



# Procesamiento y Minería de Datos

¡Bienvenidos! En este curso aprenderás habilidades para analizar datos y obtener información valiosa.

Descubre cómo el procesamiento y la minería de datos son esenciales para la astronomía.

Profesor: Alex Vera Casanova.

Contacto: alxdaniel@gmail.com ; 19301136@ucen.cl

<https://github.com/AlxDaniel>

			Lunes	Martes	Miércoles	Jueves	Viernes
1	08:15-09:35	08:15-08:55					
		08:55-09:35					
2	09:45-11:05	09:45-10:25					
		10:25-11:05					
3	11:15-12:35	11:15-11:55					
		11:55-12:35					
4	12:45-14:05	12:45-13:25					
		13:25-14:05					
5	14:15-15:35	14:15-14:55		90029-PROCESAMIENTO DE DATOS / DATA MINING Grupo : 210 - TEORIA Aula : SALA-B412 Prof : ALEX DANIEL VERA CASANOVA			90029-PROCESAMIENTO DE DATOS / DATA MINING Grupo : 211 - EJERCICIO Aula : SALA-B412 Prof : ALEX DANIEL VERA CASANOVA
		14:55-15:35					
6	15:45-17:05	15:45-16:25					
		16:25-17:05					
7	17:15-18:35	17:15-17:55					
		17:55-18:35					
8	18:50-20:10	18:50-19:30					
		19:30-20:10					
9	20:15-21:35	20:15-20:55					
		20:55-21:35					
10	21:40-23:00	21:40-22:20					
		22:20-23:00					

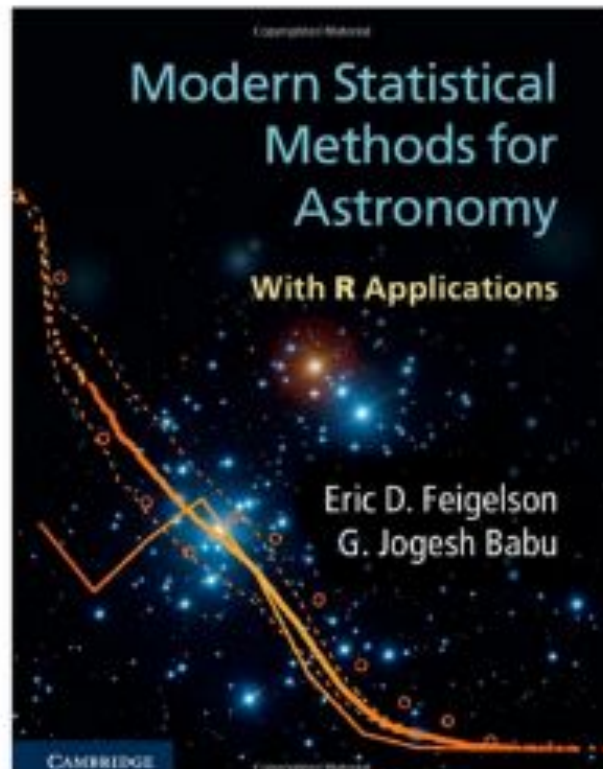
# Estructura de la Asignatura

Unidad <sup>2</sup>	Temas tratados	Resultados de aprendizaje desarrollados en la unidad.
I	Fundamentos de la astroinformática.	RA1 (cap I y II texto guía)
II	Data warehouses y surveys: Arquitecturas; implementaciones; aplicaciones en minería de datos. Ejemplos: Extracción de la información de Gaia, SDSS, Illustris TNG. Estructura de Bases de datos SQL	RA1
III	Algoritmos de minería de datos: Clustering, funciones de correlación, regresión y ajuste de modelos, clasificación, series de tiempo.	RA2
IV	Virtual Observatory: Recolección y estandarización de grandes bases de datos. Técnicas de visualización (ej: Topcat, Qfitsview)	RA3, RA4

