

Problem Set 5

Sumeet Mishra

2/26/2019

1. Fit a model that estimates average BMI as a function of height and age. Note that a purely linear model is unlikely to be adequate. Write down the line of code you used to fit your model.

```
library(ggplot2)
library(NHANES)
library(tidyverse)

## -- Attaching packages ----- tidyverse 1.2.1 --
## v tibble 2.0.1      v purrr 0.2.5
## v tidyr 0.8.2       v dplyr 0.7.8
## v readr 1.3.1       v stringr 1.3.1
## v tibble 2.0.1      v forcats 0.3.0

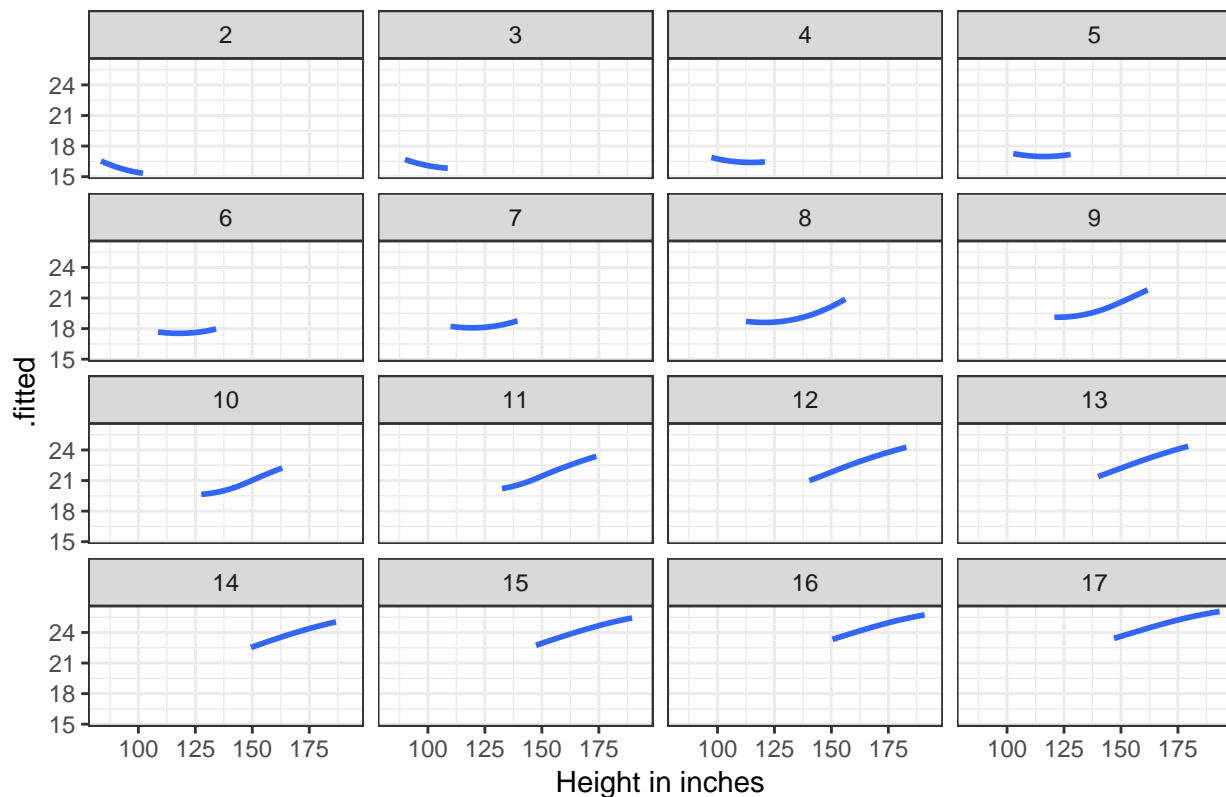
## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()     masks stats::lag()

library(broom)
nhanes<- (NHANES)
myvars1<-c('BMI', 'Age', 'Height')
nhanes<-nhanes[myvars1]
nhanes=nhanes[!(!complete.cases(nhanes)), ]
nhanes.lo = loess(BMI ~ Age*Height, data = nhanes, span = 1, normalize=TRUE)
nhanes.plot.df=augment(nhanes.lo)
nhanes.plot.df=nhanes.plot.df%>%filter(Age<=17)
```

