lm() in mutate()

Andriy Koval

Tuesday, June 24, 2014

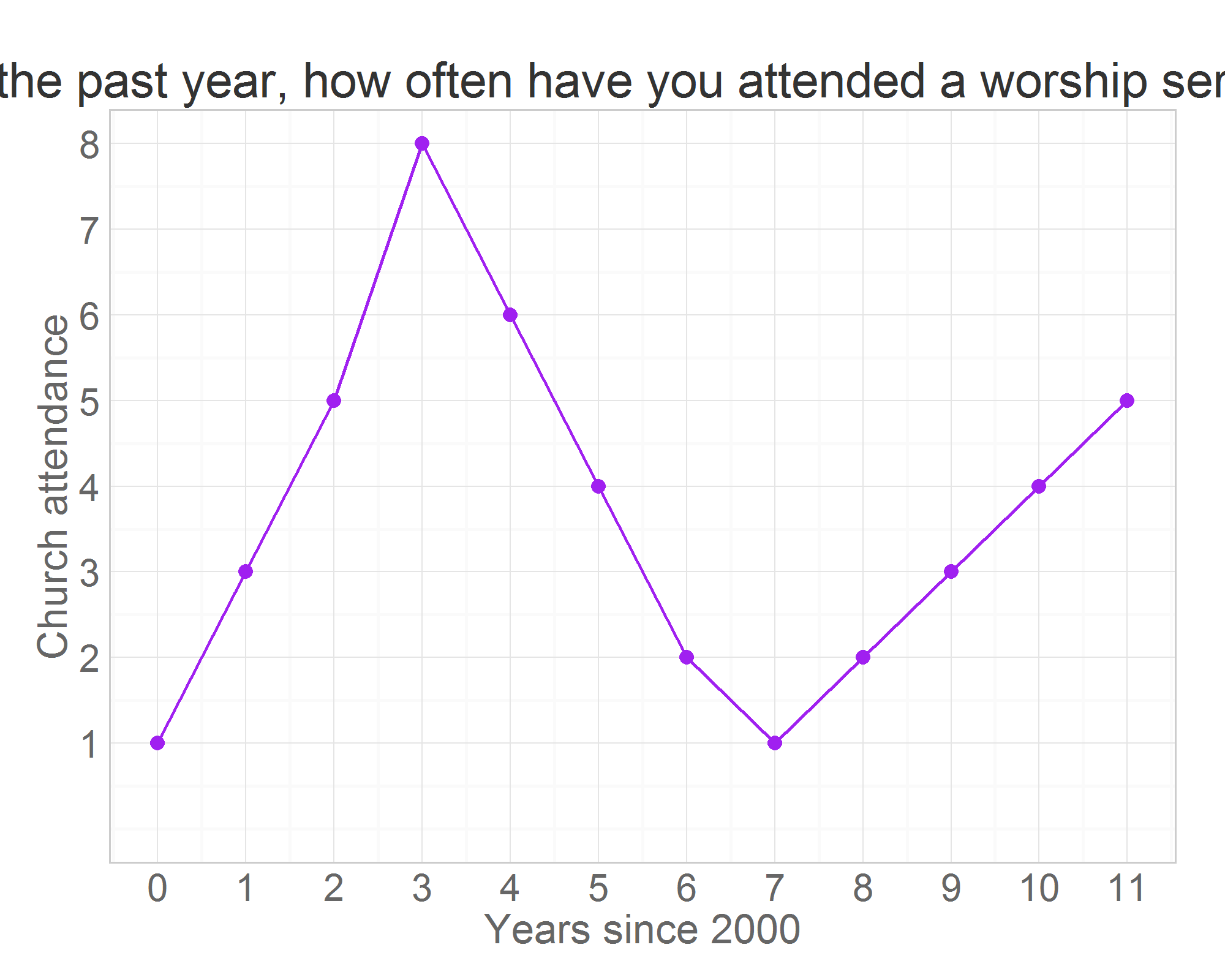
Data for a single person

ds<- dsL %>% dplyr::filter(id==1,year %in% c(2000:2011)) %>% dplyr::select(id,year,attend) %>%  
 mutate(time=year-2000)  
print(ds)

