

```
In [2]: library(ggplot2)
library(MASS)
library(broom)
library(lattice)
library(TeachingDemos)
library(repr)
```

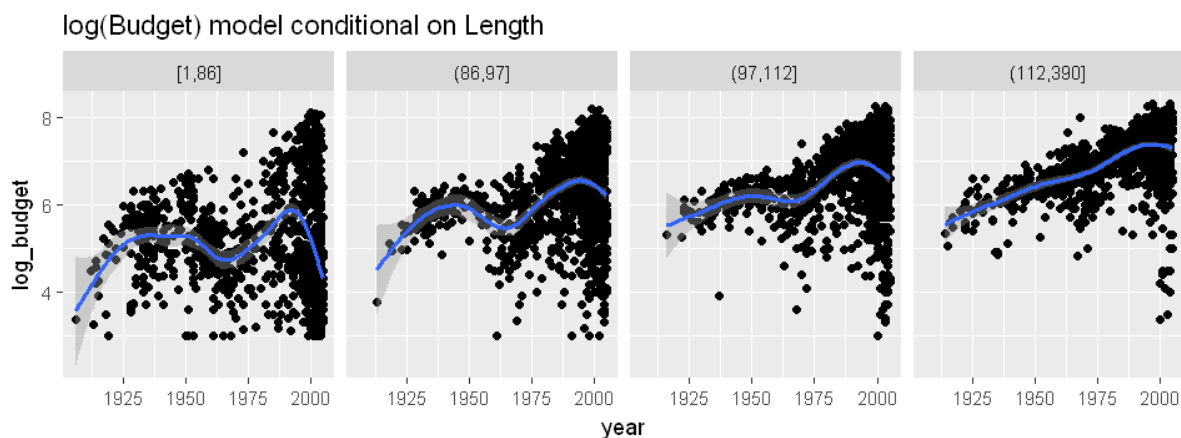
```
In [3]: movie_budgets = read.table(file.choose(),header = T)
```

```
In [4]: movie_budgets$log_budget = log10(movie_budgets$budget)
head(movie_budgets)
```

title	year	length	budget	log_budget
'G' Men	1935	85	450000	5.653213
'Manos' the Hands of Fate	1966	74	19000	4.278754
'Til There Was You	1997	113	23000000	7.361728
.com for Murder	2002	96	5000000	6.698970
10 Things I Hate About You	1999	97	16000000	7.204120
100 Mile Rule	2002	98	1100000	6.041393

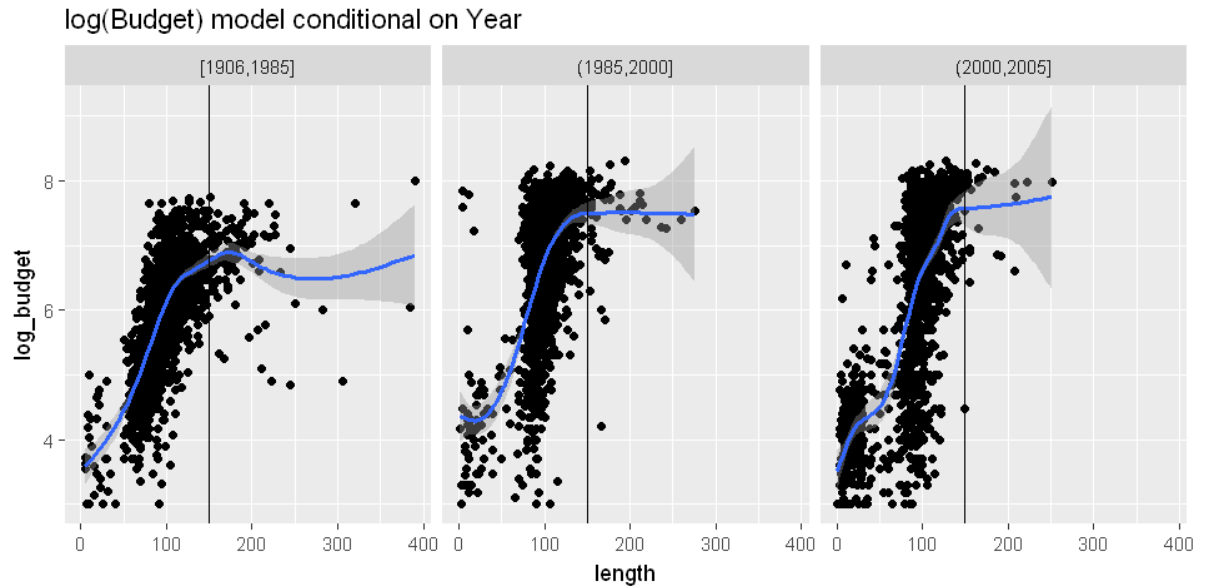
```
In [13]: options(repr.plot.width=8, repr.plot.height=3)
ggplot(movie_budgets, aes(y=log_budget, x= year)) + geom_point() + geom_smooth(
  method.args= list(degree=1),span=0.35)+
  facet_grid(~cut_number(length,n=4)) + ggtitle("log(Budget) model conditional
  on Length")
```

```
`geom_smooth()` using method = 'gam'
```



```
In [12]: options(repr.plot.width=8, repr.plot.height=4)
ggplot(movie_budgets, aes(y=log_budget, x= length)) + geom_point() + geom_smooth(
  method.args= list(degree=1),span=0.25)+
  facet_grid(~cut_number(year,n=3)) + geom_vline(data = movie_budgets,aes(xint
  ercept = 150)) +
  ggtitle("log(Budget) model conditional on Year")
```

