Assignment 4

Apurva Gupta,Shailendra Patil,Surbhi Paithankar

2/18/2018

1

log10(Budget) vs Year

In our analysis, to find the relationship between budget and year, we plotted the model fit using both "loess" and "linear" fit. Here we found that both the fits were almost similar. So in order to make a decision , we then plotted the residual plot for both the graphs and found that in both residual plots, points were scattered around zero. It was a matter of choice as both the models were similiar. We decided to go with a curved function for log10(Budget) vs year.

log10(Budget) vs Length

We plotted the model fit using both "loess" and "linear" fit. By the plot, loess seemed to be a good choice. In order to confirm that our assumption of loess is right, we plotted the residual plot for both "loess" and "linear" models. In case of "linear" fit, the residuals were not scattered along zero, and we also saw a trend in residual. But in case of "loess" model, the residuals were scattered around zero. Hence we decided to go with a curved function for log10(Budget) vs Length.

Interaction term

We plotted log10(budget) vs length over different years using cut_number. For different time range we see some nonmonotonic relationship. And hence we choose to have an interaction between lenght and year

Span

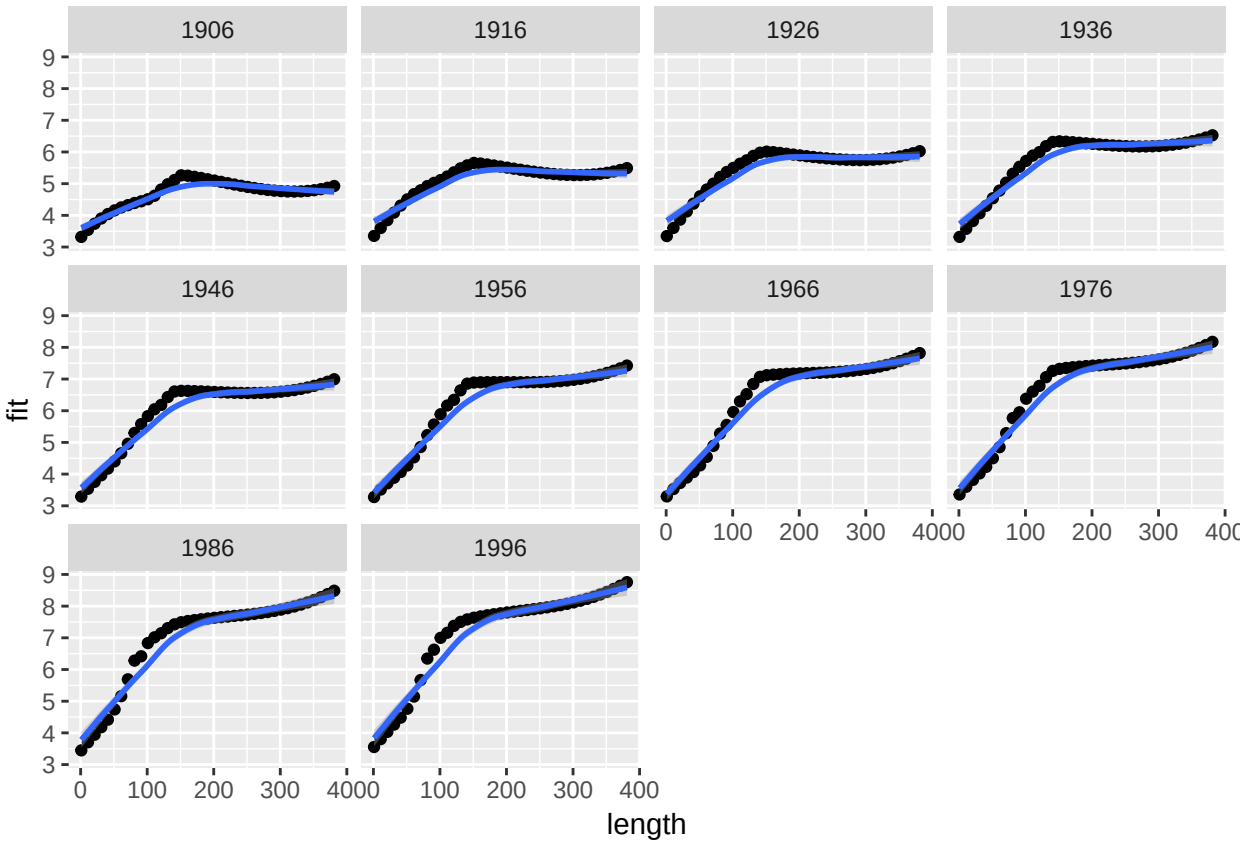
We plotted the graphs using different spans, and saw that there was not significant changes and the graphs were almost same for all spans. So in this case any choice of span would be a good. Hence we decided to go with the default span.

