# Coursera assignment Plotly

Lei

21 September 2017

In this work, I present some interactive graphs using plotly. The data is a dummy about gene mutations along with time.
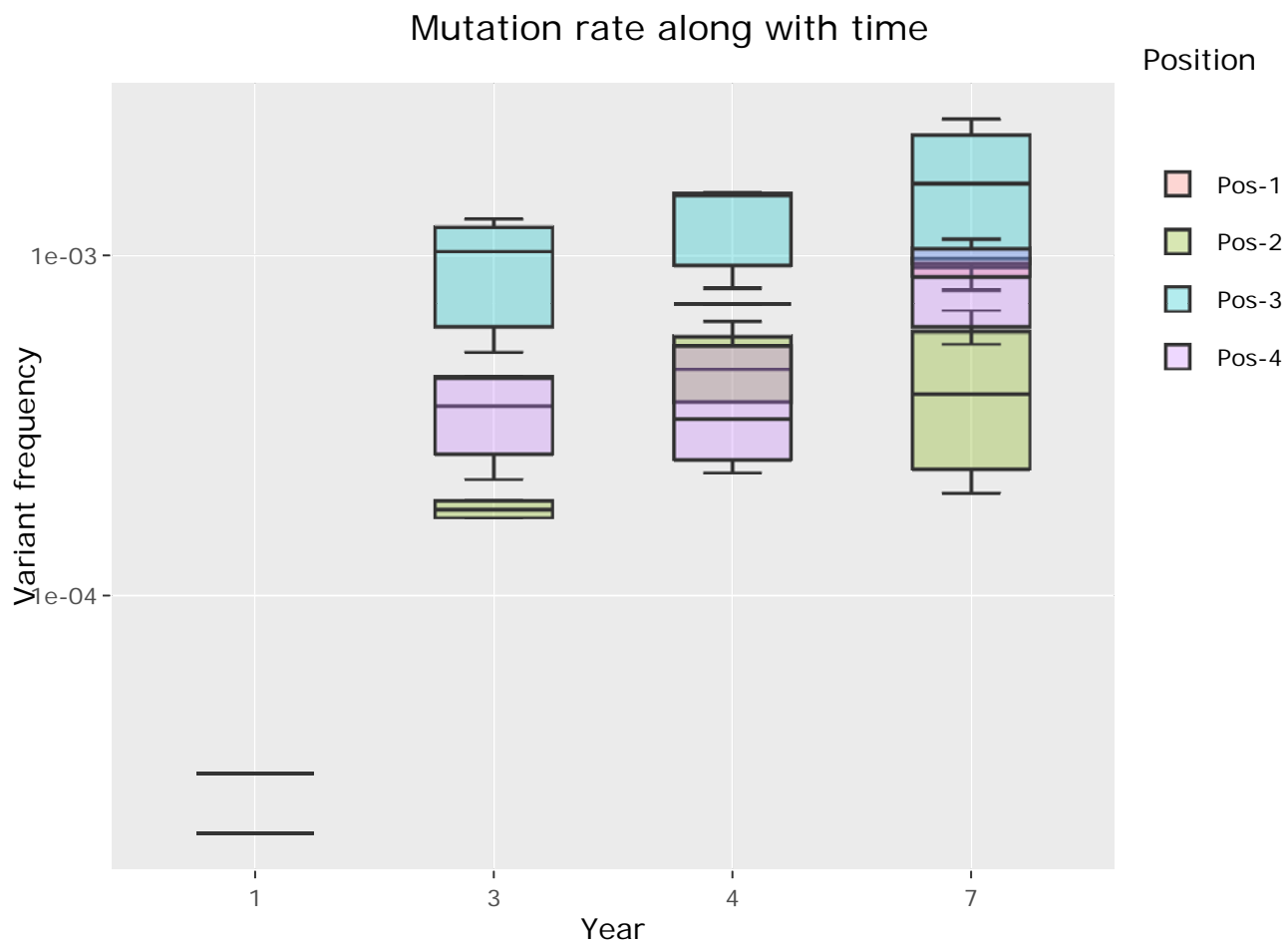
Read in data

```
setwd("~/Rstudio/")
mydata <- read.csv("mydata.csv")
```

```
library(ggplot2)
library("plotly")
# boxplot:
p <- ggplot(mydata, aes(x=as.factor(Year), y=mut_rate, fill=Position)) +
    geom_boxplot(alpha=0.3) + scale_y_log10() +
    ggtitle("Mutation rate along with time") + xlab("Year") + ylab("Variant frequen
cy")
p
```
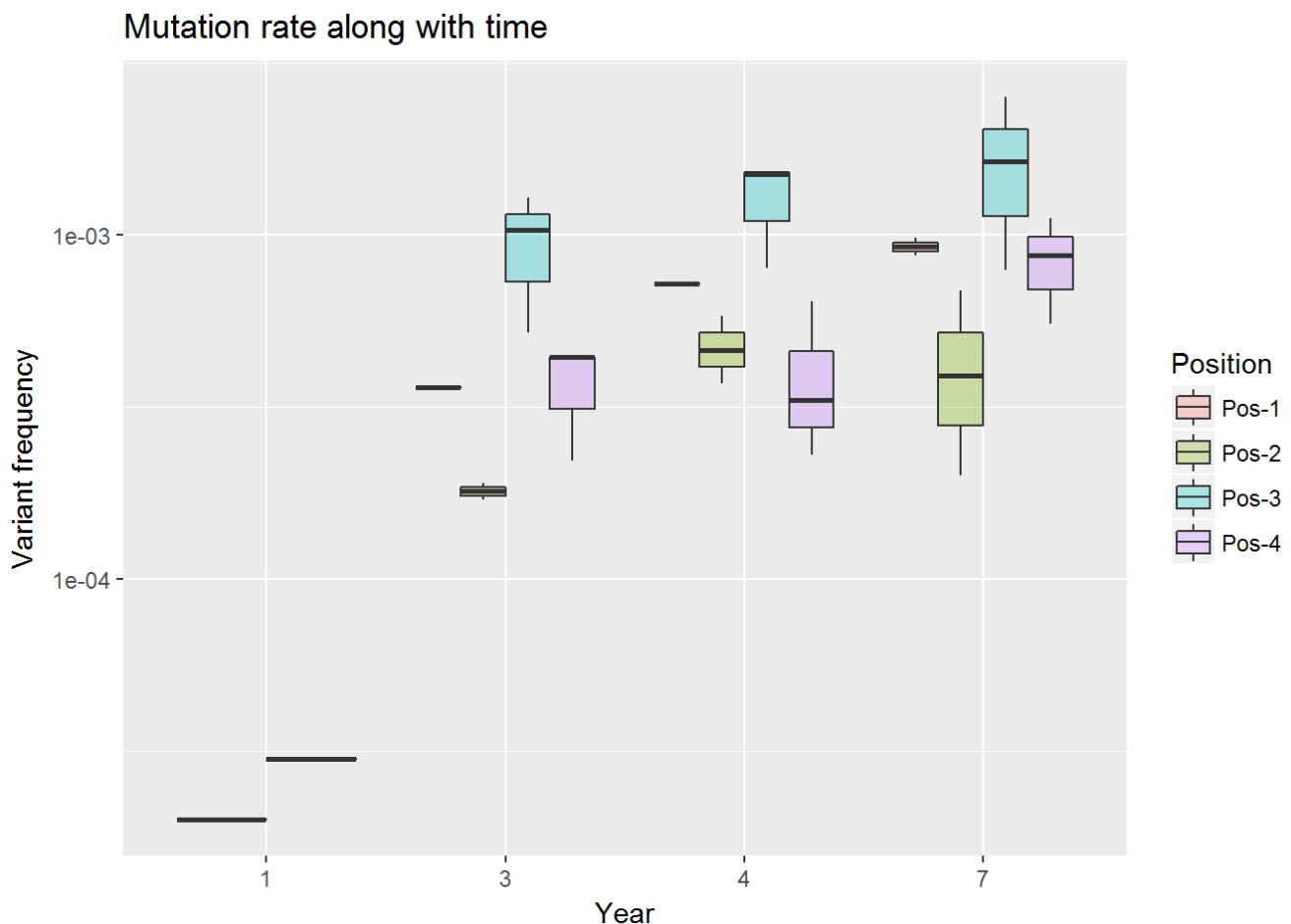


```
#interactive plot:
ggplotly(p)
```

Mutation rate along with time

Line plots:

```r
# line plot:
library(dplyr)
data_percent <- mutate(mydata, freq_percent = mut_rate * 100)

# transform data frame to data.table:
library(data.table)
dt.data_percent <- as.data.table(data_percent)

# set 2 keys for each data.table:
setkey(dt.data_percent, Position, Year)

#make a tmp data for plot only:
tmp <- dt.data_percent[,mean(mut_rate), by = key(dt.data_percent)]
setnames(tmp, "V1", "frequency_mean")


# Line plots:
p2 <- ggplot(tmp) + geom_line(aes(x=Year, y = frequency_mean, colour = Position)) +
scale_y_log10() +
    ggtitle("Mutation along with time") + xlab("Year") + ylab("Variant frequency")
ggplotly(p2)
```

Alternatively, show sepeartely for each mutation:

```
p3 <- ggplot(dt.data_percent, aes(y=freq_percent, x = Year, group = Position)) +
    stat_summary(fun.y =mean, geom="line") + facet_wrap( ~ Position)+
    ggtitle("Mutation along with time") + xlab("Year") + ylab("Variant frequency")
ggplotly(p3)
```

The plotly functionalities are amazing!