

---

**Nombre de la unidad curricular:** Seminario: Ciencia de datos con R

---

**Forma parte de la Oferta Estable:** No

---

**Licenciaturas:** Estadística, Matemática

---

**Frecuencia y semestre de la formación al que pertenece:** Primer semestre única vez

---

**Créditos asignados:** 5, Área A

---

**Nombre del/la docente responsable:** Ernesto Mordecki

---

**E-mail:** [mordecki@cmat.edu.uy](mailto:mordecki@cmat.edu.uy)

---

**Requisitos previos:** Elementos de cálculo y álgebra lineal. Elementos de estadística y/o probabilidad

---

**Ejemplos de unidades curriculares de Facultad de Ciencias u otros que aportan dichos conocimientos:** Cálculo 1, Álgebra Lineal 1, Probabilidad, Bioestadística

---

**Conocimientos adicionales sugeridos:**

Ninguno

---

**Objetivos de la unidad curricular:**

**a) Herramientas, conceptos y habilidades que se pretenden desarrollar**

Se trata de una introducción a la estadística a través de la resolución de problemas mediante el R

**b) En el marco del plan de estudios**

**Temario sintético de la unidad curricular:**

Principios de modelación estadística y aprendizaje automático

**Temario desarrollado:**

Statistical Learning

- 2.1 What Is Statistical Learning . . . . .
- 2.1.1 Why Estimate  $f$  . . . . .
- 2.1.2 How Do We Estimate  $f$  . . . . .
- 2.1.3 The Trade-Off Between Prediction Accuracy and Model Interpretability . . . . .
- 2.1.4 Supervised Versus Unsupervised Learning . . . . .
- 2.1.5 Regression Versus Classification Problems . . . . .
- 2.2 Assessing Model Accuracy . . . . .
- 2.2.1 Measuring the Quality of Fit . . . . .
- 2.2.2 The Bias-Variance Trade-Off . . . . .
- 2.2.3 The Classification Setting . . . . .
- 2.3 Lab: Introduction to R . . . . .
- 2.3.1 Basic Commands . . . . .
- 2.3.2 Graphics . . . . .
- 2.3.3 Indexing Data . . . . .
- 2.3.4 Loading Data . . . . .
- 2.3.5 Additional Graphical and Numerical Summaries

---

**Bibliografía**

**a) Básica:**

An Introduction to Statistical Learning with Applications in R.  
Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani

**b) Complementaria:**

-

---

**Modalidad cursada:** Presencial

---

**Metodología de enseñanza:** Exposiciones de los participantes y estudiantes

---

**Duración en semanas:** 15

---

**Carga horaria total:** 15

---

**Carga horaria detallada:**

**a) Horas aula de clases teóricas:** 0

**b) Horas aulas de clases prácticas:** 0

**c) Horas de seminarios:** 26

**d) Horas de talleres:** 0

**e) Horas de salida de campo:** 0

**f) Horas sugeridas de estudio domiciliario durante el período de clase:** 52

---

**Sistema de APROBACIÓN final**

**Tiene examen final:** No

**Se exonera el examen final:** Si

**Nota de exoneración (del 3 al 12):** Aprobado sin nota

**Sistema de GANANCIA**

**a) Características de las evaluaciones:**

Se trata de exposiciones del libro de referencia que deberán hacer los estudiantes.

**b) Porcentaje de asistencia requerido para ganar la unidad curricular: 80**

**c) Puntaje mínimo individual de cada evaluación y total: 50**

**d) Modo de devolución o corrección de pruebas:**

---

**Habilitada a rendir en calidad de examen libre: No**

---

**COMENTARIOS o ACLARACIONES:**

Se trata de un seminario introductorio de estadística con R, o como se está de moda decir ahora: `"Ciencia de datos"`.

Los temas se distribuyen entre los participantes al inicio y estos van realizando exposiciones en sucesivas

---