R-2 Lab

2022-04-19

Group

Archana Gowda
Akhil Gupta
Akhilesh Manikeri
Rushabh Barbhaya

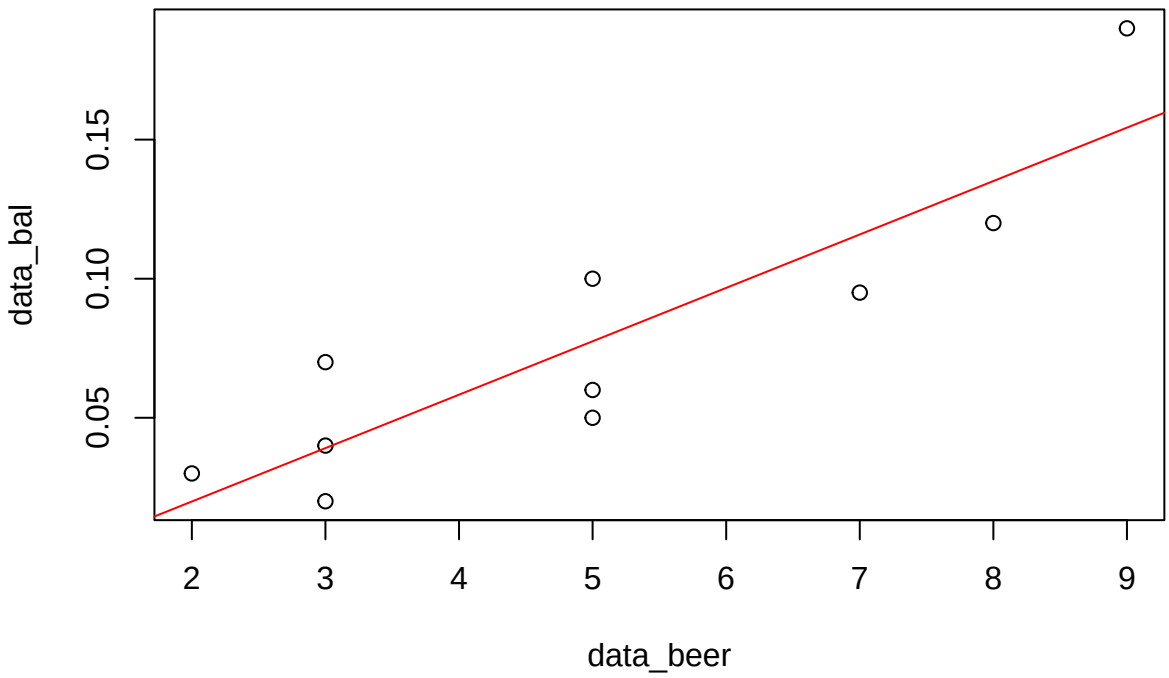
Question 1

It is well known that the more beer you drink, the more your blood alcohol level rises. Suppose we have the following data on student beer consumption

Student	1	2	3	4	5	6	7	8	9	10
Beers	5	2	9	8	3	7	3	5	3	5
BAL	0.10	0.03	0.19	0.12	0.04	0.095	0.07	0.06	0.02	0.05

Make a scatterplot and fit the data with a regression line.

