CT5102: Programing for Data Analytics

Assignment 1: Vectors

Maximum Marks (10)

The goal of this assignment is to explore how vector calculations can be performed in R. First, generate a vector set of random numbers from a normal distribution (note the seed value must be set to 111).

```
set.seed(111)

x<-rnorm(n=101,mean=72,sd=8)

> head(x)
[1] 73.88177 69.35411 69.50701 53.58123 70.63299 73.12223

> tail(x)
[1] 60.79167 82.07299 70.98018 66.16491 62.30911 76.79696
```

Based on this vector, write your own R code (no loops can be used) to perform the following (results should be stored in separate variables and printed using the cat() function).

- Calculate the mean (1 mark)
- Calculate the standard deviation (2 marks)
- Calculate the range (1 mark)
- Calculate the median (1 mark)
- Create a new vector with all values less than the mean (1 mark)
- Create a new vector with all values greater than or equal to the mean (1 mark)
- Find the top 5 values (1 mark)
- Find the bottom 5 values (1 mark)
- Find the absolute difference from the mean for each value (1 mark)

Submit your code to blackboard in a <01><*Surname*><*First Name*>*.R* file.

Deadline for submission is Wednesday September 16th 2015, 23:59.

Where appropriate, you can test your code using standard R functions, e.g. mean(), sd(), and range().

Note, the standard deviation for a sample is:

$$sd = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}}$$