INTER AMERICAN UNIVERSITY OF PUERTO RICO
METROPOLITAN CAMPUS
SCHOOL OF SCIENCE AND TECHOLOGY
NATURAL SCIENCE DEPARTMENT
MASTER'S IN SCIENCE IN MOLECULAR MICROBIOLOGY

SYLLABUS

I.   GENERAL INFORMATION

Course Title	:	Data Analysis and Bioinformatics
Code and number	:	MOMI 6400
Credits	:	2
Academic term	:	
Instructor	:	
Office hours and location	:	
Office telephone	:	
E-mail	:	

II.   DESCRIPTION

Application of bioinformatics and biostatistics to the analysis of a high volume of experimental data in microbiological investigations. Emphasis on solving practical problems through the use of open-source programs. It requires 30 hours of conference-laboratory.

III.   OBJETIVES

It is expected that at the end of the course, the student will be able to:
1. Use the main platforms for analysis of next-generation sequencing data and other sources, such as R or QIIME.
2. Apply bioinformatics and biostatistics to the analysis of a high volume of experimental data.
3. Characterize microbial diversity using different open source software.

Competencies of the graduate profile addressed in this course:

1.   Evaluate the most useful platforms for analyzing molecular microbiology data using emerging tools.

2.   Argue ideas and research results, before the scientific community, orally and in writing in Spanish and English.

IV. CONTENT
- A. Introduction to R
  - 1. Mathematical Functions
  - 2. Vectors
  - 3. Logical Values
- B. Applying Functions to Arrays or Data Frames
  - 1. Logical Operations
  - 2. Arrays
  - 3. Lists
  - 4. Read data in text format
  - 5. Objects in R
  - 6. Import/Export data
- C. Exploratory data analysis
  - 1. Pie charts
  - 2. Stem and leaf plots
  - 3. Histograms
  - 4. Scatter plots
- D. Numerical Descriptive Measures
  - 1. Measures of Centrality
    - a) Mean
    - b) Median
  - 2. Measures of Dispersion
    - a) Standard Deviation Sample Variance
    - b) Quantile
  - 3. Hypothesis, correlation, regression and variance analysis
  - 4. Significance tests
  - 5. P-value
  - 6. Hypothesis test for correlation
  - 7. Analysis of variance
- E. Ecological analysis of data using the VEGAN package in R
  - 1. Alpha diversity
  - 2. Beta diversity
  - 3. Diversity estimators
  - 4. Similarity analysis
- F. Sequences
  - 1. Sequence format
  - 2. Different types of databases
  - 3. Sequence format conversión
- G. DNA and Protein Databases
  - 1. DNA sequences in NCBI, Greengenes and SILVA databases.
  - 2. Protein Databases: NCBI, PDB and UniProt
- H. Sequence alignments
  - 1. Local and global alignments
  - 2. Paired alignment (pairwise alignment)
  - 3. Multiple sequence alignment
- I. Estimating alpha and beta diversity of general (non-sequence) data
  - 1. Creating data matrices
  - 2. Alpha and Beta diversity
- J. Microarray Data Analysis
  - 1. Sripts in R for microarray analysis 2.

K.   Quantitative Microbial Ecology using QIIME and MOTHUR for 16S and 18S sequences generated on sequencing platforms.
        1.   QIIME Tutorial
        2.   Cytoscape/Evident/ Picrust/Emperor
        3.   Mothur tutorial

V.   LEARNING ACTIVITIES

1. Illustrated lectures in power point format
2. Invited speakers
3. Exercises and computer simulations

VI.   EVALUATION

The evaluation of the course will be based on:

	Score	% of Final Grade
Class projects	400	70
Final Exam	100	30
Total	**500**	**100**

Class and Exams attendance
Class attendance is mandatory. A student who needs to be absent from a class should contact the professor prior to the class by phone or email. There will be no make-up exams, except for reasons of illness. In such case, make-ups will be offered with a proper doctor's excuse during the final exam period during the professor's office hours.

VII.   SPECIAL NOTES

A. Auxiliary services or special needs
All students who require auxiliary services or special assistance must request these at the beginning of the course or as soon as they know that they need them, through the proper registry, in the Office of Orientation with Sr. José Rodríguez.

B. Honesty, fraud, and plagiarism

Dishonesty, fraud, plagiarism and any other inappropriate behavior in relation to academic work constitutes major infractions sanctioned by the General Student Regulations. The major infractions, as stated in the General Student Regulations, may have as a consequence, suspension from the University for a definite period greater than one year or the permanent expulsion from the University, among others sanctions.

C. Use of electronic devices
Cellular telephones and any other electronic device that could interrupt the teaching and learning processes or alter the environment leading to academic excellence will be deactivated. Any urgent situation will be dealt with, as appropriate. The handling

of electronic devices that allow students to access, store or send data during evaluations or examinations is prohibited.

D. Compliance with the Provisions of Title IX

The Federal Higher Education Act, as amended, prohibits discrimination because of sex in any academic, educational, extracurricular, and athletic activity or in any other program or function, sponsored or controlled by a higher education institution, whether or not it is conducted within or outside the property of the institution, if the institution receives federal funds.

In harmony with the current federal regulation, in our academic unit an Assistant Coordinator of Title IX has been designated to offer assistance and orientation in relation to any alleged incident constituting discrimination because of sex or gender, sexual harassment or sexual aggression. The Assistant Coordinator, Sr. George Rivera, can be reached by phone at 787-250-1912, extension 2262 o 2147, or by e-mail griverar@metro.inter.edu.

The Normative Document titled Norms and Procedures to Deal with Alleged Violations of the Provisions of Title IX is the document that contains the institutional rules to direct any complaint that appears to be this type of allegation. This document is available in the Web site of Inter American University of Puerto Rico (www.inter.edu).

VIII.   EDUCATIONAL RESOURCES

Electronic Resources

1. The R Project for Statistical Computing; www.r-project.org/
2. Quantitative Insights Into Microbial Ecology; qiime.org/
3. Open source bioinformatics software platform for visualizing molecular interaction networks - www.cytoscape.org

IX.   BIBLIOGRAPHY

Textbooks

1. An Introduction to Bioinformatics Algorithms, Neil C. Jones and Pavel A. Pevzner ISBN-10: 0-262-10106-8 ISBN-13: 978-0-262-10106-6
2. Biostatistics for the Biological and Health Sciences with Statdisk, Marc Triola and Mario Triola, ISBN-13: 978-0321194367

Revisado 2022