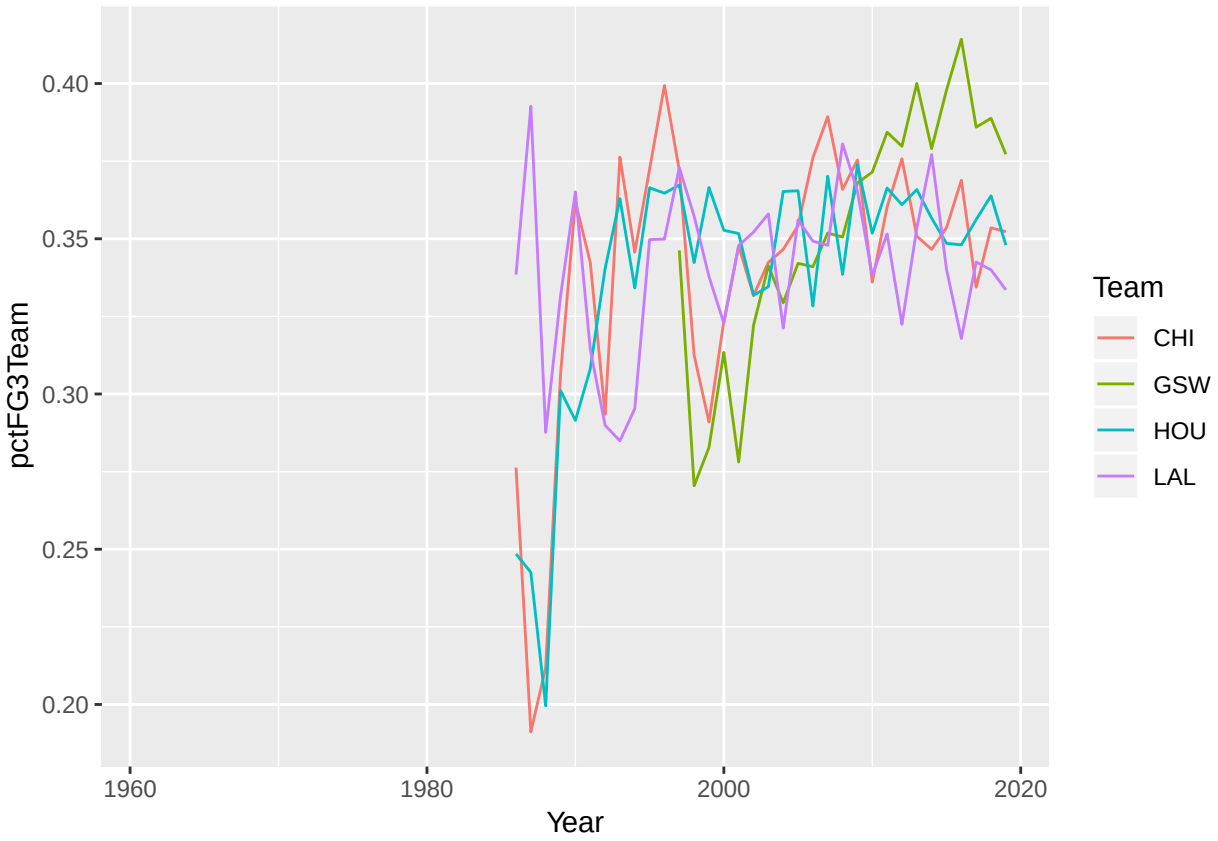
Problem Set 6

Subeom Lee

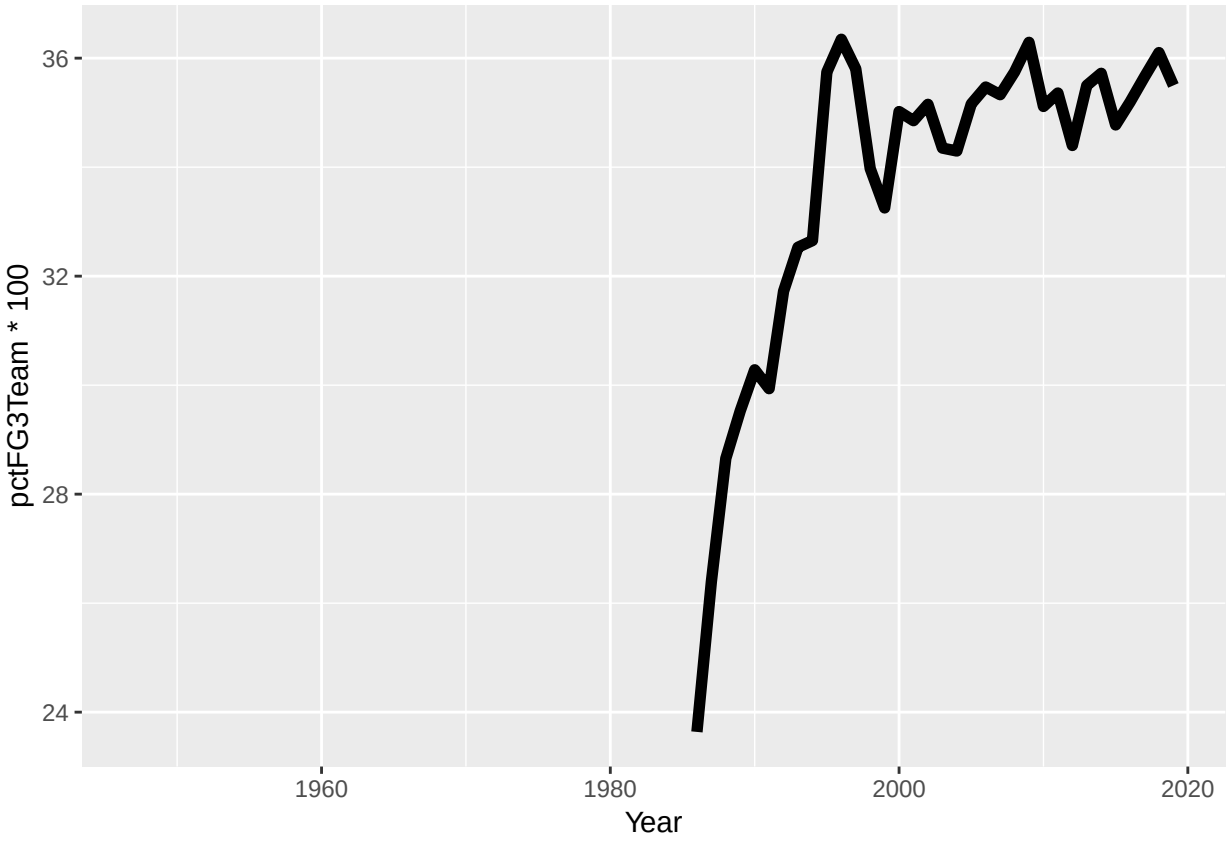
2019-03-07

1 Flight delay predictions

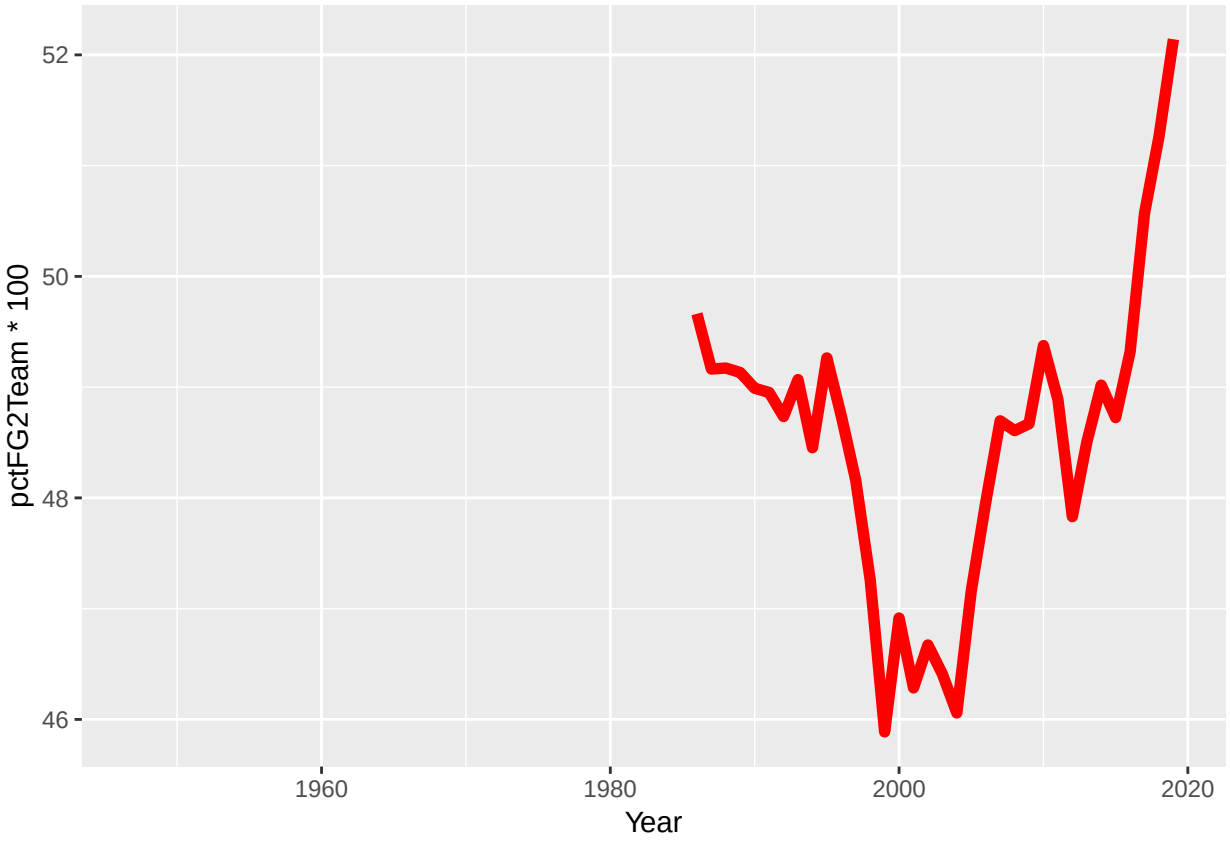
```
#http://asbcllc.com/nbastatR/index.html

library(nbastatR)
library(future)
library(stringi)
library(tidyverse)
library(lubridate)
library(texreg)
library(broom)
library(knitr)
library(ggpubr)
library(ggrepel)
library(janitor)
library(plotly)

plan(multiprocess)