id year attend time  
1 1 2000 1 0  
2 1 2001 6 1  
3 1 2002 2 2  
4 1 2003 1 3  
5 1 2004 1 4  
6 1 2005 1 5  
7 1 2006 1 6  
8 1 2007 1 7  
9 1 2008 1 8  
10 1 2009 1 9  
11 1 2010 1 10  
12 1 2011 1 11

ds <- read.table(header = TRUE, text ="  
obs id year attend  
1 47 2000 1  
2 47 2001 3  
3 47 2002 5  
4 47 2003 8  
5 47 2004 6  
6 47 2005 4  
7 47 2006 2  
8 47 2007 1  
9 47 2008 2  
10 47 2009 3  
11 47 2010 4  
12 47 2011 5  
  
")  
ds<- ds %>% dplyr::mutate(time=year-2000)  
print(ds)

obs id year attend time  
1 1 47 2000 1 0  
2 2 47 2001 3 1  
3 3 47 2002 5 2  
4 4 47 2003 8 3  
5 5 47 2004 6 4  
6 6 47 2005 4 5  
7 7 47 2006 2 6  
8 8 47 2007 1 7  
9 9 47 2008 2 8  
10 10 47 2009 3 9  
11 11 47 2010 4 10  
12 12 47 2011 5 11