2. Draw a graph faceted by age that shows how, according to your model, average BMI varies with height.

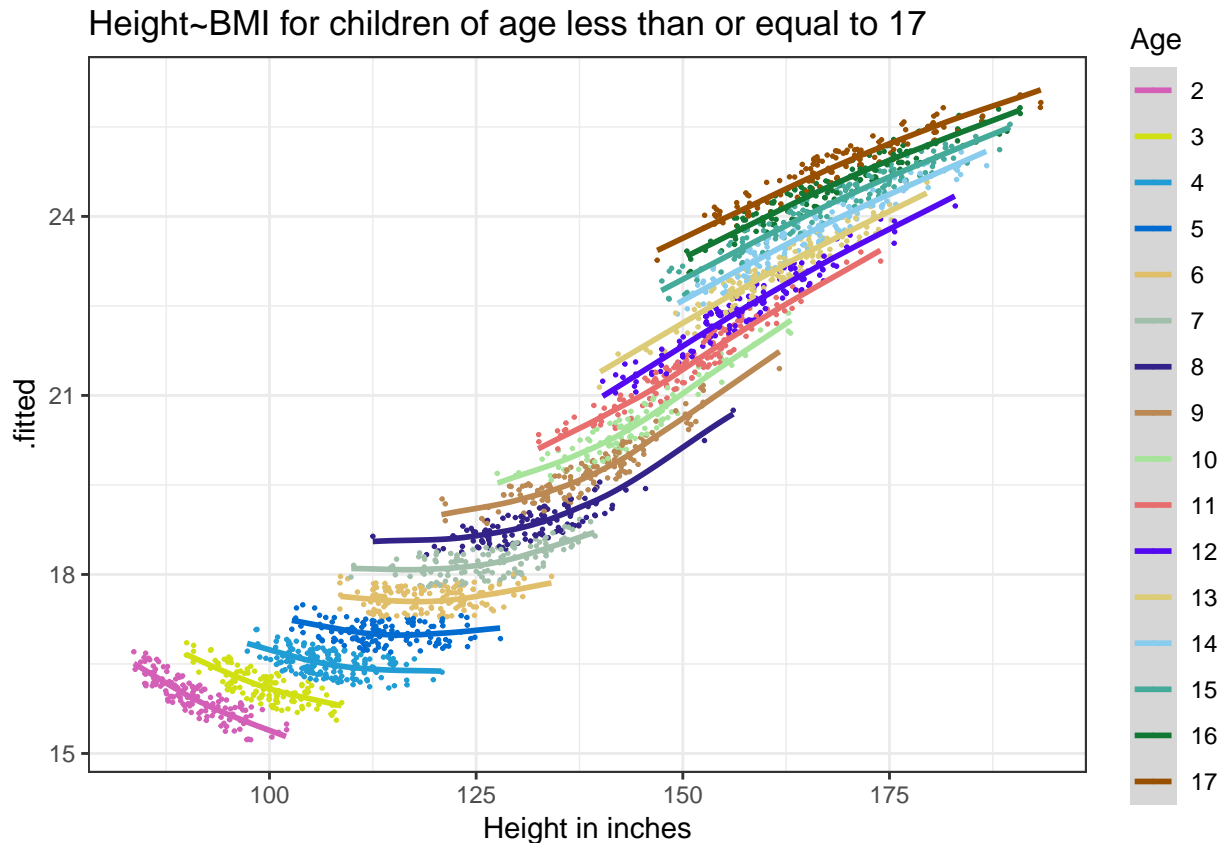
```
ggplot(nhanes.plot.df, aes(x=Height, y=.fitted)) + geom_smooth(method='loess') +
  facet_wrap(~Age, ncol = 4) + theme_bw() +
  ggtitle('Height~BMI faceted by Age of children less than or equal to 17') +
  xlab('Height in inches')
```

Height~BMI faceted by Age of children less than or equal to 17



3. Draw a graph which uses color to denote age that shows how, according to your model, average BMI varies with height. Your color-scheme must be legible and reasonably color- blind friendly.

