NEX2019_HW1

Jiri Franc and his friends

R Markdown info:

This is an R Markdown document which 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:

Simple example

Summary and some plots of dataset car

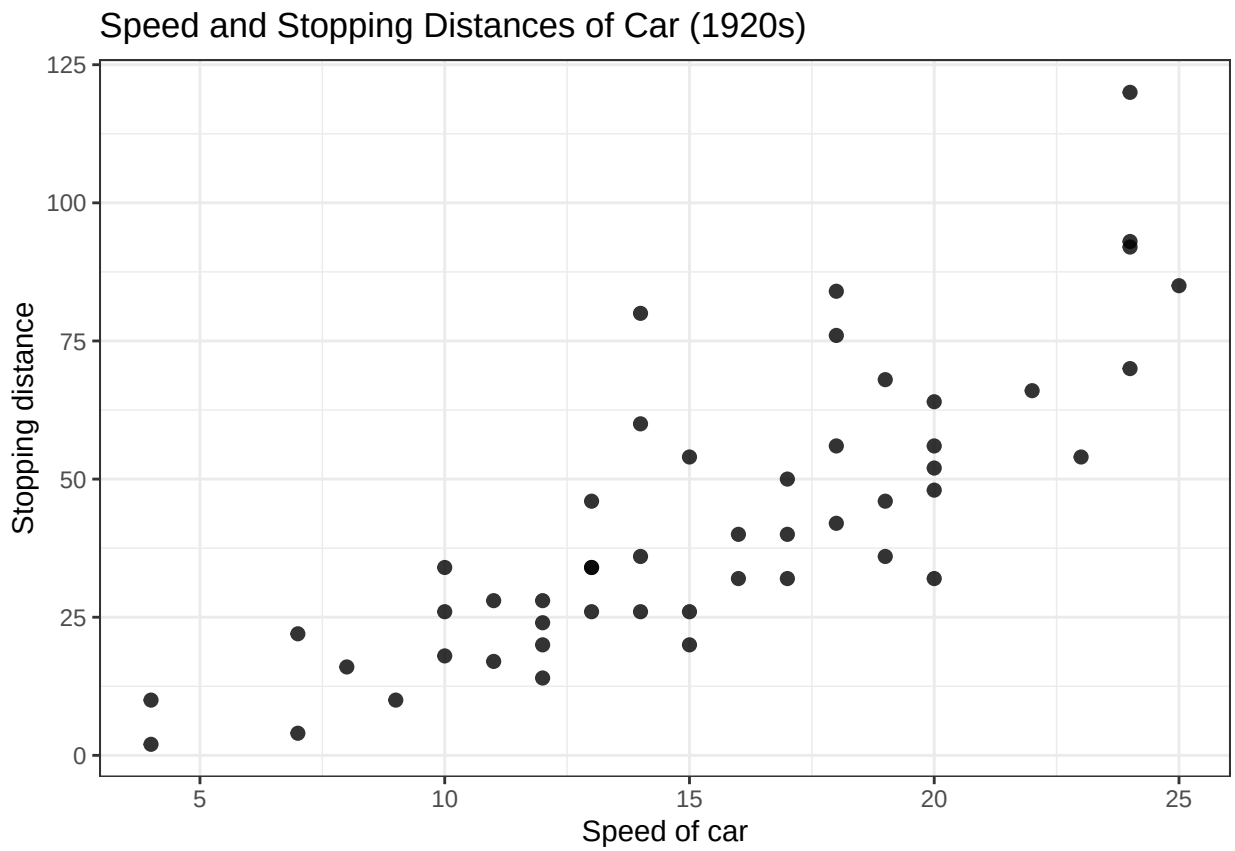
Summary of dataset cars:

```
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
```

Les's visualize the dataset by ggplot.

```
ggplot(cars, aes(x=speed, y=dist)) +
    geom_point(size=2, alpha=0.8) +
    theme_bw() +
    xlab("Speed of car") +
    ylab("Stopping distance") +
    ggtitle("Speed and Stopping Distances of Car (1920s)")
```



Question: How stopping time depends on car speed?

Answer: ...

Lets try some linear models:

```
cars_lm1 <- lm(dist ~ -1 +speed , data = cars)
cars_lm2 <- lm(dist ~  speed , data = cars)
cars_lm3 <- lm(dist ~  -1 + I(speed^2) , data = cars)
cars_lm4 <- lm(dist ~  speed + I(speed^2) , data = cars)
```

And see summary function from model with quadratic term and without intercept.

