NYU DEPARTMENT OF BIOLOGY

BIOL-UA 103 Bioinformatics in Medicine and Biology

Instructor:
Manpreet S. Katari

Course Description:
The goal of the course is to give students “hands-on” experience with biomedical data sets related to disease. The students will learn basic concepts in statistics and the programming language R, in order to analyze different forms of high-throughput genomic data. The course is divided into three sections: 1) Introduction to R 2) Doing Statistics in R, and 3) Analysis of Biomedical data using R. In the last section of the course, we will start using several packages, including Bioconductor, to analyze high throughput gene expression datasets such as Microarray and RNA-seq, array-CGH (Comparative Genome Hybridization) and also do GO-term enrichment analysis. By the end of the course students will have a good understanding of R and be able to apply their knowledge to any dataset.

Pre-requisites:
Principles of Biology II (BIOL-UA 12)
Molecular and Cell Biology I (BIOL-UA 21) although not required, this is a great course for someone who has already taken Biostatistics (BIOL-UA 42).

Textbook and Required Materials:
R in Medicine and Biology
Paul D. Lewis (2010, Jones & Bartlett Publishers)

Grading:

Assignments	40%
Midterm#1	15%
Final	15%
Final Project (presentation 5%)	20%
Attendance/Participation	10%

Topics:
Fundamentals of R: Packages and Data structures
Descriptive Statistics and Plotting Data
Inferential Statistical Analysis & Importing and Exporting Data in R
Apply function family and writing functions in R
Looking for association: Correlations and Hypergeometric test
Retrieving Public Gene Expression Datasets
Annotating Gene Lists Multivariate Analysis in R In class presentation