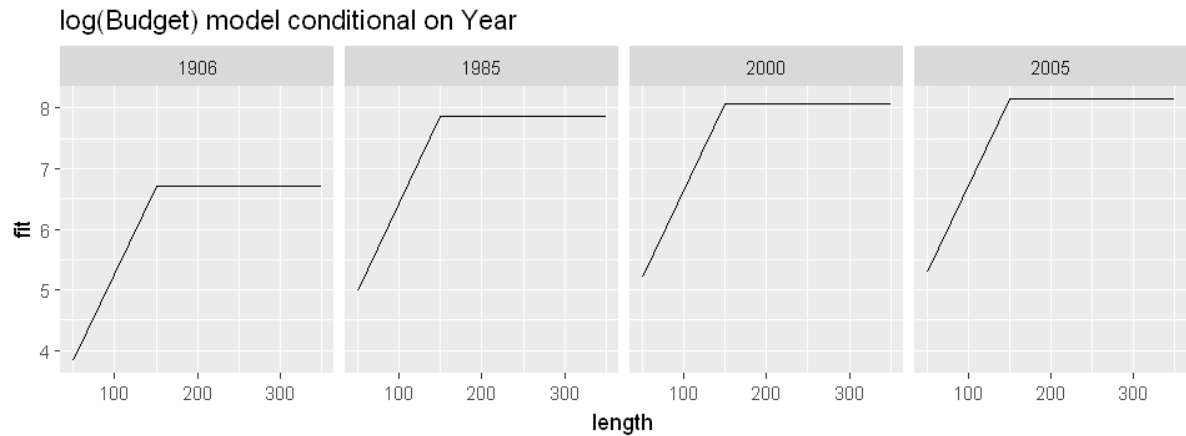
`geom\_smooth()` using method = 'gam'



When compared between length and Year, Length is more apt to be used on X-axis.

