Error in as.POSIXlt.character(x, tz, ...) : character string is not in a standard unambiguous format

* Examples include 010010-99999-1978, 1979; error is a missing row in each of these; **many examples exist where missing rows cause lots of problems, this is just one example. In some cases, total\_char or usaf\_station or wban\_station is present, but nothing else is.**

I am removing some of the issues by the following code where I attempt to find and remove rows with missing and/or incorrect USAF, WBAN, time, and total\_char data. I have performed a bind\_rows operation on all yearly data from each station to make one file per station before I do this.

data <- read.csv(fname); data <- data[order(data$date),]

data$total\_chars <- as.integer(as.character(data$total\_chars))

data$usaf\_station <- as.integer(as.character(data$usaf\_station))

data$wban\_station <- as.integer(as.character(data$wban\_station))

data$time <- as.integer(data$time)

data <- subset(data,is.finite(usaf\_station) & is.finite(wban\_station) & is.finite(time) & total\_chars > 1)

data$date <- as.character(data$date)

After I run this code, I still get the following errors.

Error in read.table(file = file, header = header, sep = sep, quote = quote, :

more columns than column names

In addition: Warning message:

In scan(file = file, what = what, sep = sep, quote = quote, dec = dec, :

EOF within quoted string

This ‘more columns than column names error’ was from 697774-99999-2002. There are extra NA’s in some rows with no headers.

Error in read.table(file = file, header = header, sep = sep, quote = quote, :

duplicate 'row.names' are not allowed

In addition: Warning message:

In scan(file = file, what = what, sep = sep, quote = quote, dec = dec, :

EOF within quoted string

This ‘duplicate 'row.names' are not allowed’ was from 467425-99999-1986. This error seems to also be caused by extra NA’s in some rows with no headers.

I hope this helps. I can send data if you need it, but my files are big. Otherwise, a list of all possible headers would be very helpful because the headers vary a lot from station to station depending on what data is available. This would make combining or comparing data from/between stations easier because we would know what headers to check for then. Otherwise, it’s up to the user to try to find what header is associated with RH (if there is one), as one example (which is one thing I’m trying to do).