Lab 18 R Script

Alexander Hernandez

11/01/2022

```
library(UsingR)
```

```
## Loading required package: MASS
```

```
## Loading required package: HistData
```

```
## Loading required package: Hmisc
```

```
## Loading required package: lattice
```

```
## Loading required package: survival
```

```
## Loading required package: Formula
```

```
## Loading required package: ggplot2
```

```
##
## Attaching package: 'Hmisc'
```

```
## The following objects are masked from 'package:base':
##
##      format.pval, units
```

```
##
## Attaching package: 'UsingR'
```

```
## The following object is masked from 'package:survival':
##
##      cancer
```

```
library(ggplot2)
library(MASS)
```

1) For the given values of x and y, fit a linear model and display it

a) Linear Model:

```r
x = c(1, 2, 2, 3, 4, 4, 5, 6, 6, 8, 9, 9, 11, 12, 12)
y = c(6, 7, 7, 9, 12, 13, 13, 15, 16, 19, 22, 23, 23, 25, 26)
values = data.frame(x, y)

model = lm(values$y ~ values$x)
model
```
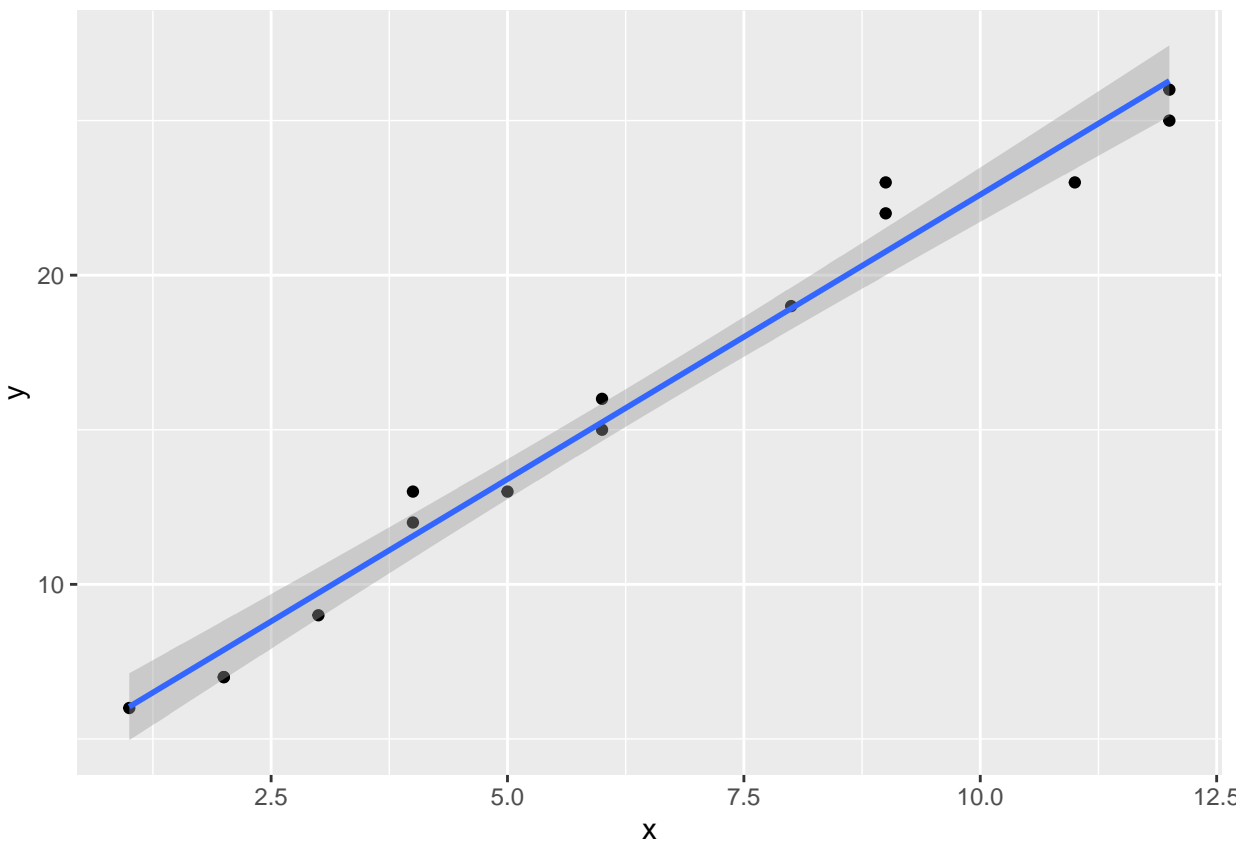
```
##
## Call:
## lm(formula = values$y ~ values$x)
##
## Coefficients:
## (Intercept)      values$x
##        4.20          1.84
```

