

Exp10-Spline Curve Models

Register Number:19BLC1186

Name:Tarun Sidhu

Lab Exercise No:9

Date:17/5/2021

Dataset: Wage

Task:Spline Curve for given Dataset

```
library(tidyverse)
```

```
## — Attaching packages — tidyverse 1.3.0 —
```

```
## ✓ ggplot2 3.3.3      ✓ purrr 0.3.4
## ✓ tibble 3.1.0       ✓ dplyr 1.0.5
## ✓ tidyr 1.1.3        ✓ stringr 1.4.0
## ✓ readr 1.4.0        ✓ forcats 0.5.1
```

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

```
library(broom)
require(splines)
```

```
## Loading required package: splines
```

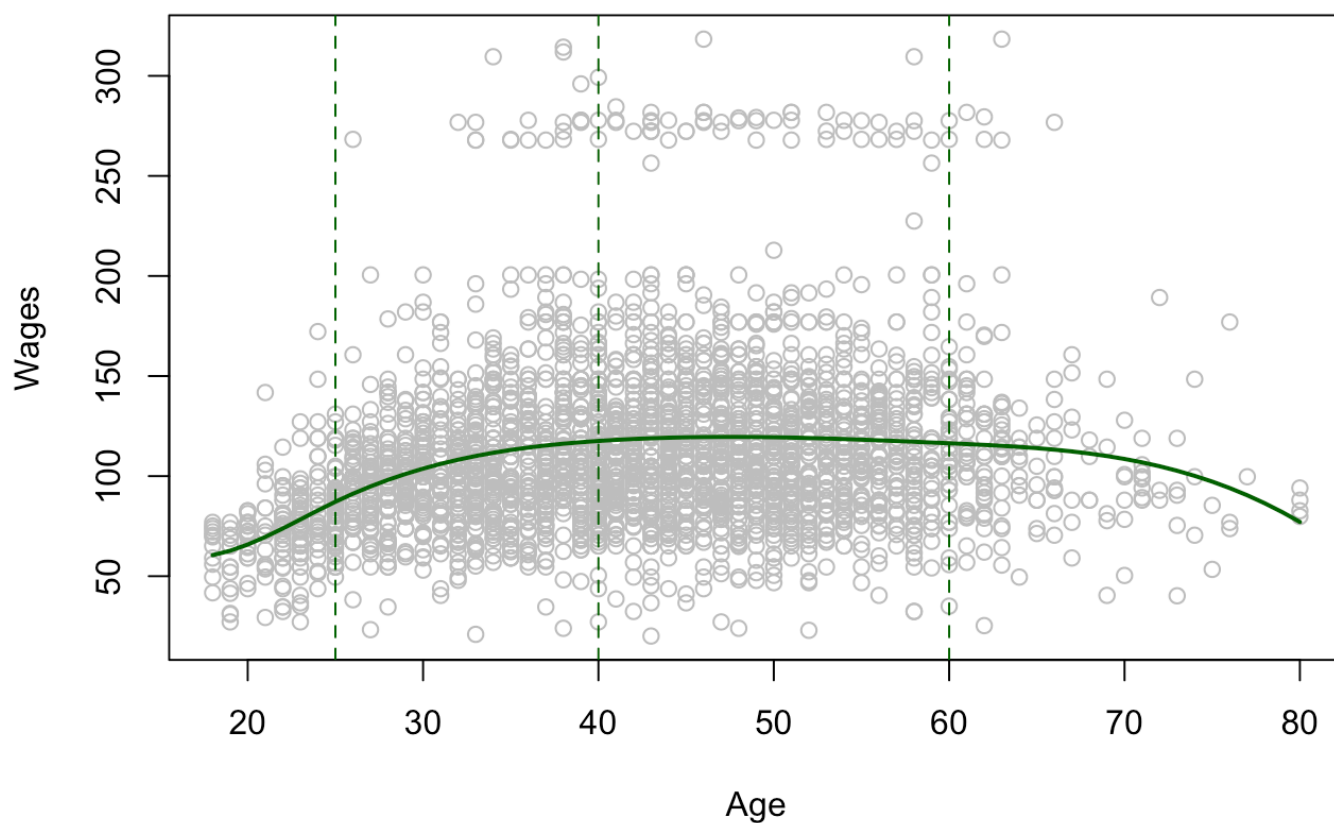
```
library(ISLR)
attach(Wage)
agelims<-range(Wage$age)
age.grid<-seq(from=agelims[1], to = agelims[2])

fit<-lm(wage ~ bs(age,knots = c(25,40,60)),data = Wage )
summary(fit)
```

```
##
## Call:
## lm(formula = wage ~ bs(age, knots = c(25, 40, 60)), data = Wage)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -98.832 -24.537  -5.049   15.209  203.207
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      60.494      9.460   6.394 1.86e-10 ***
## bs(age, knots = c(25, 40, 60))1    3.980     12.538   0.317 0.750899
## bs(age, knots = c(25, 40, 60))2   44.631      9.626   4.636 3.70e-06 ***
## bs(age, knots = c(25, 40, 60))3   62.839     10.755   5.843 5.69e-09 ***
## bs(age, knots = c(25, 40, 60))4   55.991     10.706   5.230 1.81e-07 ***
## bs(age, knots = c(25, 40, 60))5   50.688     14.402   3.520 0.000439 ***
## bs(age, knots = c(25, 40, 60))6   16.606     19.126   0.868 0.385338
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 39.92 on 2993 degrees of freedom