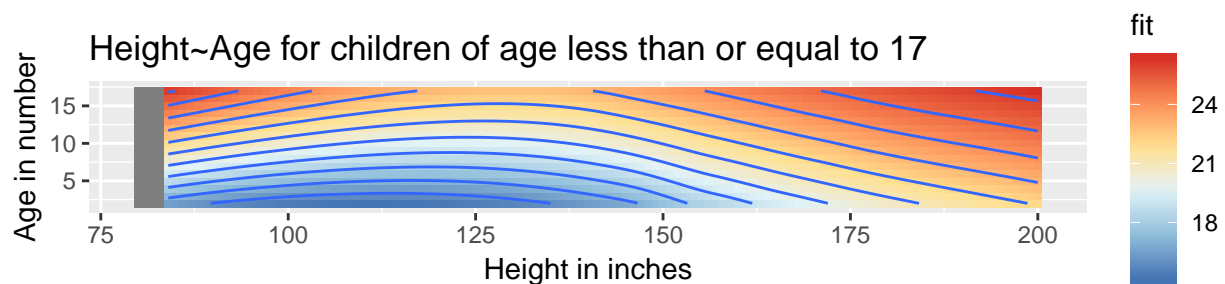
```
#nhanes1.lo = loess( avg_bmi~ Age+Height, data = nhanes1)
#nhanes1.lo.df = augment(nhanes1.lo)
palette=c('#D35FB7','#D1E013','#1F9DD4','#006CD1','#E1BE6A','#A1BFAB',
          '#332288','#BB8954','#A7E49B','#E86E6E','#5309F1','#DDCC77',
          '#88CCEE','#44AA99','#117733','#994F00','#E66100')
ggplot(nhanes.plot.df, aes(x =Height , y = .fitted,color=factor(Age)))+
  scale_color_manual(values = palette) + geom_jitter(height = 0.3, width = 0.1, size = 0.3)+
  geom_smooth(method='loess',method.args=list(degree=1))+
  theme_bw()+labs(color='Age') +
  ggtitle('Height~BMI for children of age less than or equal to 17')+xlab('Height in inches')
```



4. Draw a graph that uses contours to show how average BMI varies with both height and age.

```
nhanes.grid = expand.grid(Height=seq(80,200,1),Age=seq(2,17,1))
nhanes.predict = predict(nhanes.lo, newdata=nhanes.grid)
nhanes.plot1.df = data.frame(nhanes.grid, fit=as.vector(nhanes.predict))
ggplot(nhanes.plot1.df, aes(x=Height,y=Age,z=fit)) + geom_line() +
  geom_raster(aes(fill = fit))+
  coord_fixed() + scale_fill_distiller(palette="RdYlBu") +
  geom_contour()+ggtitle('Height~Age for children of age less than or equal to 17')+
  xlab('Height in inches')+ylab('Age in number')
```

Warning: Removed 64 rows containing non-finite values (stat_contour).



5. How does average BMI vary with height and age for children 17 and under? Is BMI a good measure of body mass for children?

We observed from the plots that-

1. The average BMI and height increases with increase in age. So, young children tends to have less BMI and height than children elder to them.

2. From the raster and contour plot, we noticed that when age is more than 15, height and BMI tends to be high.

3. BMI is not always a good measure of body mass of children as it can not distinguish a obese child from a well built child(child with muscles). There might be other variables for children to be taller or obese which BMI may not explain.