Homework2

Jun Rao

8/15/2020

Homework 2: data visualization using ggplot2

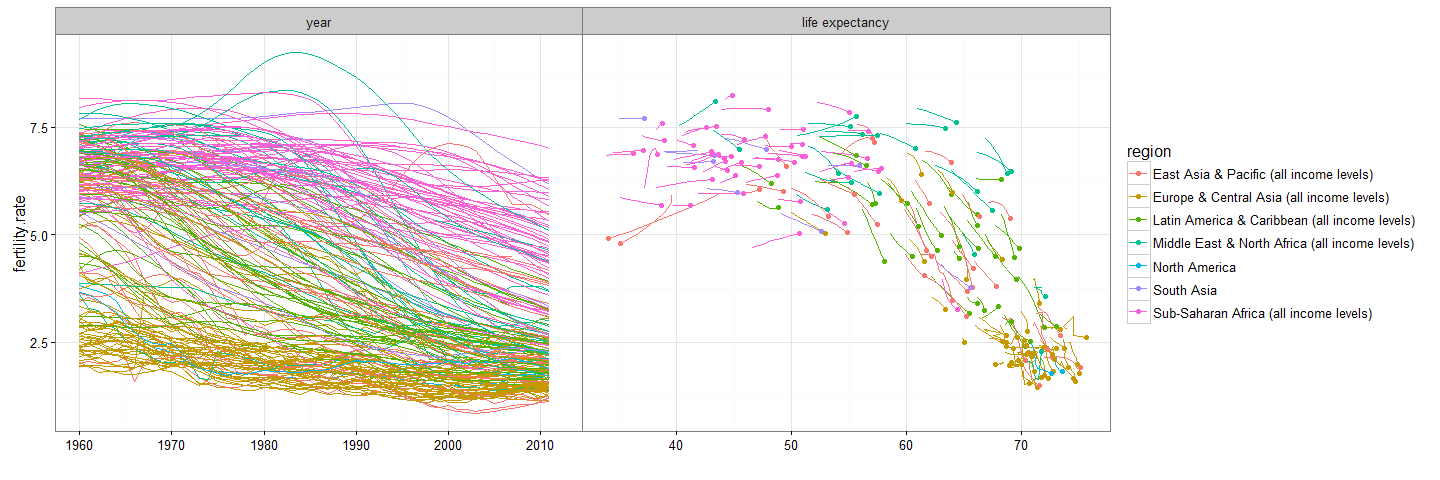
These are exercises from

<http://members.cbio.mines-paristech.fr/~thocking/animint2-manual/Ch02-ggplot2.html>

(the variable name viz.aligned below refers to ggplot code on that web page).

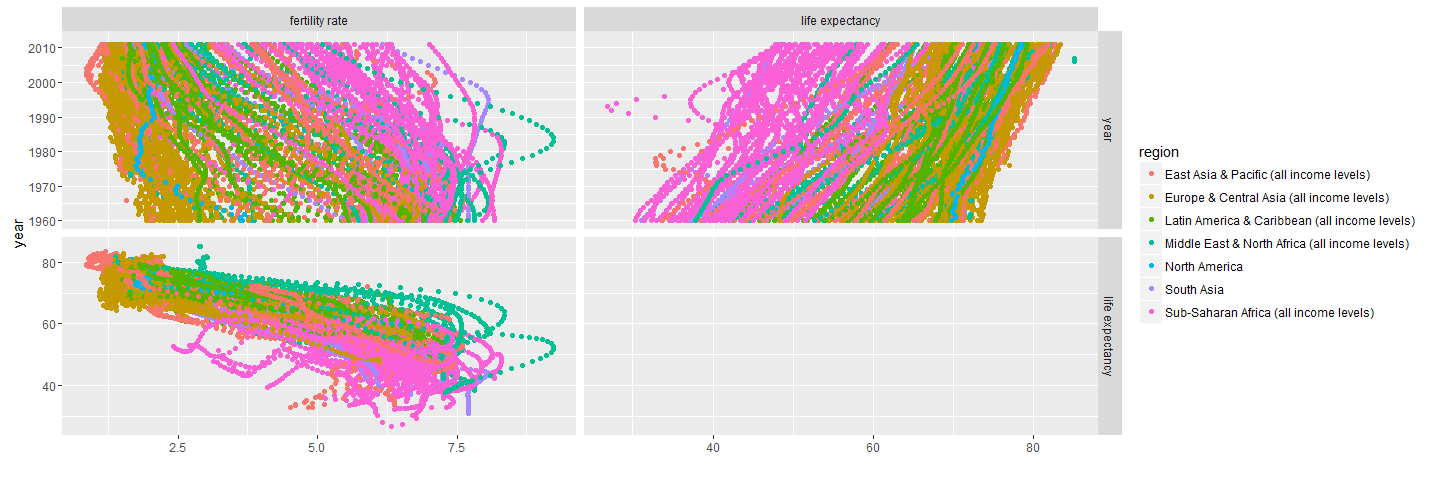
#if we didn't in stll remotes package, we install it  
if(FALSE){  
 install.packages("remotes")  
}  
  