## Multiple R-squared:  0.08642,    Adjusted R-squared:  0.08459
## F-statistic: 47.19 on 6 and 2993 DF,  p-value: < 2.2e-16
```

```
plot(Wage$age,Wage$wage,col="grey",xlab="Age",ylab="Wages")
points(age.grid,predict(fit,newdata = list(age=age.grid)),col="darkgreen",lwd=2,ty
pe="l")

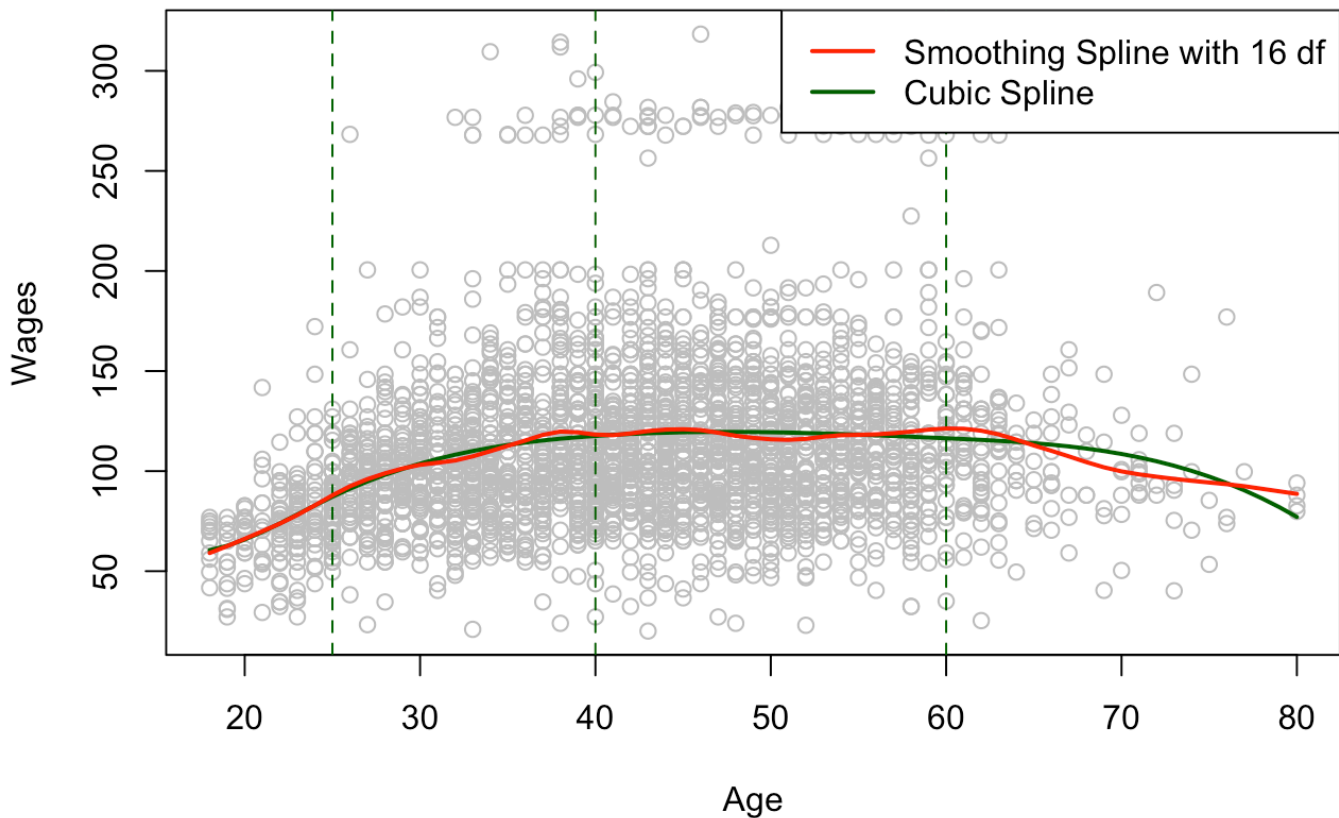
abline(v=c(25,40,60),lty=2,col="darkgreen")
```



```
fit1<-smooth.spline(Wage$age,Wage$wage,df=16) #16 degrees of freedom

plot(Wage$age,Wage$wage,col="grey",xlab="Age",ylab="Wages")
points(age.grid,predict(fit,newdata = list(age=age.grid)),col="darkgreen",lwd=2,ty
pe="l")

abline(v=c(25,40,60),lty=2,col="darkgreen")
lines(fit1,col="red",lwd=2)
legend("topright",c("Smoothing Spline with 16 df","Cubic Spline"),col=c("red","dar
kgreen"),lwd=2)
```



```
fit2<-smooth.spline(Wage$age,Wage$wage,cv = TRUE)
```

```
## Warning in smooth.spline(Wage$age, Wage$wage, cv = TRUE): cross-validation with
## non-unique 'x' values seems doubtful
```

```
fit2
```

```
## Call:
## smooth.spline(x = Wage$age, y = Wage$wage, cv = TRUE)
##
## Smoothing Parameter spar= 0.6988943 lambda= 0.02792303 (12 iterations)
## Equivalent Degrees of Freedom (Df): 6.794596
## Penalized Criterion (RSS): 75215.9
## PRESS(1.o.o. CV): 1593.383
```

```
plot(Wage$age,Wage$wage,col="grey",xlab="Age",ylab="Wages")  
  
lines(fit2,lwd=2,col="purple")  
legend("topright",("Smoothing Splines with 6.78 df selected by CV"),col="purple",lwd=2)
```