b) ggplot:

```r
ggplot(values, aes(x, y)) +
  geom_point() +
  geom_smooth(method='lm')
```

```
## `geom_smooth()` using formula 'y ~ x'
```
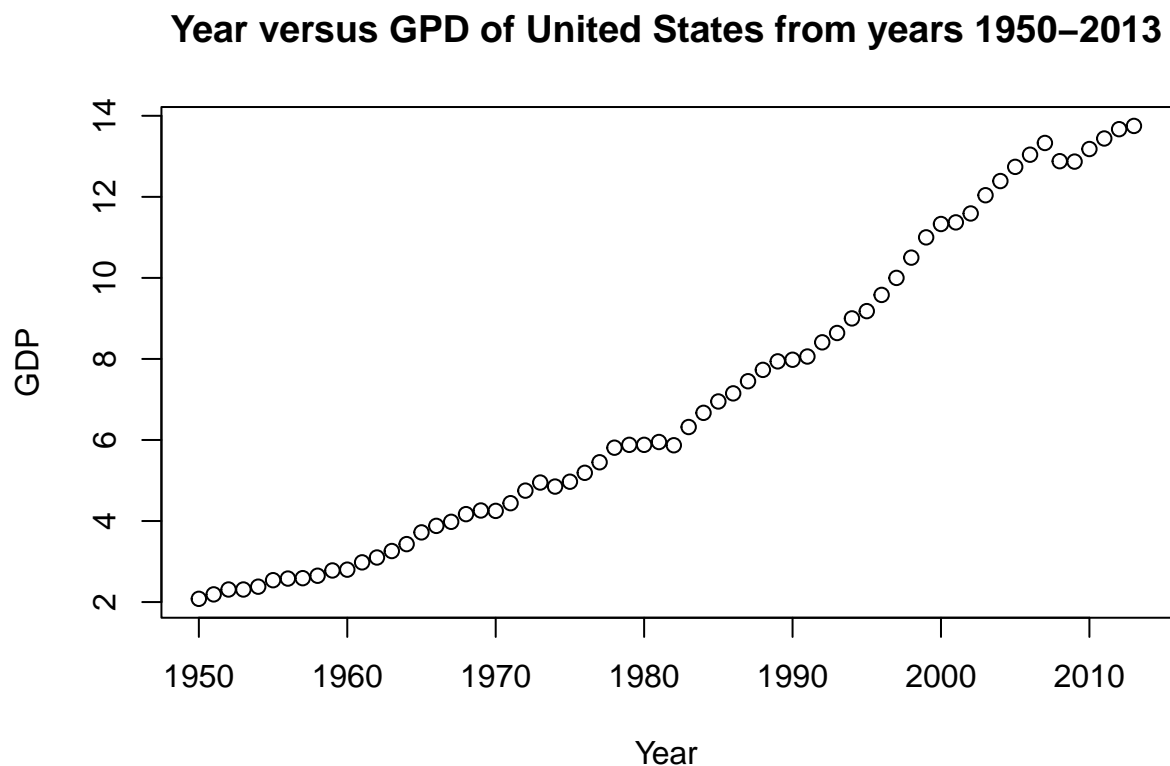
2) GDP of the US from 1950-2013

a) Display the data using a scatterplot

```
gdp = read.table("https://media.pearsoncmg.com/aw/aw_sharpe_business_3/datasets/txt/GDP_2013.txt",
                 skip=1)
colnames(gdp) = c("Year", "GDP")

plot(gdp,
     main = "Year versus GPD of United States from years 1950-2013")
```

**Year versus GPD of United States from years 1950–2013**

b) Fit a simple linear regression model

```
model = lm(gdp$GDP ~ gdp$Year)
model

##
## Call:
## lm(formula = gdp$GDP ~ gdp$Year)
##
## Coefficients:
## (Intercept)      gdp$Year
##   -387.8433        0.1993
```

c) Add the fitted line to the scatter plot

```
plot(gdp,
     main = "Year versus GPD of United States from years 1950-2013")
abline(model)
```