#if we didn't in stll animint2 package, we install it  
if(FALSE){  
 remotes::install\_github("tdhock/animint2")   
}  
  
library(animint2)

1.In viz.aligned we showed a ggplot with a scatterplot panel on the left and a time series panel on the right. Make another version of the data visualization with the time series panel on the left and the scatterplot panel on the right. Hint: the order of panels in ggplots is controlled by the factor levels of the facet variable. For example facet\_grid(. ~ x) means use the different values of the x variable in different facets, where x = factor(values, levels).

add.x.var <- function(df, x.var){  
 data.frame(df, x.var=factor(x.var, c("year", "life expectancy")))  
}  
  
  
viz.aligned <- function(df1,df2,df){  
 viz.aligned <- ggplot()+  
 geom\_path(aes(x=life.expectancy,y=fertility.rate,color=region,  
 group=country),  
 data=add.x.var(df1, "life expectancy"))+  
 geom\_point(aes(x=life.expectancy,y=fertility.rate,color=region),  
 data=add.x.var(df2, "life expectancy"))+  
  
 geom\_line(aes(x=year,y=fertility.rate, color=region, group=country),  
 data=add.x.var(df, "year"))+  
 xlab("")+  
 facet\_grid(. ~ x.var, scales="free")+  
 theme\_bw()+  
 theme(panel.margin=grid::unit(0, "lines"))  
   
  
 return(viz.aligned)  
}  
  
  
viz.aligned(subset(WorldBank, 1970 <= year & year <= 1975),subset(WorldBank, year==1975),WorldBank)

2.In viz. aligned the scatterplot displays fertility rate and life expectancy, but the time series displays only fertility rate. Make another version of the data visualization that shows both time series. There should be only one ggplot that looks something like this. Hint: add another geom\_point and use both horizontal and vertical panels in facet\_grid(row\_var ~ col\_var).

add.vars <- function(df,x.var,y.var){  
 data.frame(df,x.var=factor(x.var, c("fertility rate", "life expectancy")), y.var=factor(y.var, c("year", "life expectancy")))  
}  
  
viz.aligned2 <- function(df){  
 result <- ggplot()+  
  
 geom\_point(aes(x=life.expectancy,y=year, color=region),  
 data=add.vars(df, "life expectancy","year"))+  
 geom\_point(aes(x=fertility.rate,y=year, color=region),  
 data=add.vars(df, "fertility rate","year"))+  
 geom\_point(aes(x=fertility.rate,y=life.expectancy, color=region),  
 data=add.vars(df, "fertility rate","life expectancy"))+  
 xlab("")+  
 facet\_grid(y.var~ x.var, scales="free")  
   
 return(result)  
}  
  
  
viz.aligned2(WorldBank)



3.Create a multi-panel data visualization (scatterplot of data for one year in each panel) that shows each year of the WorldBank data set in a separate panel of one ggplot. Hint: use facet\_wrap.

viz.panels <- function(df){  
 result <- ggplot(df,aes(life.expectancy,fertility.rate,color=region,group=country)) +  
 geom\_point() +  
 facet\_wrap(~year)  
   
   
 return(result)  
}  
  
  
viz.panels(WorldBank)