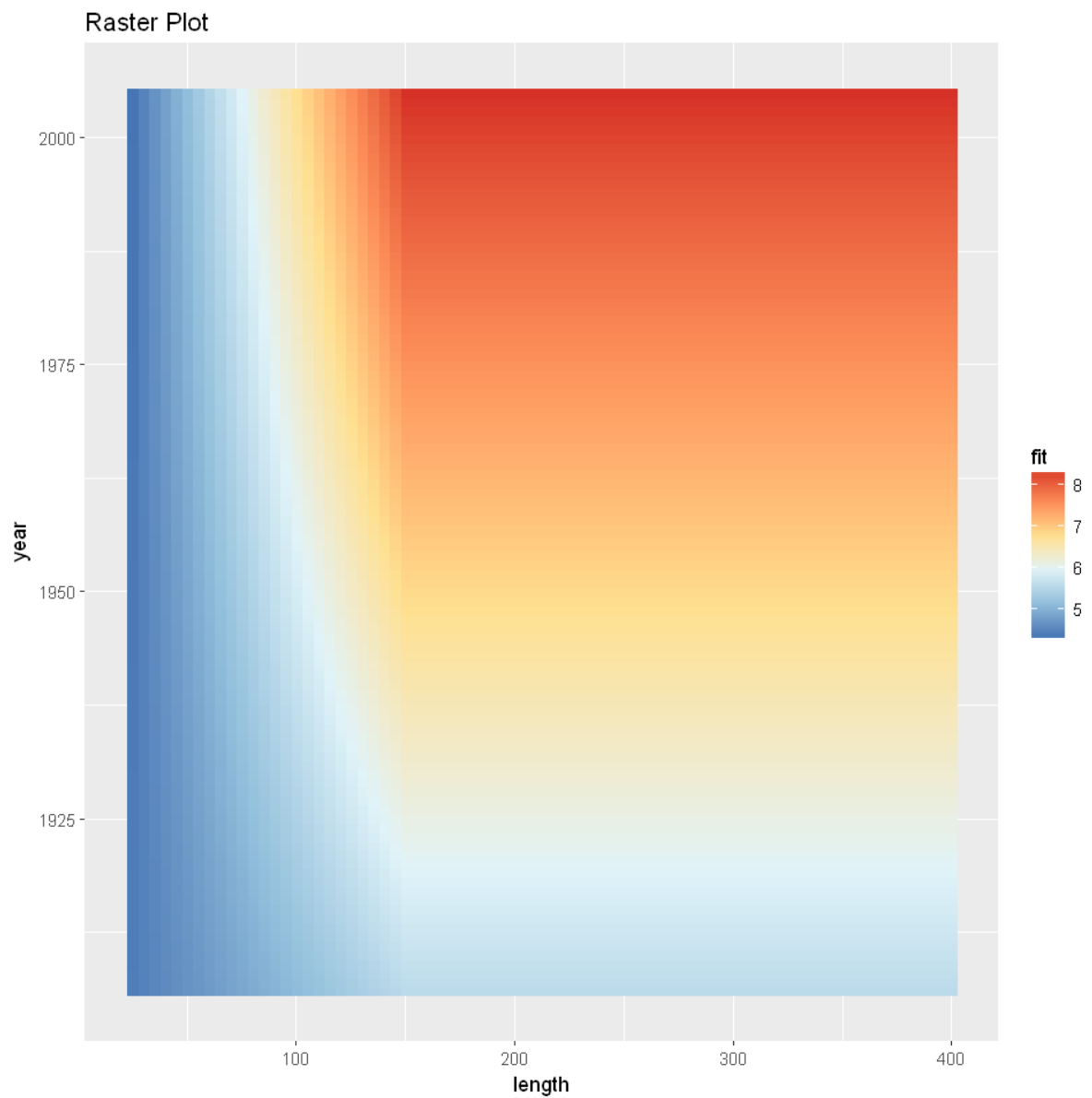
```
In [8]: length.bend = function(x) {
  return((x - 150) * (x < 150))
}
movie_budgets.rlm = rlm(log_budget ~ year + length.bend(length), data = movie_
budgets, psi = psi.bisquare)
movie_budgets.grid = expand.grid(year = c(1906, 1985, 2000, 2005), length = c(
50,150,250,350))
movie_budgets.grid.predict = predict(movie_budgets.rlm, newdata = movie_budget
s.grid)
```

```
In [11]: options(repr.plot.width=8, repr.plot.height=3)
ggplot(data.frame(movie_budgets.grid, fit = as.vector(movie_budgets.grid.predict)), aes(x = length, y = fit)) +
  geom_line() + facet_grid(~year) + ggtitle("log(Budget) model conditional on Year")
```



```
In [16]: mb.grid = expand.grid(year = seq(1906,2005,1), length = seq(25,400,5))
movie.budgets.rlm = rlm(log_budget ~ year * length.bend(length), data = movie_
  budgets,
                        psi = psi.bisquare)
mb_pr = predict(movie.budgets.rlm, newdata = mb.grid)
mb_plot_df = data.frame(mb.grid, fit = as.vector(mb_pr))
```

```
In [28]: options(repr.plot.width=8, repr.plot.height=8)
ggplot(mb_plot_df,aes(x = length, y = year, fill = fit)) + geom_raster() +
  scale_fill_distiller(palette = "RdYlBu") + ggtitle("Raster Plot")
```



```
In [27]: options(repr.plot.width=8, repr.plot.height=4)
ggplot(mb_plot_df,aes(x = length, y = year, z = fit)) + geom_raster(aes(fill =
fit))+
  scale_fill_distiller(palette = "RdYlBu") + geom_contour() + ggtitle("Contour
Plot")
```