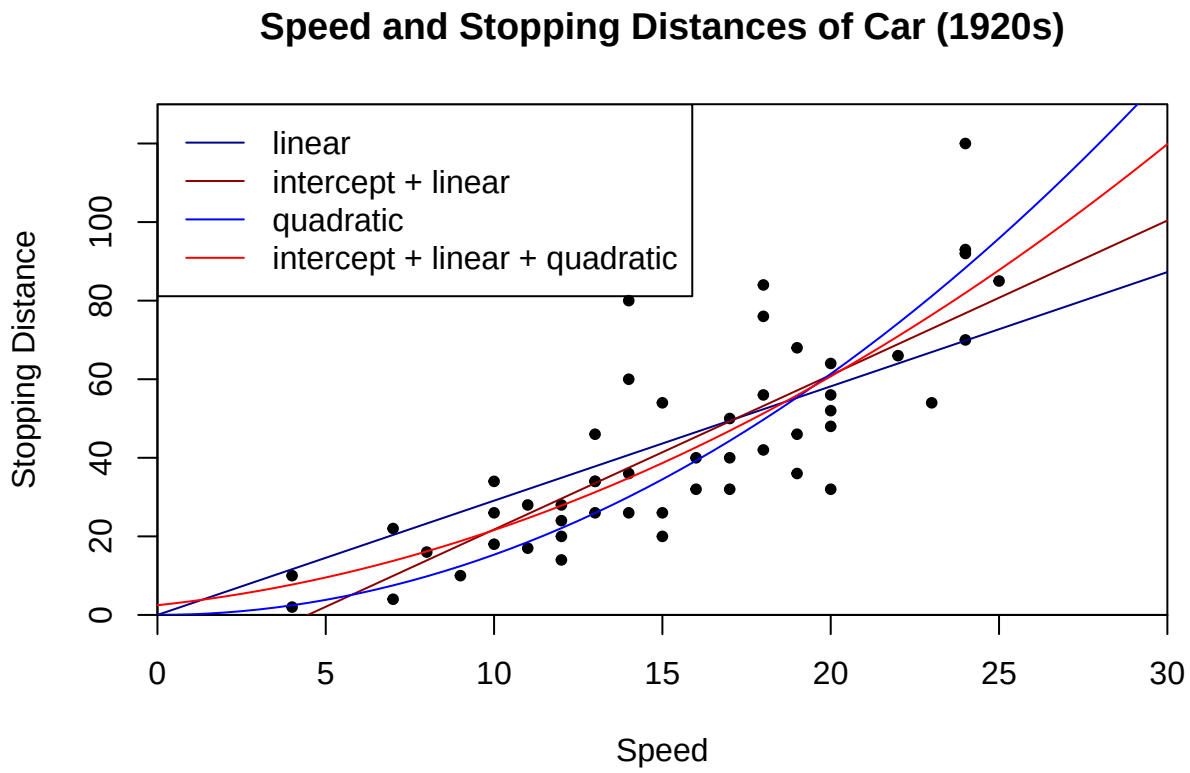
```
summary(cars_lm3)
```

```
##
## Call:
## lm(formula = dist ~ -1 + I(speed^2), data = cars)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -29.350  -7.988   1.325   8.080  49.939
##
## Coefficients:
##            Estimate Std. Error t value Pr(>|t|)
## I(speed^2) 0.153374   0.007122   21.54   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Residual standard error: 15.61 on 49 degrees of freedom
## Multiple R-squared:  0.9044, Adjusted R-squared:  0.9025
## F-statistic: 463.8 on 1 and 49 DF,  p-value: < 2.2e-16
```

Plot all models into one figure:

```
plot(dist ~ speed, data = cars, xlim = c(0,30), ylim = c(0,130),pch=20,
    col = "black", xaxs="i",yaxs="i",
    main="Speed and Stopping Distances of Car (1920s)",xlab="Speed", ylab="Stopping Distance")
abline(cars_lm1, col ="blue4")
abline(cars_lm2, col ="red4")
#lines(sort(cars$speed), fitted(cars_lm3)[order(cars$speed)], col='green')
lines(seq(0, 30, 0.5), predict(cars_lm3,data.frame(speed = seq(0, 30, 0.5))), col='blue')
lines(seq(0, 30, 0.5), predict(cars_lm4,data.frame(speed = seq(0, 30, 0.5))), col='red')
legend("topleft",legend = c("linear","intercept + linear",
                            "quadratic","intercept + linear + quadratic"),
                          lty = c(1,1,1,1),col = c("blue4","red4","blue","red"))
```



Question 02: Which model do you prefer and why?

I guess, model number 3. And why? Because I like Tesla model 3. Hahaha.

Note

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