Homework Assignment 01

Belov, Neoral, Sahan, Shulga

25.10.2018

R Markdown info:

In this assigment we use a data set "Experimentdata". The data set consisting of 36 observations of 6 variables.

-
-

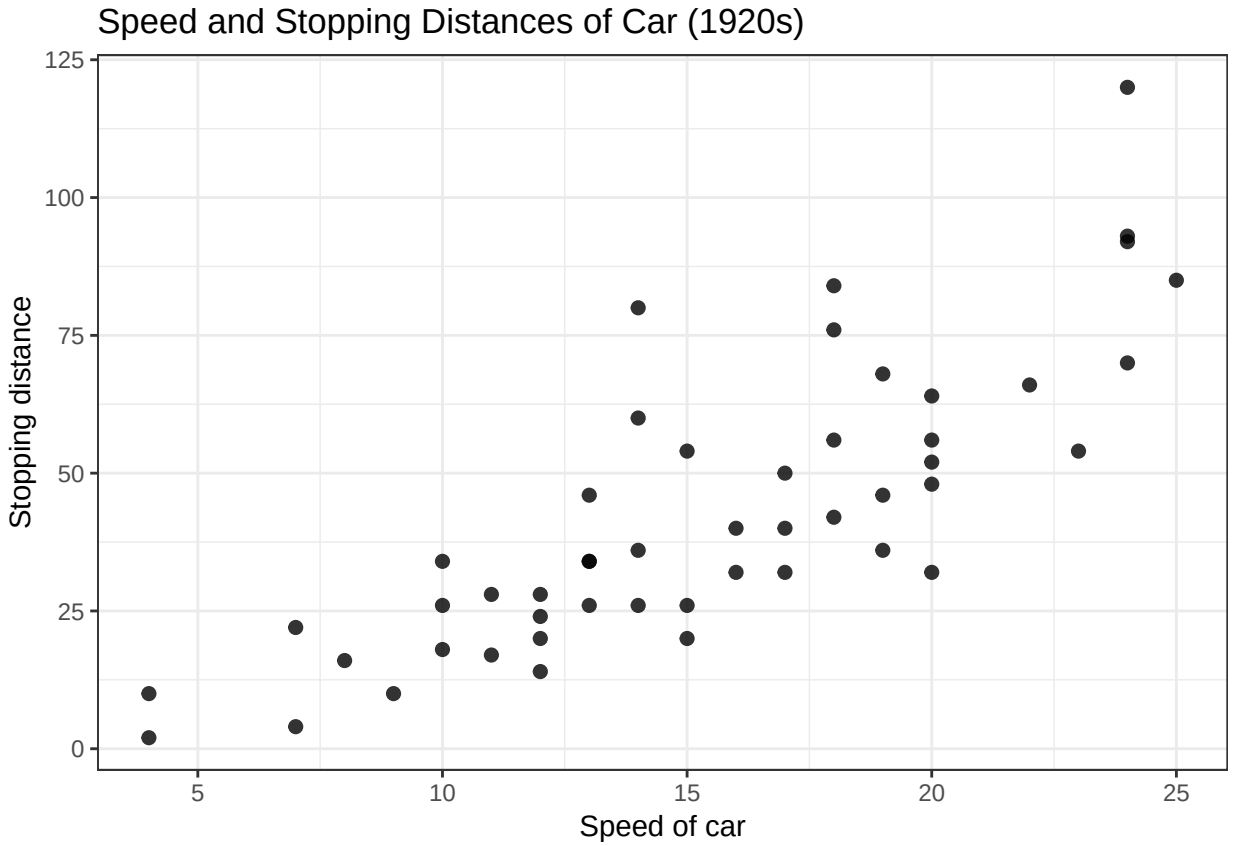
Example:

Summary of dataset cars:

```
##  BLOCK     HITS_L       HITS_R         HITS_SUM     DIAMETER HAND
##  1:9   Min.   : 5   Min.   : 4.00   Min.   : 9.00   1:12     B:12
##  2:9   1st Qu.: 8   1st Qu.: 8.00   1st Qu.:15.75   3:12     D:12
##  3:9   Median :11   Median :11.00   Median :22.00   5:12     N:12
##  4:9   Mean   :11   Mean   :10.75   Mean   :21.75
##        3rd Qu.:13   3rd Qu.:13.00   3rd Qu.:26.00
##        Max.   :23   Max.   :21.00   Max.   :44.00
```

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: What can we conclude from it?

