TMC 204
Statistical Data Analysis with R
Unit 5
Statistical function in R

Presented By : Aditya Joshi
Asst. Professor
Department of Computer Application
Graphic Era Deemed to be University
20-04-2020

Statistical functions:

Descriptive statistics consist of describing simply the data using some **summary statistics** and graphics. Here, we'll describe how to compute summary statistics using **R**

Import your data into R

- **Prepare your data**
- **Save your data** in an external .txt tab or .csv files
- **Import your data into R** as follow:

```
# If .txt tab file, use this
my_data <- read.delim(file.choose())
# Or, if .csv file, use this
my_data <- read.csv(file.choose())
```

Here, we'll use the built-in R data set named iris.

```
# Store the data in the variable my_data
my_data <- iris
```

Check your data

You can inspect your data using the functions head() and tails(), which will display the first and the last part of the data, respectively.

```
# Print the first 6 rows
head(my_data, 6)
```

R functions for computing descriptive statistics

Some R functions for computing descriptive statistics:

Description	R function
Mean	**mean**()
Standard deviation	**sd**()
Variance	**var**()
Minimum	**min**()
Maximum	**maximum**()
Median	**median**()
Range of values (minimum and maximum)	**range**()
Sample quantiles	**quantile**()
Generic function	**summary**()
Interquartile range	**IQR**()

Descriptive statistics for a single group

- Measure of central tendency: mean, median, mode
- Measure of variability

✓ Range: minimum & maximum

✓ Interquartile range

✓ Variance and standard deviation

✓ Median absolute deviation

Computing an overall summary of a variable and an entire data frame

- summary() function

✓Summary of a single variable

✓Summary of a data frame

- sapply() function
- stat.desc() function

Case of missing values