

# Analítica de Datos

Syllabus 2025 - I

Pontificia Universidad Javeriana

## Clase

- 🕒 Lunes 4:10pm–6:45pm
- 👤 Carlos Cardona Andrade
- ✉ [carlos.cardonaa@javeriana.edu.co](mailto:carlos.cardonaa@javeriana.edu.co)
- [https://github.com/ccardonaandrade/analitica\\_2025](https://github.com/ccardonaandrade/analitica_2025)
- 📖 OpenIntro Statistics, 4<sup>th</sup> ed.
- 📖 Introduction to Modern Statistics, 2<sup>th</sup> ed.
- 📖 R for Data Science, 2<sup>th</sup> ed.
- 📖 Data Visualization: A practical introduction

## Resumen

**Descripción:** La gestión empresarial moderna exige una toma de decisiones basada en datos, superando las limitaciones de métodos puramente intuitivos o basados en experiencias. Este curso introduce la estadística como herramienta fundamental para transformar datos en decisiones empresariales efectivas. Los estudiantes desarrollarán competencias en:

- Recopilación y procesamiento sistemático de datos
- Análisis exploratorio y técnicas de inferencia estadística
- Modelización e interpretación de resultados
- Evaluación crítica de argumentos basados en datos

El curso enfatiza la aplicación práctica de la estadística en contextos empresariales, investigación científica y políticas públicas. Como señaló S. Wilks (1950), 'El pensamiento estadístico será una habilidad tan esencial para la ciudadanía eficiente como la capacidad de leer y escribir'

El aprendizaje de la programación estadística es inherente a la práctica del análisis de datos. Por ello, a lo largo de este curso también enseñaremos el lenguaje de programación estadística R.

**Prerrequisitos:** Este curso no asume ningún conocimiento previo en estadística o programación. Además, sólo se requiere aritmética básica.

## Recomendaciones

1. **Sean amables.**
2. **Asuman la responsabilidad** de su propia educación e intenten **aprender** tanto como puedan.
3. **Hagan su propio trabajo.**
4. Desarrollen su **intuición**—*por ejemplo, ¿por qué la regresión funciona en una situación y falla en otra?*
5. **Aprendan R.** Sean pacientes con ustedes mientras lo intentan—y utilicen **Google** para resolver sus dudas.
6. **Pidan ayuda temprano**—no esperen hasta el final del semestre.

## Software

**R:** Usaremos el lenguaje de programación estadística **R**, y utilizaremos **RStudio** para interactuar con R.

**Aprender R:** requerirá tiempo y esfuerzo, pero es una herramienta poderosa y versátil valorada por muchos empleadores. Dediquen el esfuerzo y tiempo necesarios, y serán recompensados. Los computadores en la universidad ya tienen instalado R, pero recomiendo enfáticamente que instalen R y RStudio en su propio computador.

Si están preocupados por aprender R—o quieren aprender más rápido—les sugiero que consulten los siguientes recursos gratuitos en línea.

- [Introduction to R](#) de DataCamp
- [R Bootcamp](#) de TeamLeada
- [Beginner's guide to R](#) de Computerworld

El equipo de RStudio también ha preparado un excelente [material de aprendizaje](#).

## Evaluación

Las notas para esta clase se asignarán en función de las siguientes tareas: tareas aproximadamente quincenales, dos exámenes parciales, y un proyecto final. La nota final se determinará según los siguientes pesos para cada una de las actividades:

<b>Quices y Talleres</b>	30%
<b>Parcial I</b>	15%
<b>Parcial II</b>	15%
<b>Proyecto Final</b>	40%

## Quices, talleres y parciales

**Talleres:** Este curso incluye varios talleres que son fundamentales para aprender el material y aprobar este curso. Cada taller incluye tanto instrucción general de estadística como consejos de programación necesarios para completar los ejercicios—vinculando el material de la clase con R—así como temas que pueden no ser cubiertos en la clase. Los talleres son la mejor manera de recibir retroalimentación rápida y ayuda en este curso.

