**5 Exercises**

**Application**

Refer to the Air Quality data described previously, and the analyses we have done with Ozone as the response variable, and the five explanatory variables (including the two engineered features).

1. Use cubic splines to model the relationship between Ozone and Temp:

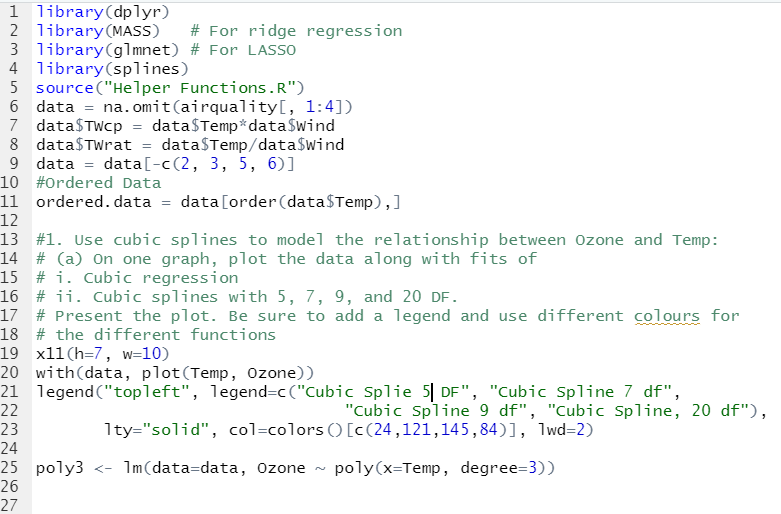
(a) On one graph, plot the data along with fits of

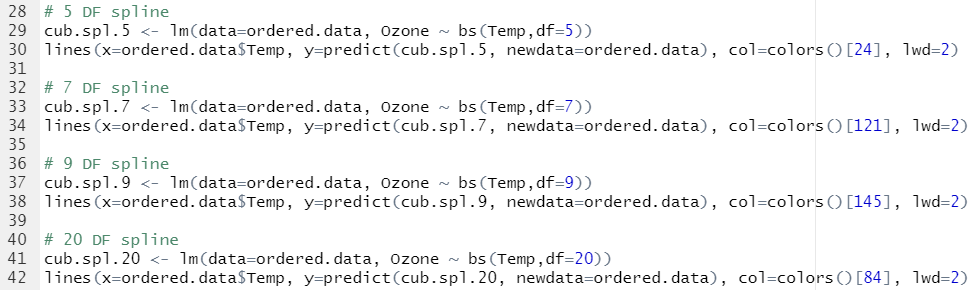
i. Cubic regression

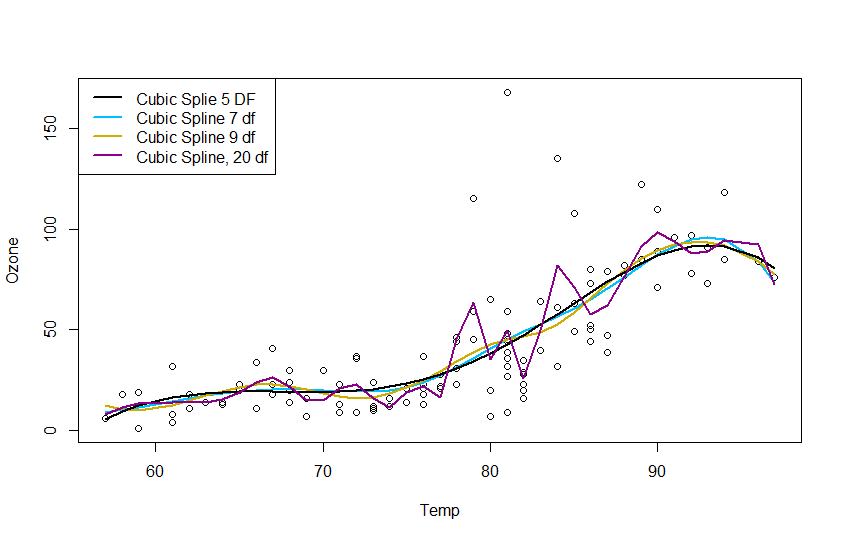
ii. Cubic splines with 5, 7, 9, and 20 DF.

**Present the plot. Be sure to add a legend and use different colours for**

**the different functions**





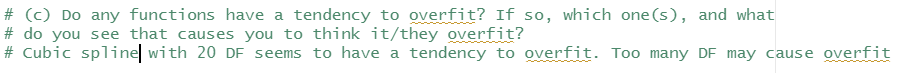


b) Which model seems to have the most bias? (**Just report the name**)



(c) Do any functions have a tendency to overfit? If so, **which one(s), and what**

**do you see that causes you to think it/they overfit?**



(d) If you had to choose one model, **which would it be? Why?**

