1.2-HDD-Test.R

frm

2023-01-12

# 1.2 HDD Test.R  
# rmarkdown::render("1.2 HDD Test.R", "word\_document")  
# HDD - heating degree day   
# "The number of degrees that a day's average temperature is below 65 degrees   
# Fahrenheit (approx. 18 degrees Celsius), the temperature below which   
# buildings need to be heated. The price of weather derivatives traded in the   
# winter is based on an index made up of monthly HDD values." Investopedia  
rm(list = ls()) # Take out the Environment "trash"  
# library(date)  
# library(stats) # lm() - linear model  
# Read a file and create data frame  
HDDFile <- read.table("ATLHDD.csv", header = TRUE, sep = ",")  
head(HDDFile, 5) # Show first 5 lines in Console

## Date HDD Estimated  
## 1 1/1/17 16.6 0  
## 2 1/2/17 11.3 0  
## 3 1/3/17 5.4 0  
## 4 1/4/17 16.4 0  
## 5 1/5/17 20.7 0

tail(HDDFile, 5) # Show first 5 lines in Console

## Date HDD Estimated  
## 377 1/12/18 13.0 0  
## 378 1/13/18 35.1 0  
## 379 1/14/18 36.1 0  
## 380 1/15/18 28.8 0  
## 381 1/16/18 26.9 0

is.data.frame(HDDFile) # TRUE: HDDFile is a data frame

## [1] TRUE

is.numeric(HDDFile$Date) # FALSE: Imported as a factor

## [1] FALSE

is.factor(HDDFile$Date) # TRUE: A categorical variable, idealy limited in number (not here)

## [1] FALSE

is.numeric(HDDFile$HDD) # TRUE: Numeric data suitable for mathematical calculations

## [1] TRUE

is.numeric(HDDFile$Estimated) # TRUE

## [1] TRUE

is.ordered(HDDFile$Date) # FALSE: It is stored as a factor variable

## [1] FALSE

is.character(HDDFile$Date) # FALSE

## [1] TRUE

max(HDDFile$Estimated) # Estimated variable is not relevant here, returns the max

## [1] 6

min(HDDFile$Estimated) # Minimum reported to the console

## [1] 0

# Often data sets have columns that need to be deleted  
HDDFile$Estimated <- NULL # Estimated column is removed  
head(HDDFile, 5)

## Date HDD  
## 1 1/1/17 16.6  
## 2 1/2/17 11.3  
## 3 1/3/17 5.4  
## 4 1/4/17 16.4  
## 5 1/5/17 20.7

HDDFile$TDate <- as.character(HDDFile$Date) # Convert factor to character  
is.factor(HDDFile$TDate) # Now FALSE

## [1] FALSE

is.character(HDDFile$TDate) # Factor Date converted to character TDate

## [1] TRUE

# Convert character date (TDate) to Julian date but display in calendar format  
# Note default output format is YYYY-MM-DD, format = "..." denotes input  
HDDFile$TDate2 <- as.Date(HDDFile$TDate, format = "%m/%d/%y")  
# Insights on dates  
head(HDDFile, 5)

## Date HDD TDate TDate2  
## 1 1/1/17 16.6 1/1/17 2017-01-01  
## 2 1/2/17 11.3 1/2/17 2017-01-02  
## 3 1/3/17 5.4 1/3/17 2017-01-03  
## 4 1/4/17 16.4 1/4/17 2017-01-04  
## 5 1/5/17 20.7 1/5/17 2017-01-05

# Insights on dates  
as.integer(HDDFile$Date) # Nonsense and definitely not a Julian date

## Warning: NAs introduced by coercion

## [1] NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA  
## [28] NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA  
## [55] NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA  
## [82] NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA  
## [109] NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA  
## [136] NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA  
## [163] NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA  
## [190] NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA  
## [217] NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA  
## [244] NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA  
## [271] NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA  
## [298] NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA  
## [325] NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA  
## [352] NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA  
## [379] NA NA NA

as.integer(HDDFile$TDate) # Cannot convert character string to integer, NAs

## Warning: NAs introduced by coercion

## [1] NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA  
## [28] NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA  
## [55] NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA  
## [82] NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA  
## [109] NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA  
## [136] NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA  
## [163] NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA  
## [190] NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA  
## [217] NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA  
## [244] NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA  
## [271] NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA  
## [298] NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA  
## [325] NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA  
## [352] NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA  
## [379] NA NA NA

