Flexible Models, Within-Country Heterogeneity, and Out-of-Sample Prediction

Carlisle Rainey

September 9, 2015

Open this in R Studio for amazingness.

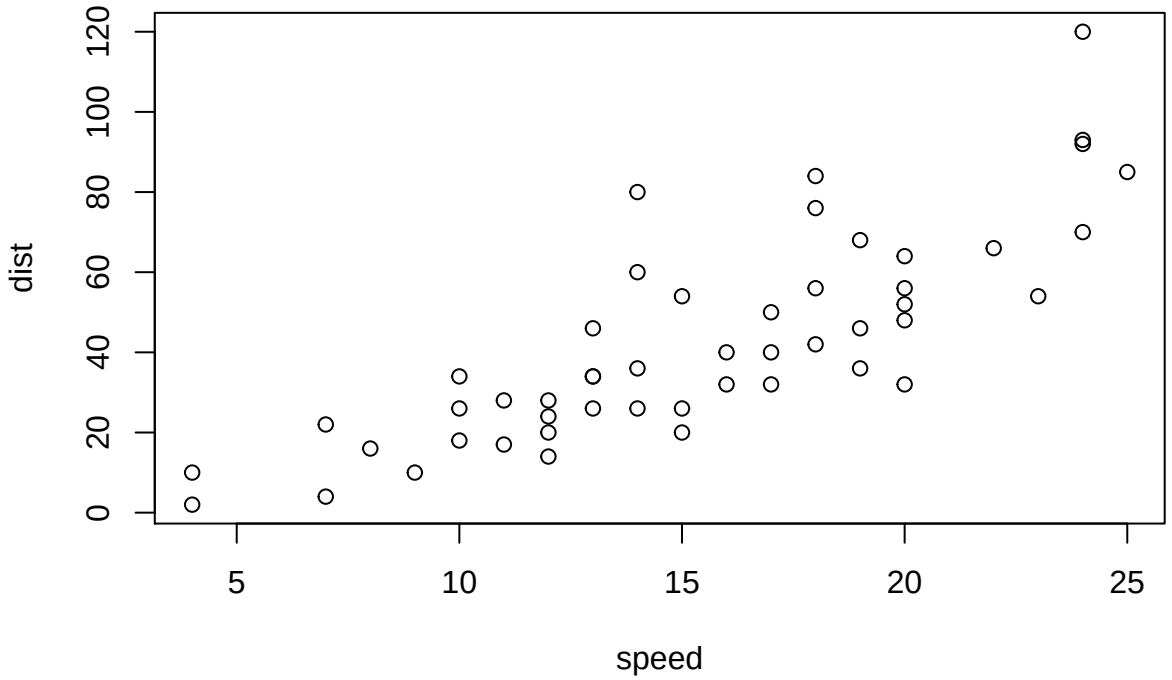
This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see http://rmarkdown.rstudio.com.

When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

```
summary(cars)
```

```
##      speed           dist
##  Min.   : 4.0   Min.   :  2.00
##  1st Qu.:12.0   1st Qu.: 26.00
##  Median :15.0   Median : 36.00
##  Mean   :15.4   Mean   : 42.98
##  3rd Qu.:19.0   3rd Qu.: 56.00
##  Max.   :25.0   Max.   :120.00
```

You can also embed plots, for example:



Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.

Simple Example with GAMs

Claim: Suppose you are interested in predicting a new observation y_{new} from x_{new} for an unobserved group. This group might simply fall outside the training set. Now suppose a true model $y_{ij} = x_{ij} + \gamma_j + \epsilon_{ij}$, where x_{ij} is fixed, $\gamma_j \sim N(0, \sigma^2_\gamma)$, and $\epsilon_{ij} \sim N(0, \sigma^2_\epsilon)$.

```
## geom_smooth: method="auto" and size of largest group is >=1000, so using gam with formula: y ~ s(x, b
```