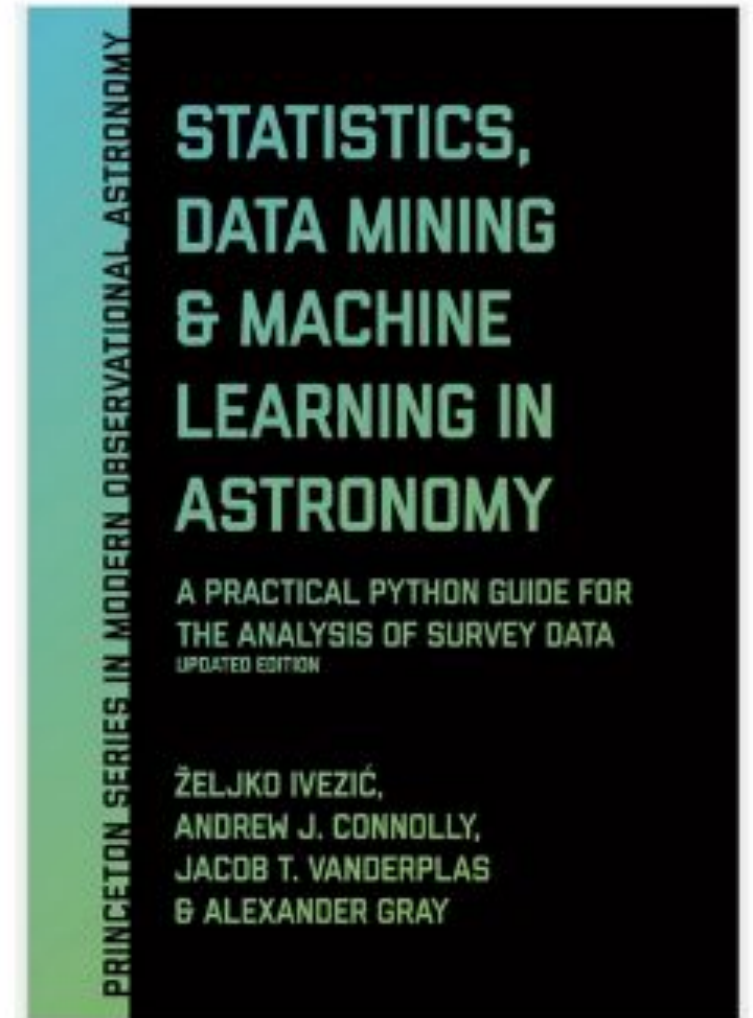


## Bibliografía:

- Statistics, Data Mining, and Machine Learning in Astronomy: A Practical Python Guide for the Analysis of Survey Data, Updated Edition

