

School of Computer Science and Engineering

Lab exercise-3

Code/Course	:	CSE3020 – Data Visualisation	Date	:	02/02/2022
Lab Experiments		i. Construct of Scatterplot, Scatterplot With Encircling, Jitter Plot, Counts Chart, Bubble plot using R ii. Statistical Model using R with Orange dataset	Slot	:	L15+L16

Pre-requisite: Moderately familiar with basic concepts in R, including variables and functions, and with RStudio, the integrated development environment for programming in R.

Use various plots to identify the Correlations among the two variables. For example, practice the given lab manual to understand all the correlation plots.

1. Assume your own dataset (may be downloaded) and apply the correct correlation techniques to identify the relationship among the variables using various visualisation techniques (**Scatterplot, Scatterplot With Encircling, Jitter Plot, Counts Chart, Bubble plot, Marginal Histogram / Boxplot, Correlogram, Diverging bars, Diverging Lollipop Chart, Diverging Dot Plot and Area Chart**)

iii. Consider the Orange dataset, which is automatically included in R. Note that the O is capitalized!

1. Look at Orange using either head or as.tibble() (you'll have to run library(tidyverse) for that second option). What type of data are each of the columns?
2. Find the mean, standard deviation, and standard error of tree circumference.
3. Make a linear model which describes circumference (the response) as a function of age (the predictor). Save it as an object with <-, then print the object out by typing its name. What do those coefficients mean?
4. Make another linear model describing age as a function of circumference. Save this as a different object.

5. Call `summary()` on both of your model objects. What do you notice?
6. Does this mean that trees growing makes them get older? Does a tree getting older make it grow larger? Or are these just correlations?
7. Does the significant p value prove that trees growing makes them get older? Why not?