

FORMULARIO DE PRESENTACIÓN DE CURSOS DE POSGRADO

1.1. Indique la denominación del curso propuesto:

Programación avanzada en R en Biología

1.2. Inserto en un carrera de posgrado

X Sí

No

1.3. En caso de que el curso ya sea dictado en otra carrera indique la siguiente información:

Carrera	Tipo de dictado	Modalidad	Carácter
Lic. Cs. Básicas con orientación en Biología	Semestral	Presencial	Electivo

2. Equipo docente.

2.1. Responsable a cargo.

Apellido: Olave

Nombre: Melisa

Documento: 32668841

Correo electrónico: molave@mendoza-conicet.gob.ar

CUIT/CUIL: 27-32668841-5

2.2. Integrantes del equipo docente (repetir cuantas veces sea necesario)

Apellido: Olave

Nombre: Melisa

Documento: 32668841

Correo electrónico: molave@mendoza-conicet.gob.ar

CUIT/CUIL: 27-32668841-5

Apellido: Chirilá

Nombre: María Virginia

Documento: 36743232

Correo electrónico: mvchirila@mendoza-conicet.gob.ar

CUIT/CUIL: 27-36743232 -8

3. Fecha probable de dictado

Semestre

X 1er

2do

mes: 16 de Mayo al 30/06

4. Número máximo y mínimo de alumnos

mínimo 5, máximo 30

5. Carga horaria propuesta

5.1. Exprese la carga horaria relacionada al dictado de la actividad en horas reloj.

Modalidad	Carga teórica	Carga práctica	Total	Porcentaje
Presencial	20hs	<input type="text" value="0"/>	20hs	40%
No presencial	<input type="text" value="0"/>	30hs	30hs	60%
Total	20hs	30hs	50hs	

6. Objetivos (2000 caracteres)

-Adquirir conocimientos prácticos en escribir códigos informáticos (scripts y programas) orientados a análisis típicos de proyectos de investigación en Cs. Biológicas, incluyendo manipulación de grandes bases de datos, automatización de procesos repetitivos, simulaciones y visualización de datos en gráficos de alta calidad. - Aprender a utilizar expresiones regulares.

-Automatizar descarga de datos desde páginas webs (por ej. GenBank).

-Aprender a optimizar el uso de recursos computacionales a través de análisis en paralelo.

-Contribuir al desarrollo de pensamiento crítico en el desarrollo de análisis de datos apropiado de acuerdo a la pregunta de investigación de interés.

7. Contenidos. (2000 caracteres)

Unidad 1: Automatización de análisis de grandes bases de datos.
Manipulación y filtrado de grandes bases de datos. Estadísticos resumen. Prueba de hipótesis univariada. Análisis multivariados. Modelos lineales. Exportar datos. Automatización de procesos repetitivos.

Unidad 2: Gráficos de alta calidad.
Gráficos simples: scatter, histogramas, caja y bigote. Gráficos personalizados.

Unidad 3: Simulaciones.

Introducción a las simulaciones. Simulador de números aleatorios. Ejemplos de aplicación de simulaciones, incluyendo simulador de mecanismos evolutivos (mutaciones, deriva génica, selección natural), captura y recaptura, crecimiento poblacional y capacidad de carga. Pruebas de hipótesis con simulaciones.

Unidad 4: Automatización de procesos y optimización de recursos.

Introducción a análisis en paralelo y uso de clusters. Optimización de recursos computacionales. Análisis en paralelo en R. Automatizar análisis de programas externos.

Unidad 5: Organización y presentación de código.

Introducción a herramientas de organización y presentación de código. Introducción a Lenguaje Markdown. RMarkdown, Jupyter notebook y Google colab.

8. Describa las actividades prácticas desarrolladas, indicando lugar donde se desarrollan y modalidad de supervisión. (Si corresponde). (2000 caracteres)

Se efectuarán dos encuentros semanales, de los cuales uno será teórico y otro de trabajos prácticos. Las actividades presenciales se realizarán en las instalaciones de la Facultad Cs. Exactas y Naturales de las Universidad Nacional de Cuyo. Los trabajos prácticos pueden realizarse de manera presenciales o no presenciales (opcional). Las actividades se basan en ejercicios que deberán ser resueltos utilizando el R, en concordancia con los temas abordados en la clase teórica anterior. Como trabajo práctico final el alumno deberá presentar un breve trabajo de investigación escrito y con defensa oral, incluyendo análisis de datos en R, de cualquier tema de interés en Cs. Biológicas. Si los tuviera, se recomienda utilización de datos propios (ej. datos de tesis). Si el alumno no cuenta con datos propios, se buscarán datos reales de libre acceso (ej. GenBank, Dryad, Figshare), o bien basados en simulaciones.

Además, para actividades no presenciales, se contará con un canal de Slack para intercambiar preguntas y respuestas con el grupo en general, ayudar a soluciones de problemas de manera grupal.

Se recomienda que el alumno cuente con computadora personal.

9. Bibliografía propuesta (2000 caracteres)

Quicke, D. L., Butcher, B. A., & Welton, R. A. K. (2020). Practical R for Biologists: An Introduction. CABI.

Gentleman, R. (2008). R programming for bioinformatics. Chapman and Hall/CRC.

Hector, A. (2021). The new statistics with R: an introduction for biologists. Oxford University Press

10. Modalidad de evaluación y requisitos de aprobación y promoción. (2000 caracteres)

Se seguirá un formato de evaluación continua, en el que se tendrá en cuenta el desempeño de los alumnos en actividades prácticas y la participación en clases teóricas (25% del puntaje total), entrega de tareas (25%) y en la elaboración y presentación oral del proyecto de investigación (50% del puntaje total). Todas las semanas se deberán presentar resoluciones de los ejercicios prácticos correspondientes. El proyecto final incluirá una instancia de revisión y corrección, una presentación final escrita breve y una exposición oral.

Aprobación y promoción: Los alumnos cuya calificación global sea de al menos 60% del puntaje máximo posible, que hayan aprobado los trabajos prácticos con evaluación o informe.

11. Tiempo de entrega de evaluaciones y calificaciones una vez finalizado el curso

15 días

12. Ingrese toda otra información que considere pertinente, incluidos requisitos específicos si corresponde. (1600 caracteres)

Alumnos en cualquier nivel de avance de cursado de postgrado son bienvenidos.

Conocimiento previo de R es necesario, y se recomienda haber cursado anteriormente "Introducción a programación en R en Biología".

Es importante tener en cuenta que este no es un curso de estadística en R. Es un curso de programación en el que se utilizarán diferentes ejemplos orientados al uso de proyectos de investigación en Cs. Biológicas.