

### 1. Información del equipo pedagógico y horarios de atención a estudiantes

**Profesor magistral:** Santiago Neira Hernández

**Correo electrónico:** s.neira10@uniandes.edu.co

**Horario de atención a estudiantes:** Martes 6:30p.m a 7:30p.m (Con cita previa)

**Lugar de atención a estudiantes:** <https://uniandes-edu-co.zoom.us/j/5493363657>

**Asistente de docencia:** Mariana Villabona Martínez

**Correo electrónico:** m.villabona@uniandes.edu.co

**Horario de atención a estudiantes:** A convenir con estudiantes, usar link.

**Lugar de atención a estudiantes:**

<https://us05web.zoom.us/j/3414359475?pwd=L3hKU2FDUHFsN1BLbGNzSUIyMGNRUT09>

ID de reunión: 341 435 9475

Código de acceso: GHM4Pn

**Enlace de la clase (Virtual):** <https://uniandes-edu-co.zoom.us/s/87120274271>

### 2. Introducción y descripción general del curso

En la era actual de la información los “datos” han sido definidos por muchas empresas como el nuevo petróleo. Así, la capacidad para extraer conocimientos significativos de conjuntos de datos es esencial para cualquier profesional y especialmente para los economistas. Este curso constituye una inmersión integral en el universo del análisis de datos mediante el empleo del lenguaje de programación R. Este lenguaje se ha consolidado como una herramienta líder en el campo del análisis de datos debido a su sintaxis intuitiva, su amplia variedad de paquetes y su sólido soporte en la comunidad científica. Además, de la capacidad de manipular y visualizar de forma agradable los resultados y procesamientos realizados.

A lo largo de este programa formativo, desde los conceptos primordiales hasta técnicas avanzadas, los participantes adquirirán las competencias necesarias para manipular datos, visualizar información de manera efectiva y llevar a cabo análisis profundos basados en evidencia. Estos corresponden a elementos cada vez más apetecidos por el mercado laboral. El curso contempla desde la introducción al entorno de RStudio, la programación en este lenguaje y la aplicación de modelos estadísticos a la solución de problemas económicos. Así, se espera que al finalizar el curso los estudiantes estén capacitados para realizar las tareas de Analista de Datos (Data Analyst) y desempeñen un conocimiento intermedio del lenguaje R.

### 3. Política de uso de la inteligencia artificial

Las herramientas de IA no son idóneas para realizar cualquier tipo de actividad y determinados tipos de uso pueden ser contraproducentes para el proceso pedagógico. En particular, los modelos largos del lenguaje (LLM) han revolucionado la experiencia de usuarios con inteligencia artificial desde el 2023. Los LLMs que pueden accederse a través de chatbots, tales como ChatGPT, Bard, Bing, y Claude, etc. democratizaron el acceso y uso de estos para tareas cotidianas. Entre las tareas más relevantes con posibilidad de uso se encuentran:

- Explorar nuevos asuntos y buscar inspiración, es decir, usar el chatbot en las etapas más tempranas de la investigación.
- Apoyar ejercicios que impliquen el reconocimiento y/o reproducción de patrones en textos, por ejemplo: traducir, resumir, modificar el tono y estilo del texto
- Resolver tareas de codificación o programación a través de la generación de texto compatible con los lenguajes de programación más usados.

Resulta ignorante omitir el potencial que estas herramientas pueden tener en la productividad de la solución de tareas de programación, no obstante, en etapas tempranas puede resultar antipedagógico y truncar el proceso de aprendizaje. Con esto en mente, en esta clase **se prohíbe el uso de LLM's para la generación de código o documentación**, salvo que explícitamente se permita en ciertos enunciados<sup>1</sup>.

#### 4. Resultados de aprendizaje

- Programar funciones que permitan solucionar problemas específicos a través del uso de variables, condicionales, loops y dataframes.
- Limpiar y manipular bases de datos haciendo uso de pipelines.
- Extraer información anárquica de fuentes no tradicionales y parametrizarla en datos limpios que puedan ser sujetos de análisis.
- Reconocer y aplicar las principales formas de visualización de datos según la naturaleza de estos y el ajuste de cada técnica a la estructura de los datos.
- Procesar bases de datos geoespaciales y generar mapas que caractericen la información geográfica junto con su simbología.
- Utilizar los principios del procesamiento del lenguaje natural para la limpieza y manipulación de caracteres.
- Aplicar los fundamentos de las buenas prácticas de programación para generar flujos de trabajo entendibles, replicables y fáciles de transmitir.
- Comunicar de manera efectiva los resultados encontrados tras los análisis, haciendo uso de los fundamentos del *storytelling* para construir narrativas que expliquen los datos o patrones.

---

<sup>1</sup> Se agradece a Juan David Gutiérrez por su colaboración en la formulación de esta política para el uso de inteligencia artificial en las aulas. Para más información sobre el uso ético de la misma, se recomienda consultar el siguiente enlace: [Lineamientos para el uso de inteligencia artificial en contextos universitarios \(versión 5.0\) – Foro Administración, Gestión y Política Pública \(forogpp.com\)](https://forogpp.com)

