Tutorial 3
Data Analytics

Course Instructor: Dr Sreeja S R
Department of Computer Science and Engineering
Indian Institute of Information Technology
IIIT Sri City

January 29 2021

R base functions for importing data

read.csv(): for reading **"comma separated value"** files (".csv").

read.csv2(): variant used in countries that use a comma "," as decimal point and a semicolon ";" as field separators.

read.delim(): for reading *"tab-separated value"* files (".txt"). By default, point (".") is used as decimal points.

read.delim2(): for reading *"tab-separated value"* files (".txt"). By default, comma (",") is used as decimal points.

R base functions for importing data

```
# Read tabular data into R
read.table(file, header = FALSE, sep = "", dec = ".")

# Read "comma separated value" files (".csv")
read.csv(file, header = TRUE, sep = ",", dec = ".", ...)

# Or use read.csv2: variant used in countries that # use a
comma as decimal point and a semicolon as field separator.
read.csv2(file, header = TRUE, sep = ";", dec = ",", ...)

# Read TAB delimited files
read.delim(file, header = TRUE, sep = "\t", dec = ".", ...)
read.delim2(file, header = TRUE, sep = "\t", dec = ",", ...)
```

- **file**: the path to the file containing the data to be imported into R.
- **sep**: the field separator character. "\t" is used for tab-delimited file.
- **header**: logical value. If TRUE, **read.table()** assumes that your file has a header row, so row 1 is the name of each column. If that's not the case, you can add the argument **header = FALSE**.
- **dec**: the character used in the file for decimal points.

Reading a local file

```
# Read a txt file, named "mtcars.txt"
my_data <- read.delim("mtcars.txt")

# Read a csv file, named "mtcars.csv"
my_data <- read.csv("mtcars.csv")

my_data <- read.table(file.choose(), sep ="|", header =
TRUE, dec =".")
```

t-test

```
t.test(x, mu = 0, alternative = "two.sided")
```

- **x**: a numeric vector containing your data values
- **mu**: the theoretical mean. Default is 0 but you can change it.
- **alternative**: the alternative hypothesis. Allowed value is one of “two.sided” (default), “greater” or “less”.