Least Square or Robust

As we saw there was nonmontonic relation, it would be better to fit the model using a robust fit

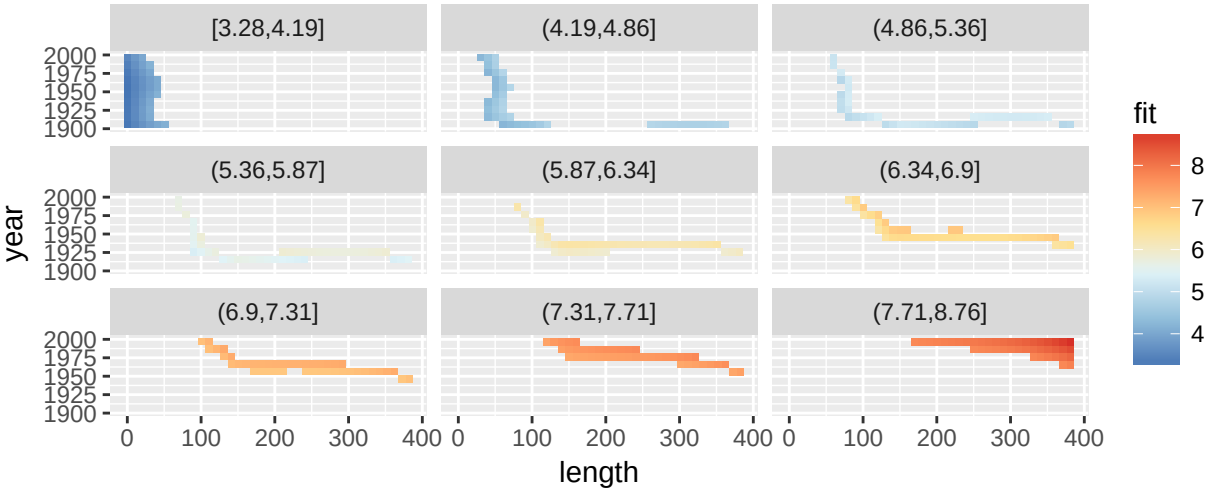
2

Displaying the fit

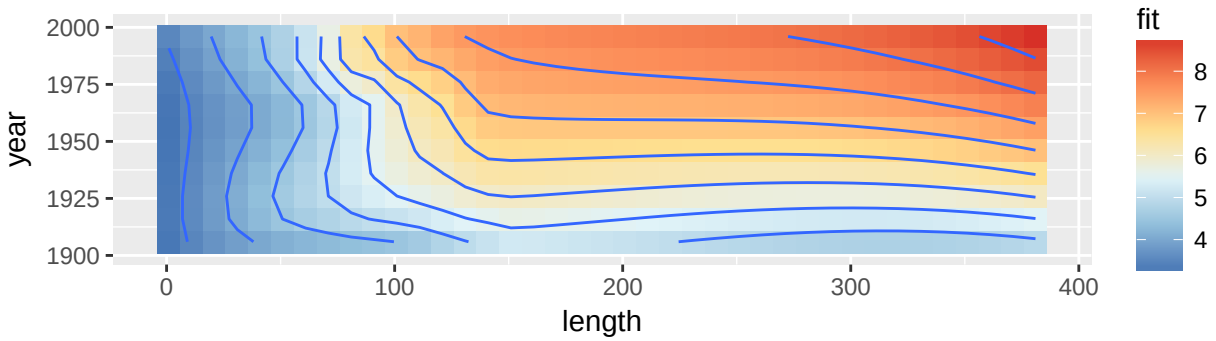


From the above fit , we can see that as length increases , the budget increases , but after a certain point i.e length greater than 150 , we see that length doesn't matter much. And also we see that for initial years before 1940 the magnitude of budget was less and we see that for later years the magnitude of budget increased with time.

3



The basic pattern is quite clear, for smaller lengths , log10(budget) is smaller and the magnitude is less for initital years as seen in the first row. In the second row we see that the curve is shifted to right compared to the right curve suggesting that as length increases the budget increases, and in third row it has further shifted and after a certain point the length the doesn't matter.



The contour plot shows the basic pattern, we can see for smaller lengths the budgets are lower and as length increase the budget also increases. And the magnitude for a given length looks greater for later years. After a certain length the increases in log10(budget) is less.