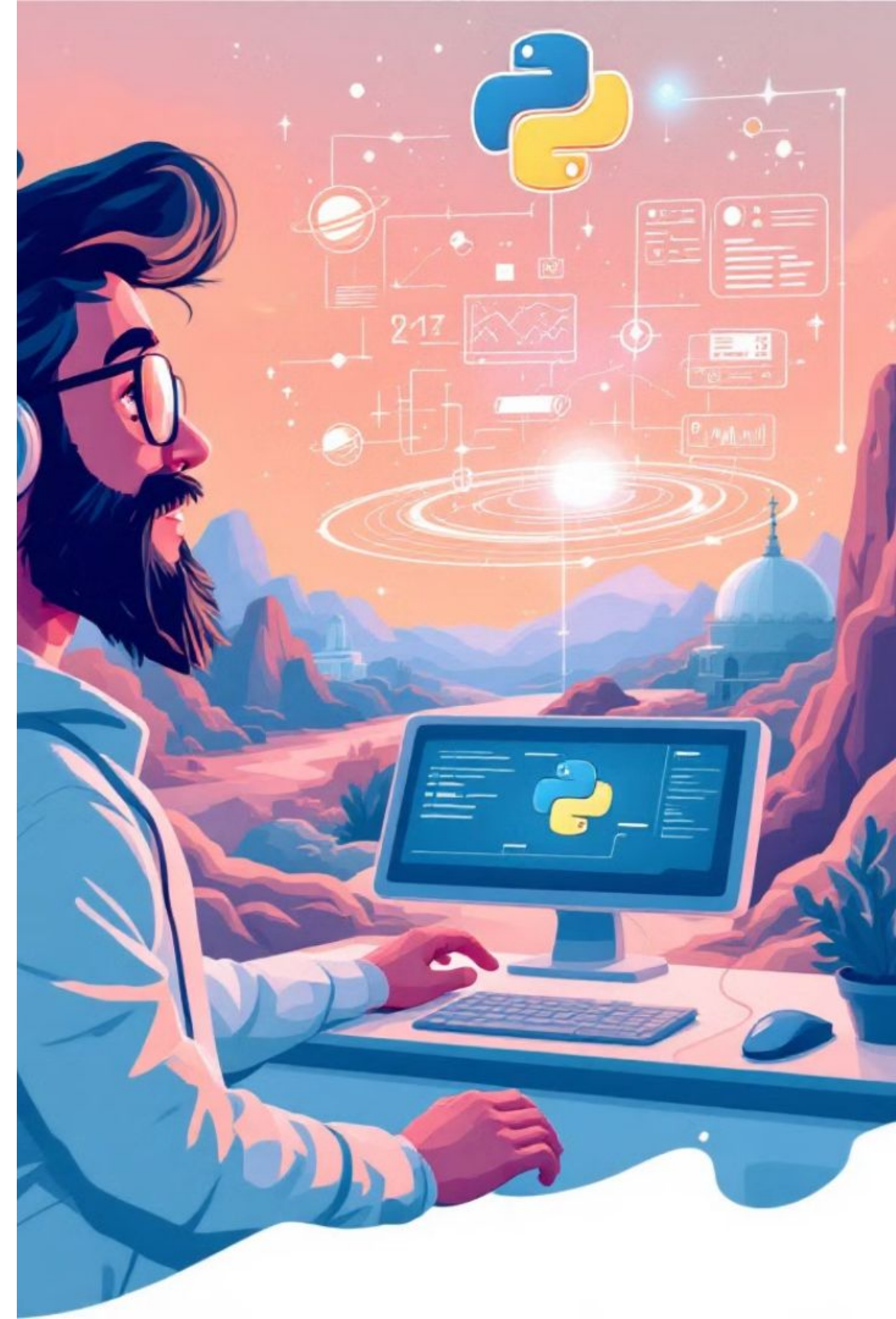


- Modern Statistical Methods for Astronomy: With R Applications



# Objetivos del curso

- Comprender los fundamentos del procesamiento de datos astronómicos.
- Aprender a utilizar herramientas como Python, Jupyter Lab, GitHub, y bibliotecas científicas.
- Usar catálogos y simulaciones astronómicas reales (GAIA, SDSS, Illustris-TNG).
- Aplicar técnicas de análisis de datos, visualización, clustering, clasificación y regresión.
- Uso y aplicación de Observatorios virtuales.



# Metodología y Evaluación

Clases teóricas y prácticas.

- Uso de Jupyter Lab y GitHub.
- **Evaluaciones:**

## Primera Prueba:

 **Martes 29 de abril (33.3%)**

## Segunda Prueba:

 **Martes 3 de junio (33.3%)**

## Tercera Prueba:

 **Martes 8 de julio (33.3%)**

Actividades prácticas:

- Tareas por unidad.
- Portafolio Final

## PAR

 **Martes 22 de julio (TBF)**

## EXAMEN

 **Martes 29 de julio (TBF)**





Número Semana	Tema a tratar	Actividad de aprendizaje
1. 01/04	I	T: Introducción al curso, desarrollo histórico, fundamentos de astro informática.
2. 04/04	I	T: Plataformas de compilación de códigos. P: Introducción a Python (JuPyter Lab) Y GitHub
3. 08/04	II	T: Warehouse y surveys P: Data mining (extracción de datos GAIA)
4. 11/04	II	T: Surveys: GAIA P: uso de datos.
5. 15/04	II	T: Surveys: SDSS P: Descarga y uso de datos.
6. 18/04	II	Feriado por Semana santa: No hay clases.
7. 22/04	II	T: Surveys: Illustris TNG P: Descarga y uso de datos.
8. 25/04	II	T: Surveys, Repaso y consultas.
9. 29/04		P: PRIMERA Prueba (33.3% Promedio final)
10. 01/05		Feriado
10. 06/05	III	T: Regresión y ajuste de modelos no lineal P1: Regresión y ajuste de modelos.
11. 09/05	III	T: Técnicas de visualización
12. 13/05	III	T: Uso de algoritmos de regresión y ajuste de modelos
13. 16/05	III	P: Ejercicios prácticos
14. 20/05	III	T: Clustering y funciones de correlación P: Aplicación Sobre datos.
15. 23/05	III	T: Clasificación de datos P: Aplicaciones de clasificación

16. 27/05	III	T: Análisis de series de tiempo P: Aplicación de análisis de series de tiempo
17. 30/05	III	T: Algoritmos, Repaso y consultas P: Repaso y consultas
18. 03/06	III	P: SEGUNDA Prueba (33.3% Promedio final)
19. 06/06	IV	T: Observatorio virtual (Parte 1)
20. 10/06	IV	T: Observatorio virtual (Parte 2)
21. 13/06	IV	T: TOPCAT: Tool for OPerations on Catalogues And Tables
22. 17/06	IV	P: Aplicaciones a la astronomía
24. 20/06 Vi:17		Feriado: no hay clases ni evaluaciones.
25. 24/06	IV	T: Preparación de portafolio
26. 27/06	IV	T: Creación de Catálogos
27. 01/07	IV	T: Machine learning P: Aplicación
28. 04/07	IV	T: OV, Repaso y consultas P: Repaso y consultas
29. 08/07	IV	P: TERCERA Prueba (33.3% Promedio final)
30. 11/07		T: Taller portafolios
31. 15/07		T: Charla Profesor invitado
32. 18/07		T: Clase de supervisión para preparación de PAR/examen
33. 23/07	PAR	PAR Martes 22/07
34. 29/07	Examen	Examen Martes 29/07

