**OVERVIEW**

In the Stack overflow blog listed in this week’s Blackboard assignment there were several posts that piqued my interest.

<http://stackoverflow.com/questions/5963269/how-to-make-a-great-r-reproducible-example>

One involved using clipboard data to in R with code below:

test <-dput(read.table("clipboard",sep="\t",header=TRUE))

A second one was the ability to anonymize data when you

install.packages("SciencePo")

anonymize(test2)

**PROBLEM**

The idea to be able to reproduce the problem without exposing confidential data while also having the ability to use copy/paste data from Windows clipboard.

When testing the clipboard code all went well, but when I tried to anoymize the test data I got below error:

> anonymize(test2)

Error in salt(.x, .seed = .seed, .chars = .chars, .n\_chars = .n\_chars) :

Vector must be an atomic vector.

**DPUT TEST DATA**

> dput(test2)

structure(list(origin = c("EWR", "EWR", "EWR", "EWR", "EWR",

"EWR"), year = c(2013, 2013, 2013, 2013, 2013, 2013), month = c(1,

1, 1, 1, 1, 1), day = c(1L, 1L, 1L, 1L, 1L, 1L), hour = 1:6,

temp = c(39.02, 39.02, 39.02, 39.92, 39.02, 37.94), dewp = c(26.06,

26.96, 28.04, 28.04, 28.04, 28.04), humid = c(59.37, 61.63,

64.43, 62.21, 64.43, 67.21), wind\_dir = c(270, 250, 240,

250, 260, 240), wind\_speed = c(10.35702, 8.05546, 11.5078,

12.65858, 12.65858, 11.5078), wind\_gust = c(NA\_real\_, NA\_real\_,

NA\_real\_, NA\_real\_, NA\_real\_, NA\_real\_), precip = c(0, 0,

0, 0, 0, 0), pressure = c(1012, 1012.3, 1012.5, 1012.2, 1011.9,

1012.4), visib = c(10, 10, 10, 10, 10, 10), time\_hour = structure(c(1357020000,

1357023600, 1357027200, 1357030800, 1357034400, 1357038000

), class = c("POSIXct", "POSIXt"), tzone = "America/New\_York"),

date = structure(c(1357020000, 1357023600, 1357027200, 1357030800,

1357034400, 1357038000), class = c("POSIXct", "POSIXt"), tzone = "America/New\_York"),

temperature = c(53.7, 53.7, 53.7, 54.2, 53.7, 53.1)), row.names = c(NA,

6L), class = "data.frame")

**SESSION INFO**

> sessionInfo()

R version 3.5.1 (2018-07-02)

Platform: x86\_64-w64-mingw32/x64 (64-bit)

Running under: Windows >= 8 x64 (build 9200)

Matrix products: default

locale:

[1] LC\_COLLATE=English\_United States.1252

[2] LC\_CTYPE=English\_United States.1252

[3] LC\_MONETARY=English\_United States.1252

[4] LC\_NUMERIC=C

[5] LC\_TIME=English\_United States.1252

attached base packages:

[1] stats graphics grDevices utils datasets methods base

other attached packages:

[1] anonymizer\_0.2.0 nycflights13\_1.0.0

loaded via a namespace (and not attached):

[1] Rcpp\_0.12.17 digest\_0.6.15 rprojroot\_1.3-2 bitops\_1.0-6

[5] backports\_1.1.2 magrittr\_1.5 evaluate\_0.10.1 pillar\_1.2.3

[9] rlang\_0.2.1 stringi\_1.1.7 rmarkdown\_1.10 tools\_3.5.1

[13] stringr\_1.3.1 RCurl\_1.95-4.10 yaml\_2.1.19 rsconnect\_0.8.8

[17] compiler\_3.5.1 htmltools\_0.3.6 knitr\_1.20 tibble\_1.4.2

**EXPECTED OUTCOME**

The objective is to find a way to resolve the error code in anonymizer or write R code to anonymize data. The expectation is output data would be different from input data. This will allow a user to review code/data without seeing confidential data.