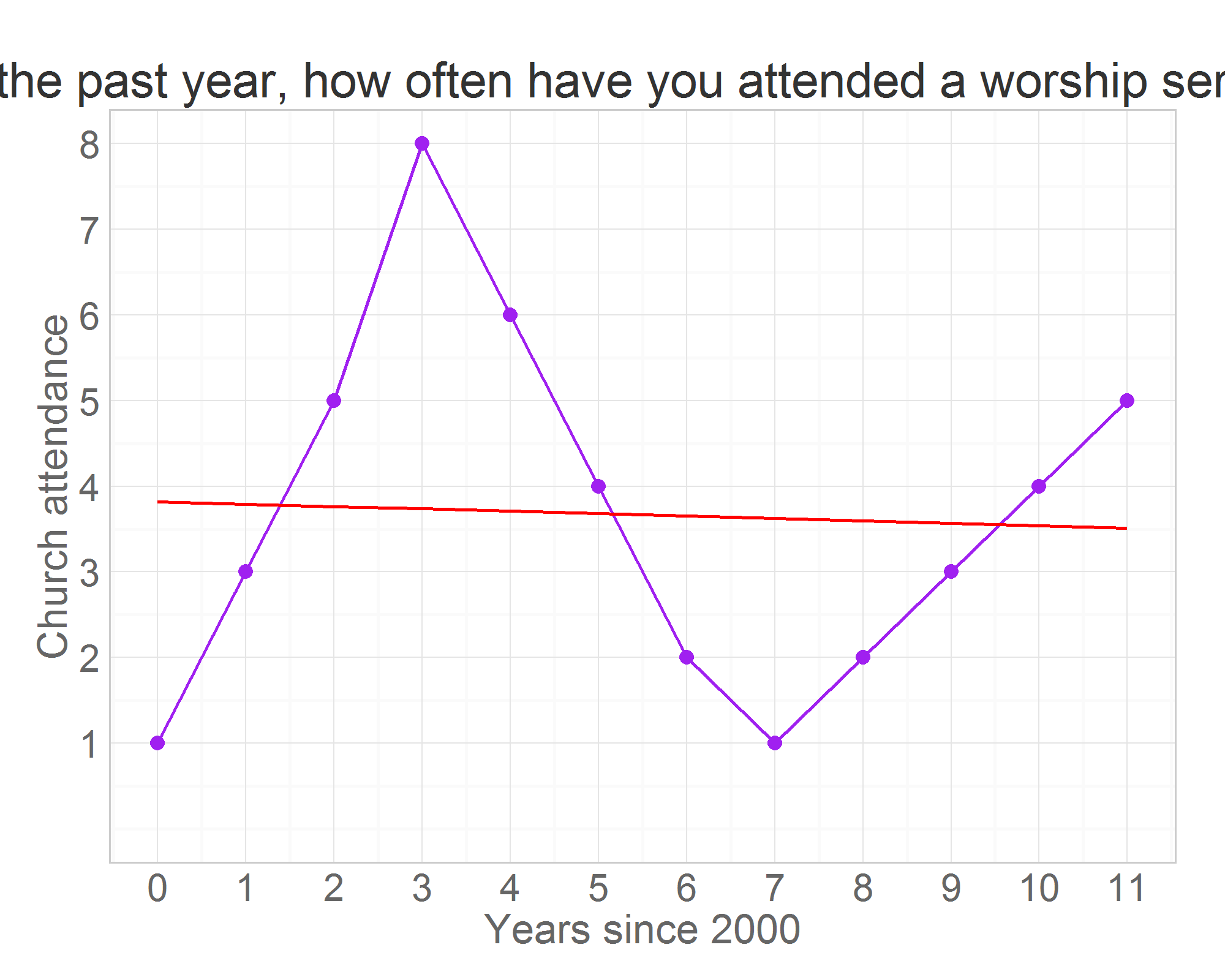


add a straight line to represent possible predition line, in this case a straight line

linear<- predict(lm(attend ~ time, ds))  
ds<- ds %>% dplyr::mutate(linear=linear)  
print(ds)

obs id year attend time linear  
1 1 47 2000 1 0 3.821  
2 2 47 2001 3 1 3.793  
3 3 47 2002 5 2 3.765  
4 4 47 2003 8 3 3.737  
5 5 47 2004 6 4 3.709  
6 6 47 2005 4 5 3.681  
7 7 47 2006 2 6 3.653  
8 8 47 2007 1 7 3.625  
9 9 47 2008 2 8 3.597  
10 10 47 2009 3 9 3.569  
11 11 47 2010 4 10 3.541  
12 12 47 2011 5 11 3.513

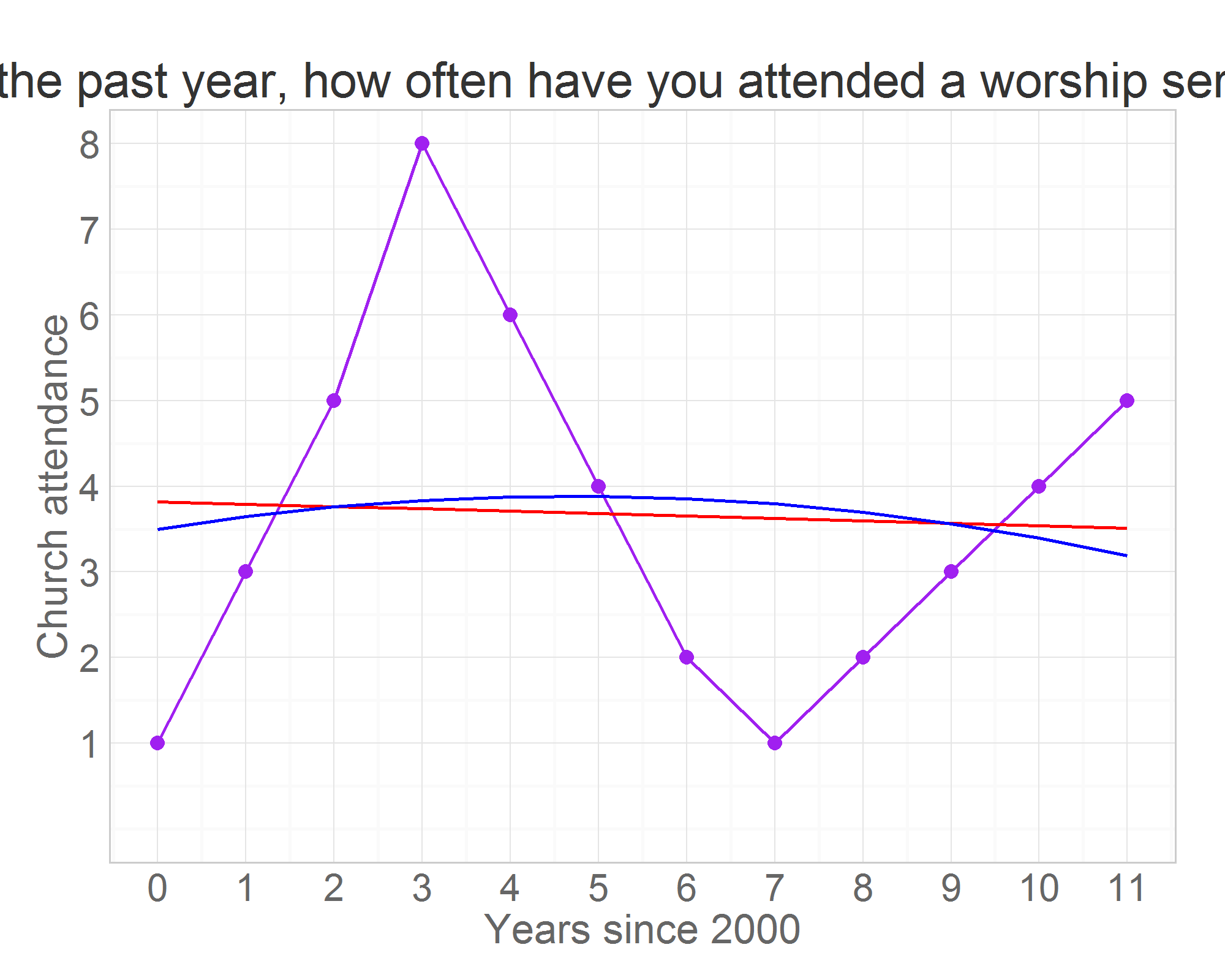
p<-p+ geom\_line(aes(y=linear),color="red", size=.5)  
p