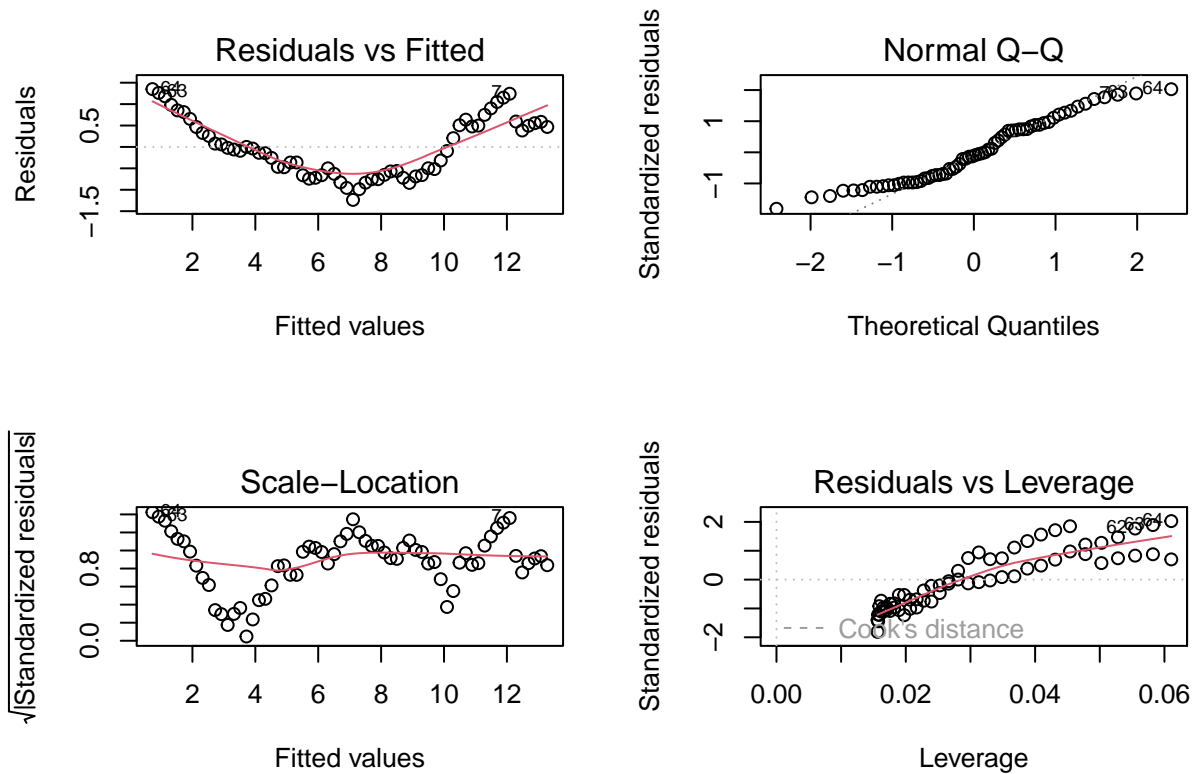**Year versus GPD of United States from years 1950–2013**
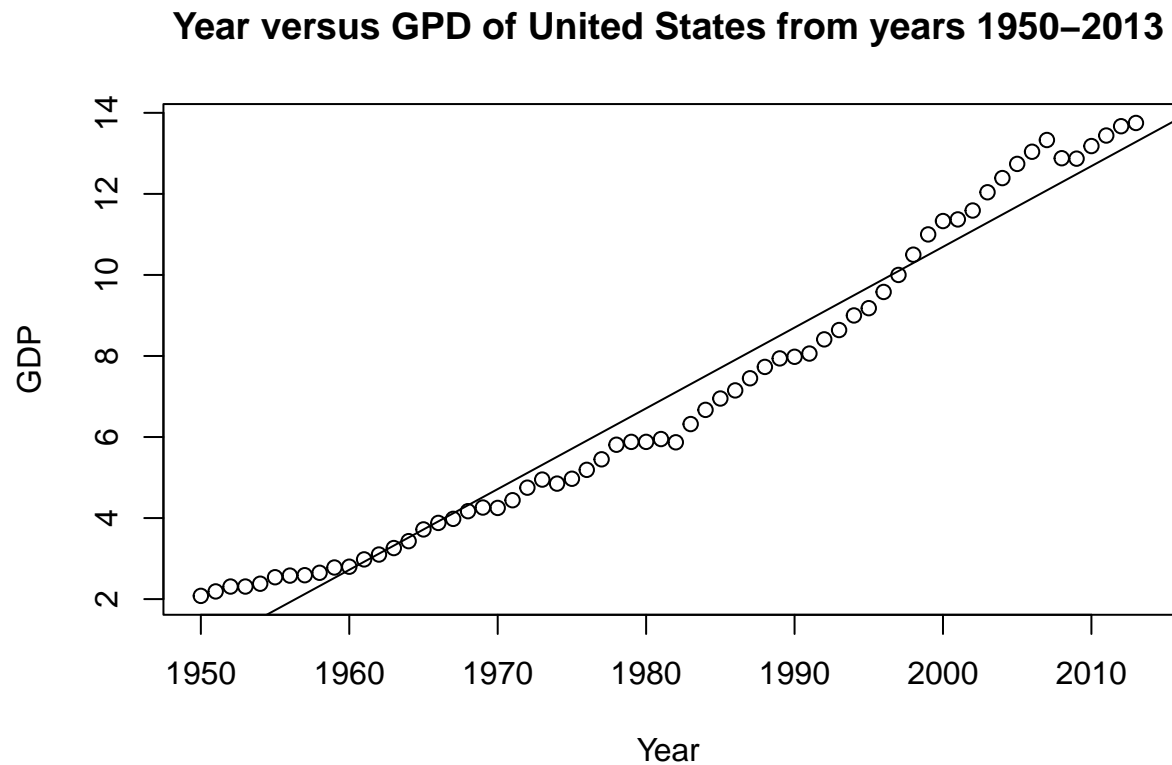
d) Determine the coefficient of determination

```
summary(model)$r.squared
```

```
## [1] 0.9673223
```

e) Analyze the residual plots. Is your model questionable?

```
par(mfrow=c(2,2))
plot(model)
```



```
# The models are questionable.
```

f) Use Box-Cox Transformation to see whether this model can be improved

```
boxcox(model)
```