# Run only when needed
# game_logs(seasons = 1947:2019, result_types = c("team", "player"))
# dataGameLogsTeam$Team = substring(dataGameLogsTeam$slugMatchup, 1, 3)

# Run when you updated data
# save(df_nba_player_dict, file='df_nba_player_dict.Rdata')
# save(dataGameLogsTeam, file='dataGameLogsTeam.Rdata')
# save(dataGameLogsPlayer, file='dataGameLogsPlayer.Rdata')

load('df_nba_player_dict.Rdata')
load('dataGameLogsTeam.Rdata')
load('dataGameLogsPlayer.Rdata')

avg <- aggregate(dataGameLogsTeam[, 24:46], list(dataGameLogsTeam$yearSeason, dataGameLogsTeam$Team), mean)
colnames(avg)[1] <- "Year"
colnames(avg)[2] <- "Team"



avgplot <- avg %>%
            filter(Team %in% c('GSW', 'CHI', 'HOU', 'LAL')) %>%
            ggplot(aes(x=Year, y=pctFG3Team, colour=Team)) +
            geom_line()
avgplot
```



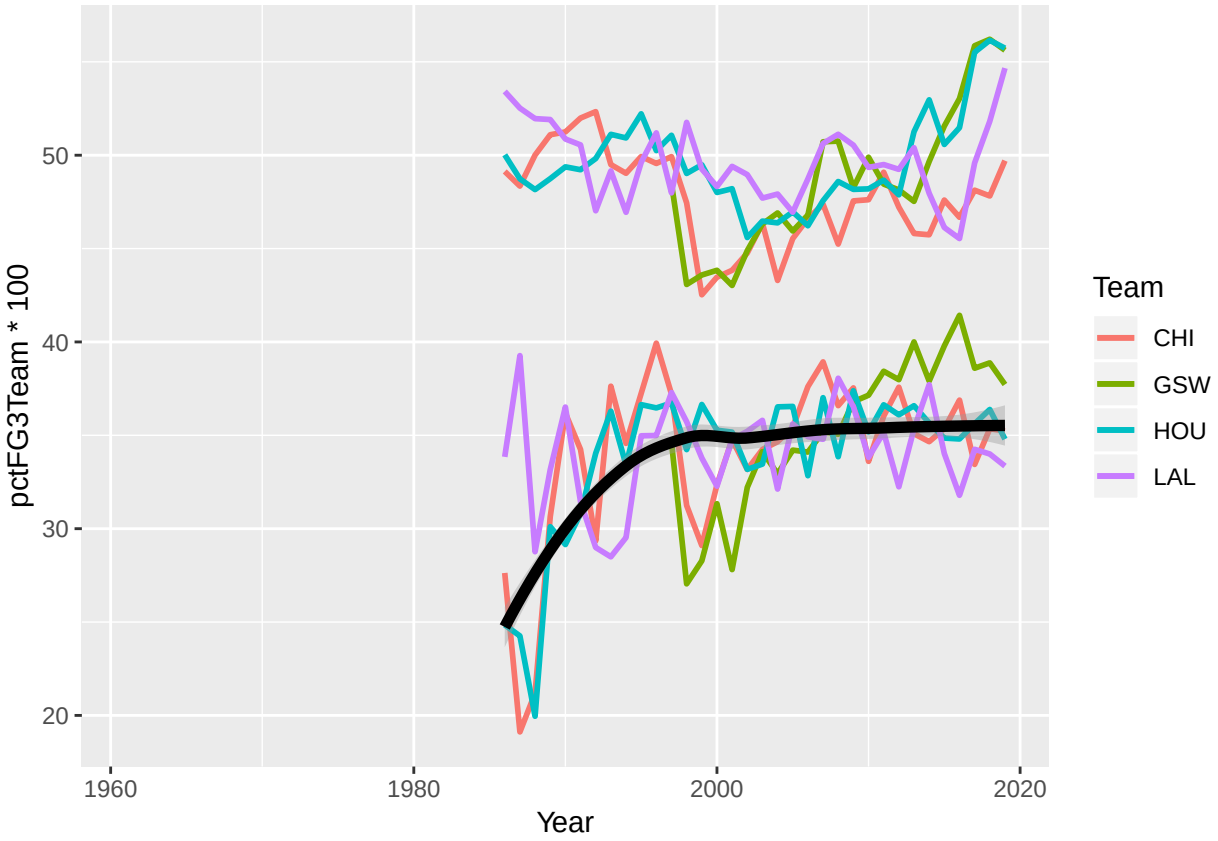

```
avg2 <- aggregate(dataGameLogsTeam[, 24:46], list(dataGameLogsTeam$yearSeason), mean)
colnames(avg2)[1] <- "Year"

avgplot2 <- avg2 %>%
