

#1

Create three variables x1, x2 and x3 by using for each the rnorm() function. Combine the three variables in a data frame df1 by using the function data.frame().

#2

Study lab 6 handout for correlation. Are the three variables x1, x2, x3 in the df1, correlated to each other? If yes, why, if no, why? Use the function cor(df1) in estimating correlations to show the correlations for combinations, x1 with x2, x1 with x3 and x2 with x3.

#3 Perform cor.test() in pairs between x1 with x2, x1 with x3 and x2 with x3 and comment on the significance results.

#4 Use the data “energy” from the library ISwR of Dalgaard book and compare energy expenditure (expand) for the two groups of stature, a factor variable with two groups lean and obese. Compare the t.test p-value with the threshold of 0.05. Is the result significant? Comment on the results.

#5 Compare p-values of the results from point 4 (t.test()) with the results from wilxoc.test().

#6 Using the function hist() check the distribution of the variable “expand” from the “energy” data.

#7 Create a table from data energy on the variable stature. Report if there are statistical differences among the frequencies of lean individuals versus those obese. Use the following commands to create the table and them to test statistically if there are differences among the proportions of the two statures:

```
table(energy$stature)
```

```
prop.test(table(energy$stature))
```

#8 Copy and paste the data provided into a text file and save it as loans.txt. Study handout 7 to read the file into an R object, by starting with library(readr) or you can call it from the library(tidyverse). Use the function read_delim() to read the file into an R object named loans. B. Having created the object loans, report the loan (dollars_outstanding) per individual per year. I would suggest first you create a

variable `loans$out_doll <- loans$dollars_outstanding * 1000000000`. Also create another variable `loans$recip <- loans$recipients * 1000000`. Then divide `loans$out_doll / loans$recip` and assign the values to `loans$loansPerPerson`. The resulting new column will be in dollars per person. Comment on what you see as a trend for the loans in dollars per person (`loansPerPerson`) in the years.

```
library(tidyverse)
```

year	dollars_outstanding	recipients
2007	516.00	28.3
2008	577.00	29.9
2009	657.00	32.1
2010	749.80	34.3
2011	848.20	36.5
2012	948.20	38.3
2013	1040.20	39.6
2014	1129.80	40.7
2015	1212.40	41.6
2016	1292.20	42.3
2017	1366.90	42.6
2018	1439.20	42.9
2019	1510.30	42.9
2020	1515.00	42.8

#9 Perform linear regression analysis of dollars_outstanding (`loans$out_doll`) by recipients (`loans$recip`).

#10 If you perform/ run `plot(obj1)`, R will provide 4 graphs which will tell the story if the data analyzed were linear or not. Run also the following plot: `plot(loans$recip, loans$out_doll)`.