Homework 2

Yiying Wu (yw3996)

R packages

```
library(tidyverse)
library(caret)
library(tidymodels)
```

Input dataset

```
dat<-read_csv("./data/College.csv")
dat <- na.omit(dat)
```

Partition the dataset into two parts: training data (80%) and test data (20%).

(a)

Fit smoothing spline models to predict out-of-state tuition (Outstate) using the percentage of alumni who donate (perc.alumni) as the only predictor, across a range of degrees of freedom. Plot the model fits for each degree of freedom. Describe the observed patterns that emerge with varying degrees of freedom. Select an appropriate degree of freedom for the model and plot this optimal fit. Explain the criteria you used to determine the best choice of degree of freedom.