              ggplot(aes(x=Year, y=pctFG3Team*100)) +
              geom_path(colour='black', size=2)
avgplot2
```



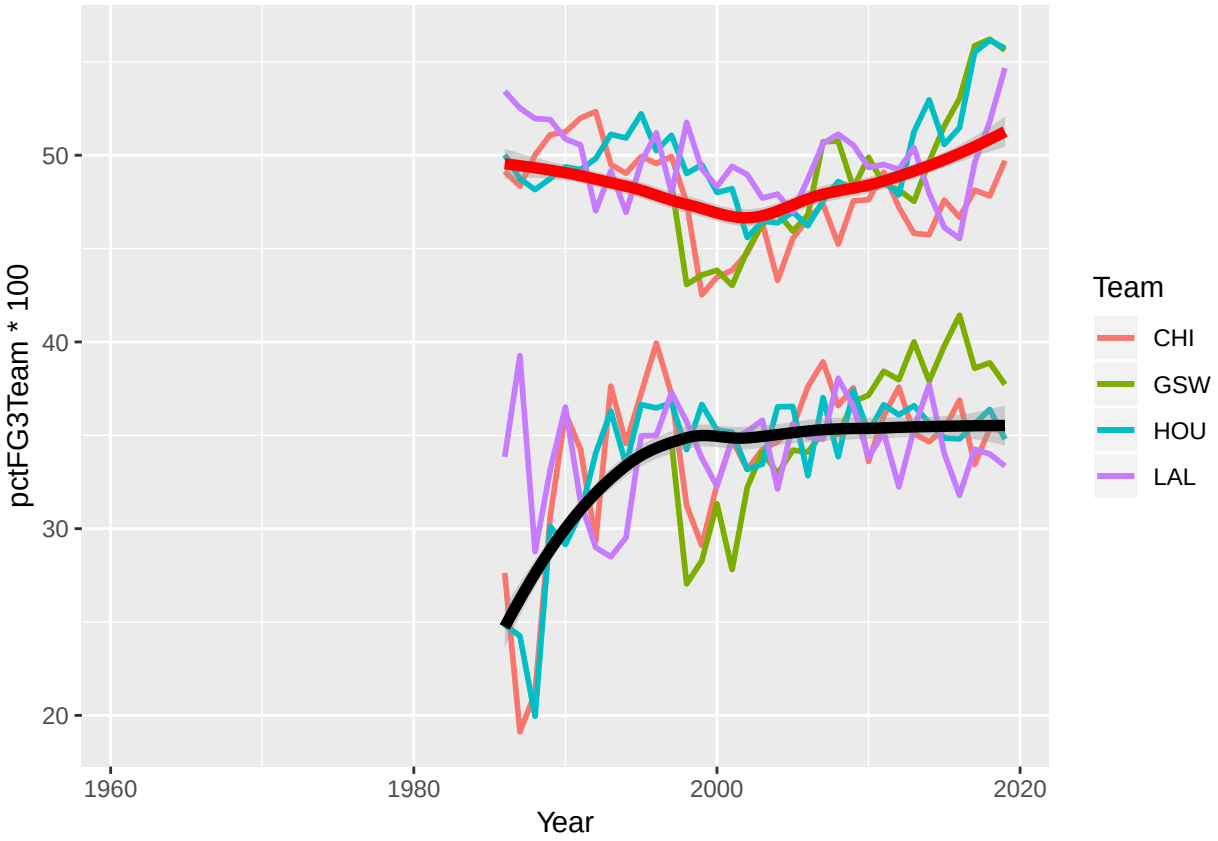

```
avgplot3 <- avg2 %>%
              ggplot(aes(x=Year, y=pctFG2Team*100)) +
              geom_path(colour='red', size=2)
avgplot3
```



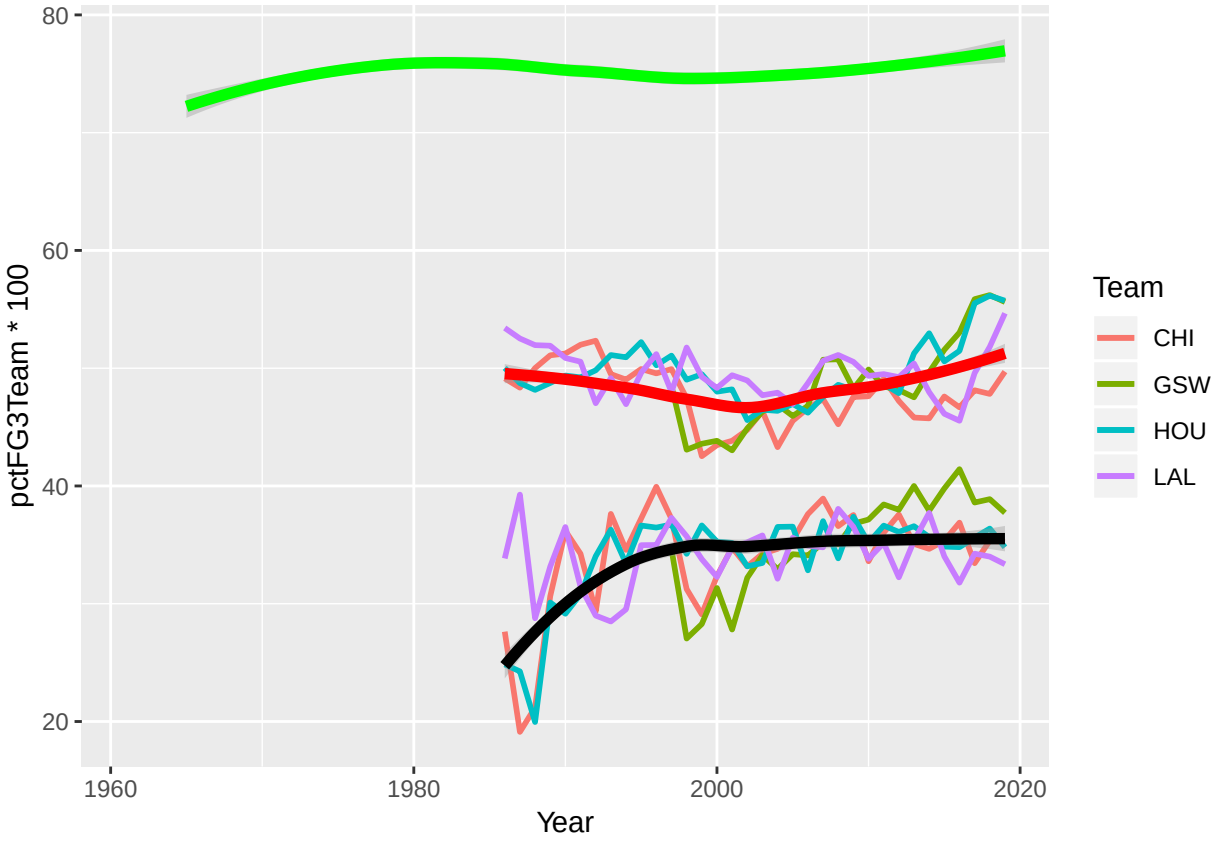

```
avgfiltered <- avg %>% filter(Team %in% c('GSW', 'CHI', 'HOU', 'LAL'))

avgcombined <- ggplot() +
                geom_line(data=avgfiltered, aes(x=Year, y=pctFG3Team*100, colour=Team), size=1) +