 Or adding the curvarture the quadratic term

quadratic<- predict(lm(attend ~ poly(time,2),ds))  
ds<- ds %>% mutate(quadratic=quadratic)  
print(ds)

obs id year attend time linear quadratic  
1 1 47 2000 1 0 3.821 3.500  
2 2 47 2001 3 1 3.793 3.647  
3 3 47 2002 5 2 3.765 3.759  
4 4 47 2003 8 3 3.737 3.836  
5 5 47 2004 6 4 3.709 3.878  
6 6 47 2005 4 5 3.681 3.885  
7 7 47 2006 2 6 3.653 3.857  
8 8 47 2007 1 7 3.625 3.794  
9 9 47 2008 2 8 3.597 3.696  
10 10 47 2009 3 9 3.569 3.563  
11 11 47 2010 4 10 3.541 3.395  
12 12 47 2011 5 11 3.513 3.192

p<-p+ geom\_line(aes(y=quadratic),color="blue", size=.5)  
 p



# p<-p+ geom\_line(aes(y=cubic),color="green", size=.5)

or the cubic term

cubic<- predict(lm(attend ~ poly(time,3),ds))  
ds<- ds %>% mutate( cubic=cubic)  
print(ds)

obs id year attend time linear quadratic cubic  
1 1 47 2000 1 0 3.821 3.500 0.7436  
2 2 47 2001 3 1 3.793 3.647 3.8974  
3 3 47 2002 5 2 3.765 3.759 5.5128  
4 4 47 2003 8 3 3.737 3.836 5.9239  
5 5 47 2004 6 4 3.709 3.878 5.4646  
6 6 47 2005 4 5 3.681 3.885 4.4693  
7 7 47 2006 2 6 3.653 3.857 3.2720  
8 8 47 2007 1 7 3.625 3.794 2.2067  
9 9 47 2008 2 8 3.597 3.696 1.6076  
10 10 47 2009 3 9 3.569 3.563 1.8089  
11 11 47 2010 4 10 3.541 3.395 3.1445  
12 12 47 2011 5 11 3.513 3.192 5.9487

p<-p+ geom\_line(aes(y=cubic),color="green", size=.5)  
p