as.integer(HDDFile$TDate2) # Julian dates, base = 1/1/1970

## [1] 17167 17168 17169 17170 17171 17172 17173 17174 17175 17176 17177 17178 17179  
## [14] 17180 17181 17182 17183 17184 17185 17186 17187 17188 17189 17190 17191 17192  
## [27] 17193 17194 17195 17196 17197 17198 17199 17200 17201 17202 17203 17204 17205  
## [40] 17206 17207 17208 17209 17210 17211 17212 17213 17214 17215 17216 17217 17218  
## [53] 17219 17220 17221 17222 17223 17224 17225 17226 17227 17228 17229 17230 17231  
## [66] 17232 17233 17234 17235 17236 17237 17238 17239 17240 17241 17242 17243 17244  
## [79] 17245 17246 17247 17248 17249 17250 17251 17252 17253 17254 17255 17256 17257  
## [92] 17258 17259 17260 17261 17262 17263 17264 17265 17266 17267 17268 17269 17270  
## [105] 17271 17272 17273 17274 17275 17276 17277 17278 17279 17280 17281 17282 17283  
## [118] 17284 17285 17286 17287 17288 17289 17290 17291 17292 17293 17294 17295 17296  
## [131] 17297 17298 17299 17300 17301 17302 17303 17304 17305 17306 17307 17308 17309  
## [144] 17310 17311 17312 17313 17314 17315 17316 17317 17318 17319 17320 17321 17322  
## [157] 17323 17324 17325 17326 17327 17328 17329 17330 17331 17332 17333 17334 17335  
## [170] 17336 17337 17338 17339 17340 17341 17342 17343 17344 17345 17346 17347 17348  
## [183] 17349 17350 17351 17352 17353 17354 17355 17356 17357 17358 17359 17360 17361  
## [196] 17362 17363 17364 17365 17366 17367 17368 17369 17370 17371 17372 17373 17374  
## [209] 17375 17376 17377 17378 17379 17380 17381 17382 17383 17384 17385 17386 17387  
## [222] 17388 17389 17390 17391 17392 17393 17394 17395 17396 17397 17398 17399 17400  
## [235] 17401 17402 17403 17404 17405 17406 17407 17408 17409 17410 17411 17412 17413  
## [248] 17414 17415 17416 17417 17418 17419 17420 17421 17422 17423 17424 17425 17426  
## [261] 17427 17428 17429 17430 17431 17432 17433 17434 17435 17436 17437 17438 17439  
## [274] 17440 17441 17442 17443 17444 17445 17446 17447 17448 17449 17450 17451 17452  
## [287] 17453 17454 17455 17456 17457 17458 17459 17460 17461 17462 17463 17464 17465  
## [300] 17466 17467 17468 17469 17470 17471 17472 17473 17474 17475 17476 17477 17478  
## [313] 17479 17480 17481 17482 17483 17484 17485 17486 17487 17488 17489 17490 17491  
## [326] 17492 17493 17494 17495 17496 17497 17498 17499 17500 17501 17502 17503 17504  
## [339] 17505 17506 17507 17508 17509 17510 17511 17512 17513 17514 17515 17516 17517  
## [352] 17518 17519 17520 17521 17522 17523 17524 17525 17526 17527 17528 17529 17530  
## [365] 17531 17532 17533 17534 17535 17536 17537 17538 17539 17540 17541 17542 17543  
## [378] 17544 17545 17546 17547

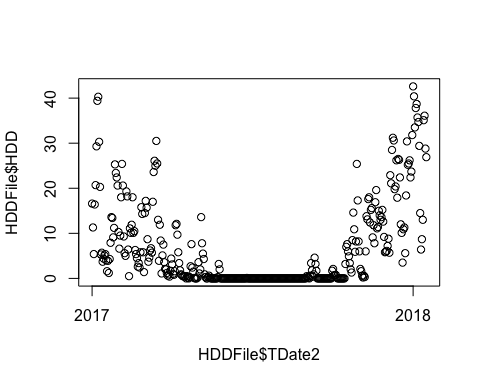
BaseDate = as.Date("1970-01-01")  
as.integer(BaseDate)

## [1] 0

head(HDDFile, 5)

## Date HDD TDate TDate2  
## 1 1/1/17 16.6 1/1/17 2017-01-01  
## 2 1/2/17 11.3 1/2/17 2017-01-02  
## 3 1/3/17 5.4 1/3/17 2017-01-03  
## 4 1/4/17 16.4 1/4/17 2017-01-04  
## 5 1/5/17 20.7 1/5/17 2017-01-05

# Some simple plots  
# plot(HDDFile$Date, HDDFile$HDD) # x-axis not formatted correctly, nonsense  
# plot(HDDFile$TDate, HDDFile$HDD) # TDate not formatted for plotting  
plot(HDDFile$TDate2, HDDFile$HDD)



# Thus, when programming, managing dates is tricky, and you must be careful