As of writing this blog-post, today is February 2nd, 2020. Or as I would say it, 2nd of February, 2020. There is nothing magical about it, it is just a sequence of numbers. On a boring Sunday evening, what could be more thrilling to look into this little bit further

Let’s kick R Studio and start writing a lot of useless stuff. First, we don’t need a function, but since, this is all about useless stuff, let’s make a useless [palindrome](https://en.wikipedia.org/wiki/Palindrome) function:

palindrome <- function(date){

identical(date, paste(rev(strsplit(date, "")[[1]]), collapse=""))

}

Once we have the function, let’s create a range of dates we want to explore. The range will be set from 1st of January 1000 until 1st of January 2500.

dd <- data.frame(seq(as.Date("1000/01/01"), as.Date("2500/01/01"), "days"))

I don’t want to get into useless debate whether the [Gregorian Calendar](https://en.wikipedia.org/wiki/Gregorian_calendar) was already accepted worldwide or not, but if you want to read the 16th Century history, click the link.

Now, the most useless part is a loop, that will take us from the pre-historic age all the way to the future:

#empty df

df <- data.frame(paldate=as.character())

#loop through all the dates

for (i in 1:nrow(dd)){

dat <- dd[i,1]

#Year Month Day format

dat <- format(dat, "%Y%m%d")

#Year Day month format

#dat <- format(dat, "%Y%d%m")

if (palindrome(dat) == TRUE) {

df <- rbind(df, data.frame(paldate=as.character(dat)))

}

}

**Issues**

Following are the Issues I am having with this pretty useless Palindrome fact.

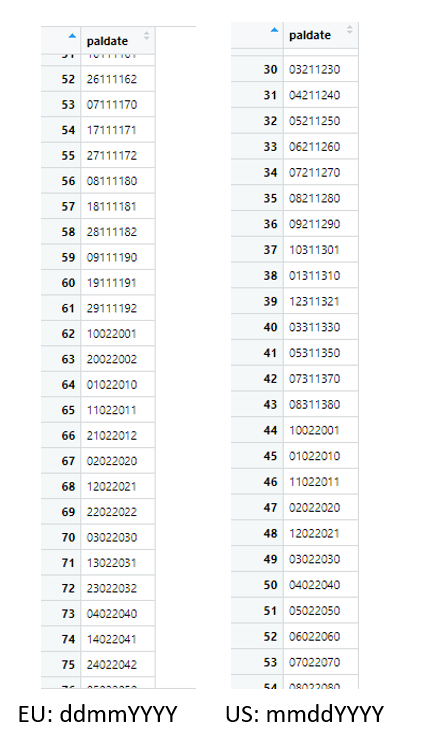
1. Year Format  – Abbreviated or non-abbreviated format; ’20 or 2020?
2. Leading zeros – Today’s date (02.02.) would have never been a palindrome if not for the leading zeros. Otherwise, it would have just been 2.2.
3. American vs. European Date format. Today’s date is fully interchangeable, but what if the date would have been February 22nd, of 2022?
   1. European: 22.02.2022 -> Is Palindrome
   2. American: 02.22.2022 -> Is not Palindrome.

**Useless statistics**

Issues aside, let’s do the comparison for the fun between American and European date format.

Date range will remain the same; from 01/01/1000 until 01/01/2500.

Number of Palidromes between the US and EU date format is 79 vs. 121:



Let’s check the distribution between the two formats and put them on graphs for easier visualization with ggplot:

library(plyr)

library(ggplot2)

library(gridExtra)

df\_ALL\_m\_d <- ddply(df\_ALL, "Region", summarise, grp.mean=median(Day))

ggplot(df\_ALL, aes(x=Day, color=Region, fill=Region)) +

geom\_histogram(fill="white", position="identity")+

theme(legend.position="top")

pd<-ggplot(df\_ALL, aes(x=Day, color=Region,fill=Region)) +

geom\_histogram(fill="white", position="identity", binwidth = 2)+

geom\_vline(data=df\_ALL\_m\_d, aes(xintercept=grp.mean, color=Region),

linetype="dashed")+

theme(legend.position="top")

pd

grid.arrange(p, pm, pd, ncol = 1)

And the distribution comparison is:

Is is clear from the graphs that the time formatting plays significant role for a particular date to be a palindrome or not.

Since we can see from the graphs that days and months are significantly different between the EU and US formats (months for EU format are in appearing only on January, February, November and December, where US format are ranging through all the months, and days are the exact opposite).

With following R code:

## Get distribution for the days in the year (13th day of the year,

## 241st day of the year, etc)

v <- ggplot(df\_ALL, aes(y=DaysDiff, x=Region)) +

geom\_violin(trim=FALSE) +

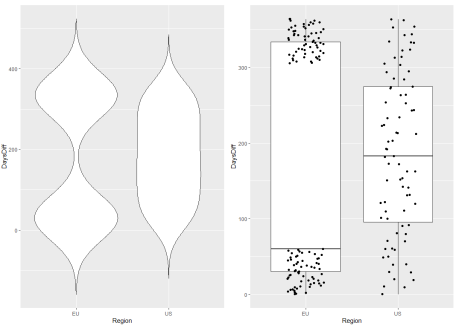
geom\_dotplot(binaxis='DaysDiff', stackdir='center', dotsize=1)

b <- ggplot(df\_ALL, aes(x=Region, y=DaysDiff)) + geom\_boxplot()

b <- b + geom\_jitter(shape=16, position=position\_jitter(0.2))

grid.arrange(v, b, nrow = 1)

We can generate this:



We can again see, that the US distribution of the differences from the Palindrome date until the end of the year (difference is the number of days between December, 31st and the palindrome date), is in EU date format bimodal where as in US format evenly distributed.

Pretty useless, I guess.