(names(met) %in% c("Year"))]

# reorder by column name
met <- met[c("date", "Median_Temp_Difference")]
head(met)
#write to csv
```

```
colnames(met) <- c("Date", "Temp_Diff")
write.csv(met, "met_clean.csv", row.names=FALSE)

#for Kingston!
#Load the data
king <- read.csv("./Assignment//Team/Assignment 1//Dataset//KingstonCRU.csv")
head(king)

#removing the last 5 columns
king = subset(king, select = -c(14,15,16,17,18) )
head(king)
df2 <- melt(king, id=c("Year"))
df2
final <- separate(data = df2, col = variable, into = c("Month"), sep = "\\.")
final

#write to csv
#write to csv
colnames(final) <- c("Year","Month", "Temp_Diff")
write.csv(final, "kingston_clean.csv", row.names=FALSE)
```