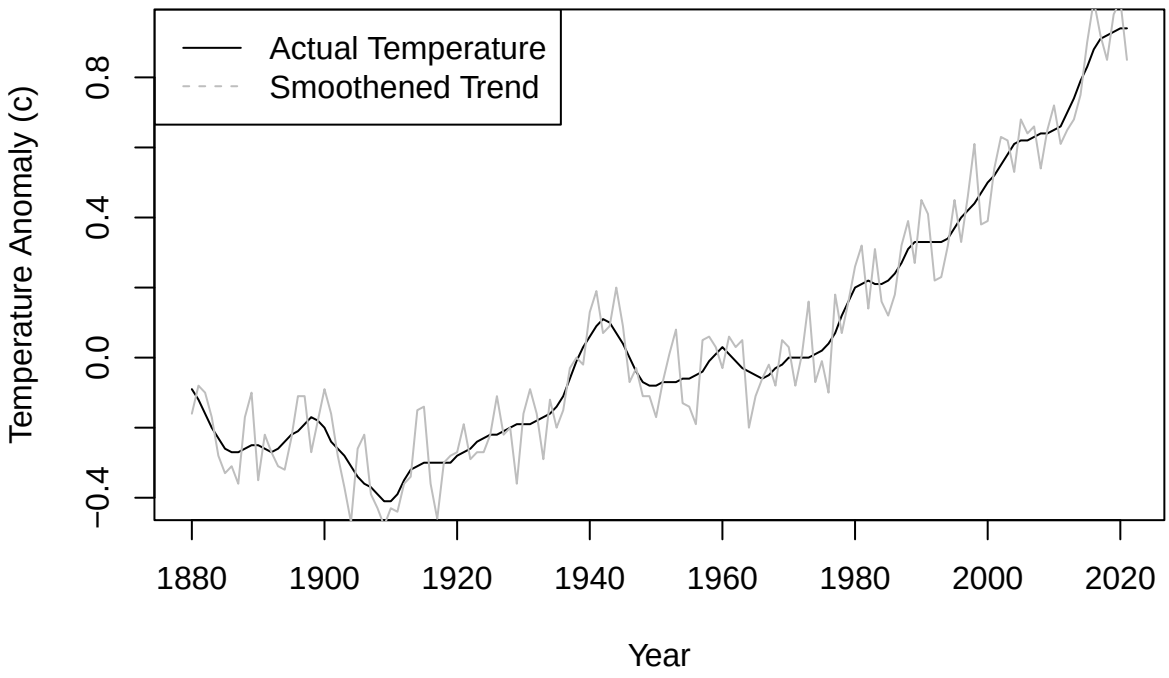
```
data_beer <- c(5,2,9,8,3,7,3,5,3,5)
data_bal <- c(0.1,0.03,0.19,0.12,0.04,0.095,0.07,0.06,0.02,0.05)
df <- data.frame(data_beer,data_bal)
sp <- plot(data_beer,data_bal)
abline(lm(data_bal~data_beer), col="red")
```



Question 2

Find a graphic of bivariate data (any type) from the newspaper, Internet or other media source. Use R to generate a similar figure (please cite your source, be clear on your R code and the functions you are using) Source = https://climate.nasa.gov/vital-signs/global-temperature/

```
URL = "https://climate.nasa.gov/vital-signs/global-temperature/"
graph <- read.csv("~/Documents/ANLY-560/R_lab2/graph.txt", sep="")
plot(graph$Year, graph$Lowess.5., col='black', type = 'l', xlab = "Year", ylab = 'Temperature Anomaly (
lines(graph$Year, graph$No_Smoothing, col='grey')
legend('topleft', legend = c("Actual Temperature", "Smoothened Trend"), col = c('black', 'grey'), lty =
```



Question 3

The normal plot is a fancy way of checking if the distribution looks normal. A more primitive one is to check the rule of thumb that 68% of the data is 1 standard deviation from the mean, 95% within 2 standard deviations and 99.8% within 3 standard deviations. Create 100 random numbers when the Xi are normal with mean 0 and standard deviation 1. What percent are within one standard deviation of the mean? Two standard deviations, three standard deviations? Is your data consistent with the normal? ((Hint: The data is supposed to have mean 0 and variance 1) (please use R commands and functions to respond)

```
random = rnorm(100, 0, 1)
k = 1
sigma = 1
n = length(random)
sig1 = (sum(-k*sigma < random & random < k*sigma)/n)
print(paste("Values within 1 standard deviation =", sig1))
```

```
## [1] "Values within 1 standard deviation = 0.63"
```

```
k = 2
sig2 = (sum(-k*sigma < random & random < k*sigma)/n)
print(paste("Values within 2 standard deviation =", sig2))
```

```
## [1] "Values within 2 standard deviation = 0.92"
```

```
k = 3
sig3 = (sum(-k*sigma < random & random < k*sigma)/n)
print(paste("Values within 3 standard deviation =", sig3))
```

```
## [1] "Values within 3 standard deviation = 1"
```

Question 4

Create 100 random numbers following the normal distribution (mean = 500, sd = 50), and draw the normal curve through the plotted histogram of the data.

```
x = rnorm(100, mean=500, sd=50)
hist(x, probability = TRUE)
curve(dnorm(x, mean = mean(x), sd = sd(x)), col='red', add = TRUE)
```