- Deben **enviar los códigos de los talleres por correo electrónico**.
- Los talleres serán también corregidos con base en el estilo del código.
- Habrá talleres aproximadamente cada 2 semanas.
- Las soluciones **deben estar escritas con sus propias palabras. No copien**.
- Consulten abajo la **política de entrega tardía**.

Los talleres y el proyecto final se entregan en parejas. A menos que se indique explícitamente lo contrario, **cada grupo debe escribir y enviar respuestas de forma independiente**. Esto significa que no se aceptarán copias literales y serán consideradas como deshonestidad académica. En otras palabras: deben escribir las respuestas **con sus propias palabras. Copiar de otras personas o de talleres anteriores se considera trampa**.

### Política de entregas tardías

- Todos los talleres y proyectos tienen que ser entregados en las fechas estipuladas. Si no se entrega el taller o proyecto a tiempo, la calificación será de 0.
- Igualmente, si el estudiante no presenta un quiz, la nota del quiz es de 0.
- En caso de presentarse alguna emergencia, necesita sustentarla con evidencia apropiada.

### Exámenes

- El primer **examen parcial se realizará el 17 de marzo de 2025**.
- El segundo **examen final se realizará el 28 de abril de 2025**.

## Libros y otras lecturas

Este curso busca destacar la amplia gama de recursos educativos gratuitos disponibles en línea y fomentar una mentalidad proactiva y autodidacta, esencial para el aprendizaje efectivo de la programación. Todos los libros los encuentran gratis en línea. Estos textos ofrecen otra perspectiva complementaria sobre el material que cubrimos en las clases. El cronograma del curso (más abajo) incluye lecturas sugeridas para cada tema. Recomiendo que lean las lecturas asignadas previo a cada clase.

**Libros de Estadística:** Hay dos libros de estadística recomendados para este curso:

1. **OpenIntro Statistics, 4<sup>th</sup> ed.** de Diez, Cetinkaya-Rundel y Barr (**OIS**)
2. **Introduction to Modern Statistics, 2<sup>th</sup> ed.** de Cetinkaya-Rundel y Hardin (**IME**)

**Libros de R:** Para aprender R, seguiremos los siguientes textos que también son gratuitos:

1. **R for Data Science, 2<sup>th</sup> ed.** de Grolemund y Wickham (**R4DS**)
2. **Data Visualization: A practical introduction** de Healy (**DV**)

¿Quieren profundizar más? Consulten **Advanced R**.

## Cronograma tentativo del curso

La tabla a continuación presenta el plan actual para el esquema del curso y las lecturas asignadas de los libros de texto. Ocasionalmente, asignaremos artículos para que los lean para la clase, un quiz o los talleres. Les enviaré esos artículos por correo. Como sugiere el título de esta sección, el cronograma y los temas pueden estar sujetos a cambios.

### Cronograma tentativo

Clase	Fecha	Tema	Lectura sugerida
01	27/01	Intro & Preliminares	DV 2
02	03/02	ggplot & tidyverse	R4DS 1-3
03	10/02	Explorando datos numéricos	IME 5
04	17/02	Explorando datos categóricos & Distribución Normal	IME 4
05	24/02	Distribución muestral & Intervalos de Confianza	OIS 5.1 - 5.2
06	03/03	Pruebas de Hipótesis I	OIS 5.3
07	10/03	Pruebas de Hipótesis II	OIS 5.3
08	17/03	<b>Parcial I</b>	
	24/03	Festivo	
09	31/03	Muestreo y Encuestas	IME 2
10	07/04	Intro al Modelo de Regresión Lineal	IMO 7; OIS 8
	14/04	Semana Santa	
11	21/04	Más sobre Regresión Lineal	IMO 8; OIS 9.1-9.4
12	28/04	<b>Parcial II</b>	
13	05/05	Experimentos	
14	12/05	Regresión Logística	IMO 9; OIS 9.4
15	19/05	Análisis de Texto	
16	26/05	<b>Presentación Trabajo Final</b>	