

CLASE 01 - PRESENTACIÓN DEL CURSO

**DBT 845 - Investigación reproducible y análisis de datos
biotecnológicos con R.**

Dr. José Gallardo Matus | <https://genomics.pucv.cl/>

Pontificia Universidad Católica de Valparaíso

19 March 2022

PLAN DE LA CLASE

- ▶ Palabras de Bienvenida.
- ▶ Revisión del Programa del curso.
- ▶ Revisión de recursos de comunicación y aprendizaje.
- ▶ Introducción.
- ▶ Trabajo práctico con R.

PROFESORES

Dr. José Gallardo

jose.gallardo@pucv.cl

Profesor de genética y genómica aplicada

Biólogo marino

Doctor en Ciencias

Invitada

Dr. María Angélica Rueda

maria.rueda.c@pucv.cl

Investigadora Postdoctoral

Licenciada en Matemáticas

Doctor en Ciencias Agropecuarias

OBJETIVOS DE APRENDIZAJE

Al final de curso los alumnos serán capaces de usar el lenguaje de programación R para:

- ▶ Aplicar los fundamentos de la investigación reproducible y del análisis exploratorio de datos biotecnológicos.
- ▶ Aplicar los fundamentos de la inferencia estadística en el análisis de datos biotecnológicos.
- ▶ Aplicar diferentes modelos lineales y generalizados para el análisis de datos de biotecnológicos.
- ▶ Aplicar algunas técnicas de análisis multivarido para el análisis de datos de biotecnológicos.

REQUISITOS

- ▶ **Bioestadística:** curso aprobado de nivel universitario.
- ▶ **Inglés:** Los softwares R, Rstudio, Rmarkdown y todas las librerías de análisis estadístico que se usarán en el curso solo están disponibles en inglés. Alumnos sin competencias de lectura en inglés no deberían tomar el curso.
- ▶ **Programación:** Deseable pero no es excluyente. Los alumnos sin experiencia previa en programación deben considerar 1-2 horas de estudio y autoaprendizaje adicional por día de clases para poder alcanzar una comprensión avanzada de los objetivos de aprendizaje del curso.

CONTENIDOS

- ▶ **UNIDAD 1. Investigación reproducible y análisis exploratorio de datos.**

Palabras clave: Reproducibilidad, R, Rstudio, Rmarkdown, Github, variables aleatorios, distribución de probabilidad, análisis exploratorio de datos.

- ▶ **UNIDAD 2. Contrastes de hipótesis paramétrica y no paramétrica.**

Palabras clave: Parámetro, estadístico, correlación, inferencia estadística, contraste de hipótesis y análisis de supervivencia.

- ▶ **UNIDAD 3. Modelos lineales y análisis multivariado**

Palabras clave: Regresión lineal, regresión lineal múltiple, Regresión cuadrática, regresión logística, análisis multivariado.

EVALUACIÓN DEL CURSO

- ▶ La evaluación del curso consiste en el desarrollo de un proyecto personal de investigación y análisis de datos biotecnológicos con R.
- ▶ Se dará énfasis a que los alumnos resuelvan un problema de análisis de datos con R usando datos relacionados a su investigación de tesis doctoral.
- ▶ El trabajo se desarrolla en dos etapas, la primera pondera un 40% y la segunda un 60% de la nota final.

CONDICIONES DE APROBACIÓN

- ▶ **Nota mínima:** La nota mínima de aprobación es un 4,0 en una escala de 1-7 y con un 60% de exigencia.
- ▶ **Asistencia a clases:** Es causal de reprobación de la asignatura, no cumplir con el mínimo de asistencia de un 80%. Esto es independiente de que las calificaciones parciales o totales sean mayores de 4,0.
- ▶ **Plazo de entrega de trabajos:** No entregar los trabajos en los plazos establecidos para ello será calificado con la nota mínima (1,0).

RECURSOS DE APRENDIZAJE Y COMUNICACIÓN

- ▶ **Repositorio:** Todo el material disponible en Google Drive.
- ▶ **Comunicación:** Foro de preguntas y respuestas por SLACK.
- ▶ **Material docente:** Diapositivas de clases, videos, guías de aprendizaje y códigos de programación para el análisis de datos con R.
- ▶ **Rstudio:** Acceso a Rstudio.cloud por 5 meses.

PRESENTACIÓN DE LOS PARTICIPANTES

