Visualization case

# http://www.grenswetenschap.nl/images/artikelfoto/OVLemminglente111.jpg Lemming sightings in Rotterdam

# Student: Jeroen van der Steen Student number: 0867254 Minor: Data Science

# Report Lemming sightings in Rotterdam

I computed the mean, standard deviation and standard error of the mean of lemming sightings in Rotterdam.  
You can see the results in this table below:

|  |  |
| --- | --- |
|  | **Lemming sightings** |
| **mean** | 25.47741 |
| **standard deviation** | 8.373527 |
| **standard error of the mean (sem)** | 0.8373527 |

If you observe the plot on figure: “Lemming sightings in Rotterdam”, you can see that spotters of the lemmings saw each time a different amount of lemmings between each hour. They spotted the lemmings in 100 hours. It could mean this is sample date, and that there is more data where they spotted longer.

## Lemming weights vs lifetimes I computed the mean, standard deviation and standard error of the mean of lemming sightings in Rotterdam. You can see the results in this table below:

|  |  |  |
| --- | --- | --- |
|  | **Lemming weights** | **Lemming lifetimes** |
| **mean** | 52.98623 | 89.51 |
| **standard deviation** | 30.25457 | 50.30623 |
| **standard error on the mean (sem)** | 3.025457 | 5.030623 |

If you observe the plot on figure: “Lemming weights vs lifetimes”, you can see a strong linear relationship between the weights and lifetimes.  
  
I found a fit for the model “Lemming weights vs lifetimes”.  
You can see the parameters in this table below:

|  |  |
| --- | --- |
| **Parameters** | **Values** |
| **intercept** | 2.769 |
| **slope of weights** | 1.637 |

So, the function of the fit model is:  
***slope\*lifetime + intercept  
1.637\*lifetime + 2.769***

# 

**Discuss**  
So, this is the final point of discussion about my lemmings research.  
Here I answer if my results make sense, first I start with a fact and then compare this with my results.  
  
***Fact:*** *A Lemming can become on average 2 years old and maximum 4 years old.*

The mean I found was 89.51 years, this means the lifetime data of the dataset is very likely incorrect.

***Fact:*** *The weight of a Lemming can be from 30 to 110 or up to 130 grams.*The mean I found was 52.98623 kg, this is way too high so incorrect.  
If the unit of measurement was grams instead of kilograms it could be the mean is correct.

# R Code

ds <- read.csv("lemming\_data\_21.csv")

library(plyr)

ds <- rename(ds,c(

"x" = "sightings", #Lemming sightings per hour in Rotterdam

"y" = "weights", #Lemming weight in kg

"z" = "lifetimes" #Lemming lifetime in years

))

pairs(ds)

**#1.**  
plot(ds$sightings, main="Lemming sightings in Rotterdam", xlab="Hours", ylab="Amount of Lemmings")

abline(h = mean(ds$sightings), col="red")

legend( "bottomleft", "Mean", lty=1, col="red" )

sem <- function(x) sd(x)/sqrt(length(x)) #Function to find the "standard error on the mean" of a sample

#25.47741

mean\_sightings <- mean(ds$sightings)

#8.373527

sd\_sightings <- sd(ds$sightings)

#0.8373527

sem\_sightings <- sem(ds$sightings)

#52.98623

mean\_weights <- mean(ds$weights)

#30.25457

sd\_weights <- sd(ds$weights)

#3.025457

sem\_weights <- sem(ds$weights)

#89.51

mean\_lifetimes <- mean(ds$lifetimes)

#50.30623

sd\_lifetimes <- sd(ds$lifetimes)

#5.030623

sem\_lifetimes <- sem(ds$lifetimes)

**#2.**  
plot(ds$weights, ds$lifetimes, main="Lemming weights vs lifetimes", xlab="Weights (kg)", ylab="Lifetime (years)")

**#3.**

model <- lm(ds$lifetimes ~ ds$weights)

#(Intercept) ds$weights

#2.769 1.637

plot(ds$weights, ds$lifetimes, main="Lemming weights vs lifetimes", xlab="Weights (kg)", ylab="Lifetime (years)")

abline(model, col="red")

legend( "topleft", "Linear fit", lty=1, col="red")

# Bibliography

*Norway Lemming*. (n.d.). Retrieved October 21, 2015, from University of Helsinki: http://www.helsinki.fi/science/metapop/species/lemmings.html

Vervloed, O. (n.d.). *Lemmingen*. Retrieved October 21, 2015, from Grenswetenschap: http://www.grenswetenschap.nl/permalink.asp?i=4384