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

And see summary function from model with intercept and quadratic term.

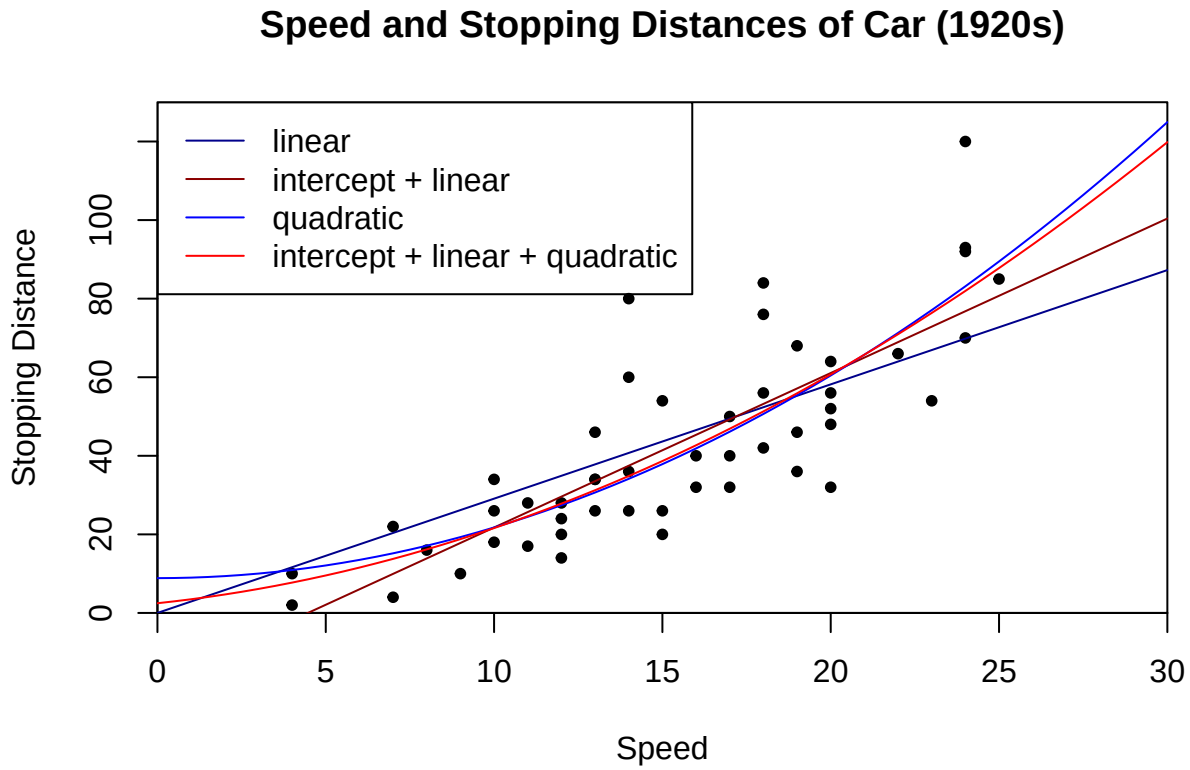
```
summary(cars_lm3)
```

```
##
## Call:
## lm(formula = dist ~ I(speed^2), data = cars)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -28.448  -9.211  -3.594   5.076  45.862
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  8.86005    4.08633   2.168   0.0351 *
## I(speed^2)   0.12897    0.01319   9.781  5.2e-13 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Residual standard error: 15.05 on 48 degrees of freedom
## Multiple R-squared:  0.6659, Adjusted R-squared:  0.6589
## F-statistic: 95.67 on 1 and 48 DF,  p-value: 5.2e-13
```

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:

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