

# DATA605 - Assignment 11

Nick Oliver

## Assignment 11

### Question

Using the “cars” dataset in R, build a linear model for stopping distance as a function of speed and replicate the analysis of your textbook chapter 3 (visualization, quality evaluation of the model, and residual analysis.)

### Solution

Loaded the `tidyverse` library

View the data set to find the column names

```
glimpse(cars)
```

```
## Rows: 50
## Columns: 2
## $ speed <dbl> 4, 4, 7, 7, 8, 9, 10, 10, 10, 11, 11, 12, 12, 12, 12, 13, 13, 13~
## $ dist <dbl> 2, 10, 4, 22, 16, 10, 18, 26, 34, 17, 28, 14, 20, 24, 28, 26, 34~
```

Use the `lm` function to fit a linear regression model

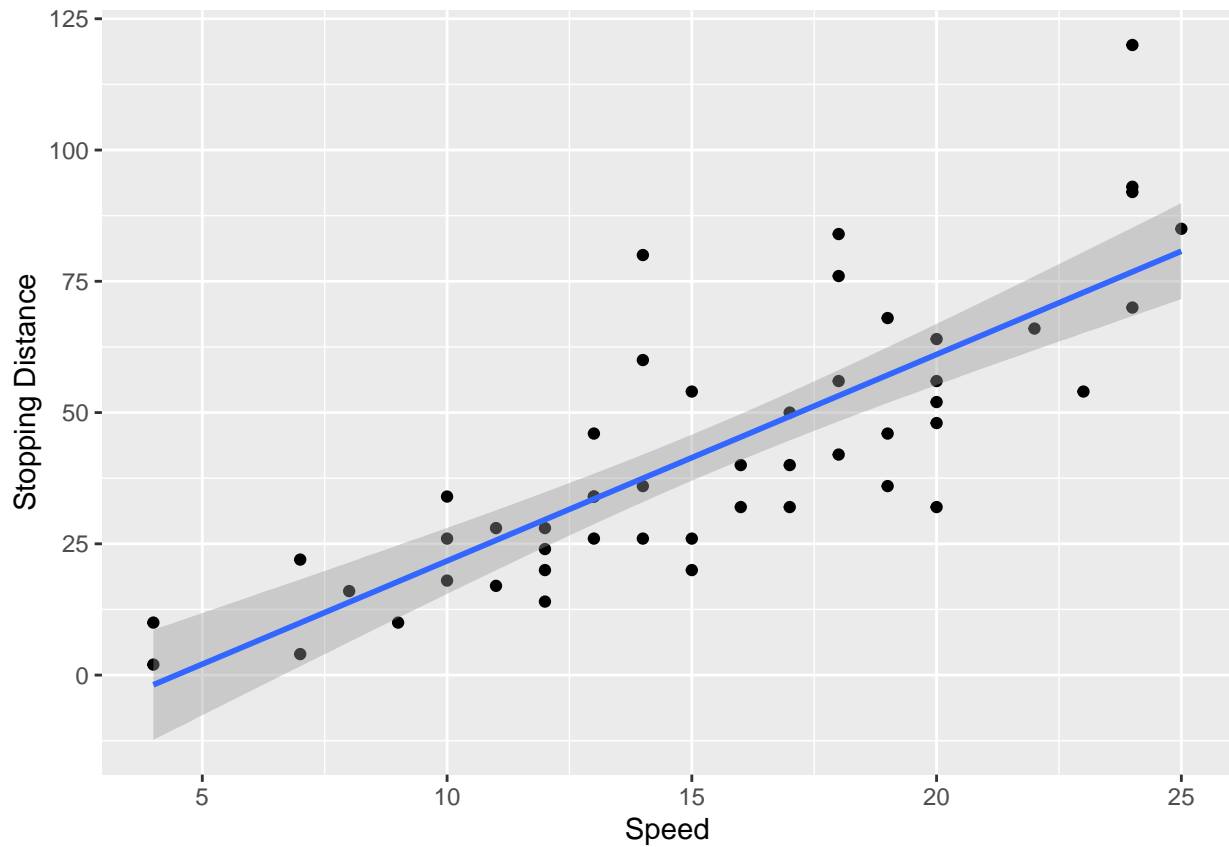
```
model <- lm(data = cars, dist ~ speed)
```

### Visualizing the data

Use `ggplot` to plot the scatter plot of speed vs stopping distance. Use `stat_smooth` to add a linear fitted line to the plot representing a linear regression model.

```
cars %>% ggplot(aes(x = speed, y = dist)) +
  geom_point() +
  stat_smooth(method = "lm", se = T) +
  xlab('Speed') +
  ylab('Stopping Distance')
```

```
## `geom_smooth()` using formula 'y ~ x'
```



You can also use the plot function and add the linear regression line using `abline`

```
plot(cars)
abline(model)
```

