Lecture 4 - **R** Software

Ashutosh Rajput
Department of Mathematics

Rajdhani College
University of Delhi

How to explore and describe data

R includes descriptive statistics and plots for summarizing data. It can compute measures of center, like mean and median, as well as measures of spread, like (sample) standard deviation and range.

Examples: Compute some statistics for a set of exam scores.

```
> scores = scan()
1 :  81  81  96  77
5 :  95  98  73  83
9 :  92  79  82  93
13 :  80  86  89  60
17 :  79  62  74  60
21 :
Read 20 items
> range(scores)
[1]  60  98
> median(scores)
[1]  81
```

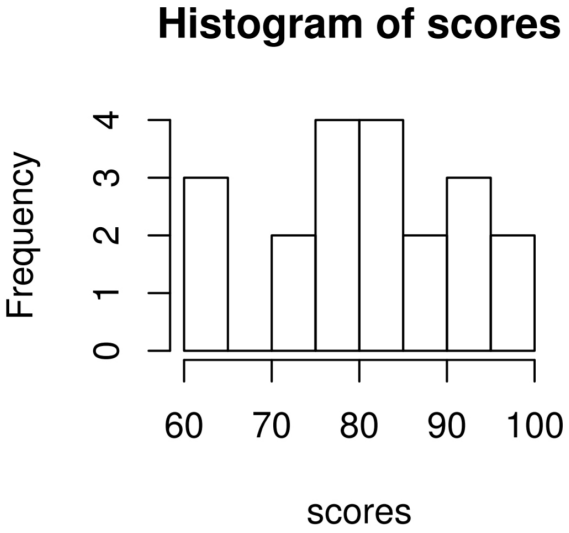
```
> mean(scores)
[1]  81
> sd(scores)
[1]  11.3555
```

Other common ways to explore a set of data are through histograms or box plots.
In **R**, the hist() function produces histograms, and the boxplot() function creates box plots.

Example: A group of exam scores to learn about the shape of the distribution.

```
> scores = scan()
1 :   81  81  96  77
5 :   95  98  73  83
9 :   92  79  82  93
13 :   80  86  89  60
17 :   79  62  74  60
21 :
Read 20 items
```

```
> hist(scores)
```



Example: The same set of exam grades with a histogram that we have customized with breaks on the "fives."

> hist(scores,breaks=c(55,65,75,85,95,105))