TOWARDS BIG DATA
INFORMATIC TOOLS FOR DATA MANIPULATION (PART 1)

Dott. Marco Milanesi
Faculty of Agriculture, Food and Environmental Sciences, Università Cattolica del Sacro Cuore

Course Aims
The course "Towards Big Data - Informatic Tools for Data Manipulation (Part 1)" aims at providing the theoretical foundations needed to manipulate and extract information from large datasets using informatics tools.

Methodology
Recent technological innovations in data production and collection have accelerated the generation of new data at an unprecedented pace. The Big Data revolution offers exciting opportunities and challenges to researchers, who now require specific informatics skills to fully leverage these ever-growing datasets. The lectures are designed to allow participants to apply each new concept introduced by the instructor. The highly interdisciplinary and practical nature of the topics, combined with the use of freeware platforms and software, will enable participants to integrate their newly acquired knowledge into their daily research routines from the very first lesson.

Course description

1. Assessment of students' bioinformatics skills
2. The big data era: fundamental concepts and implications for life sciences
3. Core bioinformatic skills:
    a. Key concepts, applications, and current challenges
    b. Introduction to programming
        i. Command-line tools, Linux environment, and HPC clusters
        ii. Overview of programming languages: Bash, R and Python
4. Fundamentals of programming
    a. Data structures, data management, and data exploration and visualization
    b. Dataset preparation for practical

Recommended texts
Course materials