                geom_line(data=avgfiltered, aes(x=Year, y=pctFG2Team*100, colour=Team), size=1) +
                geom_smooth(data=avg2, aes(x=Year, y=pctFG3Team*100), size=2, colour='black')
avgcombined
```



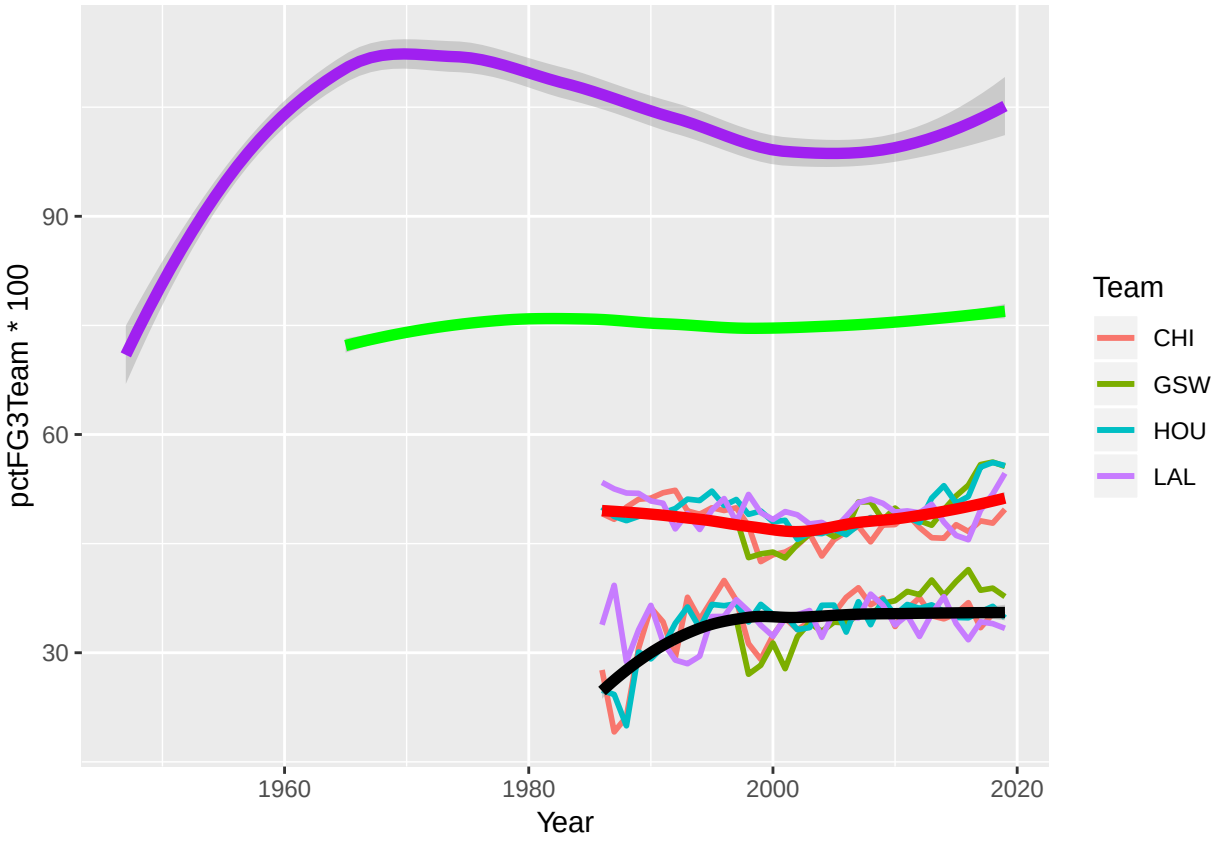

```
avgcombined2 <- avgcombined +
                geom_smooth(data=avg2, aes(x=Year, y=pctFG2Team*100), size=2, colour='red')
avgcombined2
```




```
avgcombined3 <- avgcombined2 +
                geom_smooth(data=avg2, aes(x=Year, y=pctFTTeam*100), size=2, colour='green')
avgcombined3
```




```
avgcombined4 <- avgcombined3 +
                geom_smooth(data=avg2, aes(x=Year, y=ptsTeam), size=2, colour='purple')
avgcombined4
```




```
# ggplotly(p=ggplot2::last_plot())

# library(ggplot2)
# library(ggpubr)
# theme_set(theme_pubr())
#
# figure <- ggarrange(avgplot, avgplot2,
#                     labels = c("Each Team", "All Teams"),
#                     ncol = 1, nrow = 2)
# figure

# climate <- read.csv('ps5_data.csv')
# a <- ggplot(climate) +
#       xlab('Year') +
#       ylab('Temperature(°C)') +
#       theme(panel.border=element_rect(colour="black", fill=NA), panel.background=element_rect(fill=NA),
#             panel.grid=element_line(color="grey")) +
#       geom_smooth(aes(Year, Lowess.5.), colour="blue", size=1) +
#       geom_line(aes(Year, No_Smoothing), colour="grey", size=1) +
#       geom_point(aes(Year, No_Smoothing), shape=1, size=3)
#
```