

Curso de Programación en Biológica Computacional

Clase I. Configuración inicial

Introducción a la Bioinformática. Desarrollo de software para ciencias biológicas.

Unix/Linux: organización, arquitectura, sistema de archivos..

La terminal de Unix.

Navegación y manejo de archivos.

Instalación de software y configuración de entorno de trabajo local y online.

Python: fundamentos, versiones y distribuciones, instalación y ejecución, instalación de librerías, ayuda disponible.

Primer programa.

Clase II. Scripting y control de versiones

Scripting en Linux.

Grep, sed, awk.

Control de versiones con Git.

Git local y remoto.

Trabajo colaborativo.

Clase III. Python: Variables y control de flujo

Sintaxis general.

Guías de estilo: PEP8

Definición y uso de variables.

Operadores: matemáticos y relacionales.

Tipos de objetos básicos: int, float, string, lista, diccionario, set, tupla.

Estructuras de control: if, elif, for, while.

Clase IV. Python: Funciones, manejo de archivos e interpretación de errores

Definición y uso de funciones.

Manipulación de texto:

- Operaciones y métodos útiles sobre cadenas de texto
- Conversión entre tipos de datos

- Expresiones regulares
- Lectura y escritura de archivos.
Pase de argumentos a programas.
Ejecución de programas externos.
Interpretación de errores.
Documentación de programas.
Entornos de trabajo.

Clase V. Python: Manejo de datos

Tablas de datos con Pandas:

- Lectura y escritura de datos estructurados.
- Datos en grupos.
- Datos faltantes.
- Estadística descriptiva.

Clase VI. Python: Visualización de datos

Generación de gráficos con Matplotlib, Seaborn, ggplot:

- Gráficos de dispersión
- Gráficos de barra
- Gráficos de líneas
- Gráficos de caja y bigote
- Histogramas
- Mapas de calor

Personalización de gráficos.

Gráficos para publicación.

Clase VII. Python para Biología

Análisis de datos biológicos: Biopython. Manejo de datos genómicos (genomas, anotaciones, lecturas de secuenciación en formato fastq, etc) para el cálculo de estadísticas básicas (estándares).

Scipy: métodos estadísticos.

Filogenia: Bio.Phylo; ETE Toolkit.

Clase VIII. Proyecto supervisado: Análisis e Interpretación de datos

Planteo de proyectos individuales.

Formulación de hipótesis y esquema de trabajo.

Obtención de datos crudos.

Limpieza de datos.

Conversión a formatos procesables.
Estadística descriptiva y avanzada.
Generación e interpretación de gráficos.
Interpretación de resultados.
Puesta en común de proyectos: factibilidad de análisis, sugerencias.
Conclusiones generales.
Directivas para informe final.

Modo de evaluación

En el marco de esta propuesta pedagógica se entiende al aprendizaje como un proceso con esquemas individuales y característicos de cada estudiante. La evaluación, por lo tanto, será de manera continua y teniendo en cuenta el aporte de cada alumno en clase, su progreso y dedicación. La nota final del curso dependerá de la aprobación de un trabajo final de programación y análisis de datos desarrollado durante las últimas clases.