**RATHNAVEL SUBRAMANIAM COLLEGE OF ARTS AND SCIENCE (AUTONOMOUS) SULUR, COIMBATORE.**

**ASSIGNMENT QUESTIONS**

**UNIT III**

1. A. Compare and contrast the following operations on a data.frame and equivalent tibble. What is different? Why might the default data frame behaviors cause you frustration?

df <- **data.frame**(abc = 1, xyz = "a")

df$x

*#> [1] "a"*

df[, "xyz"]

*#> [1] "a"*

df[, **c**("abc", "xyz")]

*#> abc xyz*

1. What option controls how many additional column names are printed at the footer of a tibble?
2. Generate the correct format string to parse each of the following dates and times:

d1 <- "January 1, 2010"

d2 <- "2015-Mar-07"

d3 <- "06-Jun-2017"

d4 <- **c**("August 19 (2015)", "July 1 (2015)")

d5 <- "12/30/14" *# Dec 30, 2014*

t1 <- "1705"

t2 <- "11:15:10.12 PM"

1. Practice referring to non-syntactic names in the following data frame by:

a. Extracting the variable called 1.

b. Plotting a scatterplot of 1 versus 2.

c. Creating a new column called 3, which is 2 divided by 1.

d. Renaming the columns to one, two, and three:

annoying <- tibble(

`1` = 1:10,

`2` = `1` \* 2 + rnorm(length(`1`)) )

1. Identify what is wrong with each of the following inline CSV files.

What happens when you run the code?

read\_csv("a,b\n1,2,3\n4,5,6")

read\_csv("a,b,c\n1,2\n1,2,3,4")

read\_csv("a,b\n\"1")

read\_csv("a,b\n1,2\na,b")

read\_csv("a;b\n1;3")