## 5. Cronograma

<b>Semana</b>	Semana 1 - Introducción al Curso
<b>Objetivos de la semana</b>	<ul style="list-style-type: none"> <li>• Introducción al curso</li> <li>• Comprender qué es R y cuáles son sus características como lenguaje de programación.</li> <li>• Entender qué son los Integrated Development Environments (IDEs).</li> <li>• Entender qué es R Studio.</li> <li>• Aprender a instalar R y R studio</li> <li>• Aprender a manipular los objetos básicos de R: tipos numéricos, conjuntos y booleanos.</li> <li>• Instalación y cargue de librerías.</li> </ul>
<b>Enero 20</b>	
<b>Semana</b>	Semana 2 / 3 – Objetos
<b>Objetivos de la semana</b>	<ul style="list-style-type: none"> <li>• Programar Listas.</li> <li>• Programar Vectores.</li> <li>• Programar Arrays</li> <li>• Realizar operaciones matriciales y vectoriales.</li> <li>• Acceder y extraer valores dentro de las matrices, vectores y array's</li> </ul>
<b>Enero 27 / Febrero 3</b>	
<b>Semana</b>	Semana 4 – Funciones, Control de flujo y buenas prácticas
<b>Objetivos de la semana</b>	<ul style="list-style-type: none"> <li>• Programar funciones</li> <li>• Programar condicionales: if, elif y else.</li> <li>• Programar loops (ciclos): for y while.</li> <li>• Conocer y aplicar los fundamentos de las buenas prácticas de programación y documentación.</li> </ul>
<b>Febrero 10</b>	
<b>Semana</b>	Semana 5 – Paquetes y Bases de datos I <b>TALLER 1</b>
<b>Objetivos de la semana</b>	<ul style="list-style-type: none"> <li>• Instalación y carga de paquetes</li> <li>• Identificar las características de los dataframes</li> <li>• Importar dataframes de fuentes externas</li> <li>• Inspección inicial de la información</li> <li>• Aplicar funciones a conjuntos de datos (apply, lapply, sapply y tapply)</li> </ul>
<b>Febrero 17</b>	
<b>Semana</b>	Semana 6 – Bases de datos II, EDA
<b>Objetivos de la semana</b>	<ul style="list-style-type: none"> <li>• Identificar las etapas de la limpieza de datos</li> <li>• Caracterizar los fundamentos de la exploración inicial de la base de datos (EDA)</li> <li>• Crear y alterar variables al interior de una base de datos.</li> </ul>
<b>Febrero 24</b>	

<b>Semana</b>	Semana 7 – Visualización básica I
<b>Objetivos de la semana</b>	<ul style="list-style-type: none"> <li>• Identificar los principios de la visualización de datos y del storytelling para reconocer patrones presentes en los datos.</li> <li>• Reconocer las principales formas de visualización de datos como gráficos de barras, series de tiempo, gráfico de torta, etc.</li> <li>• Reconocer qué visualización se adapta mejor al tipo de datos y problema a solucionar.</li> </ul>
<b>Marzo 3</b>	
<b>Semana</b>	Semana 8 – Visualización básica II <b>TALLER 2</b>
<b>Objetivos de la semana</b>	<ul style="list-style-type: none"> <li>• Reconocer las principales formas de visualización de datos como gráficos de barras, series de tiempo, gráfico de torta, etc.</li> <li>• Introducción al paquete ggplot2 para la programación de distintas visualizaciones de datos.</li> </ul>
<b>Marzo 10</b>	
<b>Semana</b>	Semana 9
<b>Objetivos de la semana</b>	<ul style="list-style-type: none"> <li>• Clase destinada al trabajo en grupos para la elaboración de la propuesta del trabajo final.</li> </ul>
<b>Marzo 24</b>	
<b>Semana</b>	Semana 10 – Visualización Avanzada
<b>Objetivos de la semana</b>	<ul style="list-style-type: none"> <li>• Reconocer las principales formas de visualización de datos como gráficos de barras, series de tiempo, gráfico de torta, etc.</li> <li>• Elaborar gráficos avanzados combinando ggplot2 + dplyr + tidyverse.</li> </ul>
<b>Marzo 31</b>	
<b>Semana</b>	Semana 11 – Datos geoespaciales
<b>Objetivos de la semana</b>	<ul style="list-style-type: none"> <li>• Definir los distintos tipos de “features” o características existentes en la composición de datos espaciales.</li> <li>• Cargar bases de datos espaciales al programa R a través de la librería “std”.</li> <li>• Manipular y extraer información haciendo uso de la librería “std”, como extraer la distancia entre features.</li> <li>• identificar las principales formas de visualización de datos espaciales, como la simbología, texto, entre otros.</li> </ul>
<b>Abril 7</b>	

<b>Semana</b>	Semana 12 – Procesamiento de lenguaje (NLP) <b>TALLER 3</b>
<b>Objetivos de la semana</b>	<ul style="list-style-type: none"> <li>• Identificar los principios del procesamiento del lenguaje natural.</li> <li>• Manipular variables en formato de texto para extraer información relevante.</li> </ul>
<b>Abril 21</b>	<ul style="list-style-type: none"> <li>• Visualizar información en formato de texto para reconocer patrones y tendencias en los datos.</li> <li>• Aprender pipeline de limpieza: eliminar stopwords, lematización, stemming, homogenización de formato.</li> <li>• Construir nubes de palabras</li> <li>• Expresiones regulares</li> </ul>

<b>Semana</b>	Semana 13 - Webscrapping
<b>Objetivos de la semana</b>	<ul style="list-style-type: none"> <li>• Extraer información de fuentes de datos digitales a través de Web Scrapping</li> <li>• Utilizar API's para la extracción de datos</li> <li>• Manejar la API de Open Street Maps (OSM)</li> </ul>
<b>Abril 28</b>	

<b>Semana</b>	Semana 14 – Intro a Econometría
<b>Objetivos de la semana</b>	<ul style="list-style-type: none"> <li>• Introducción a la econometría en R</li> <li>• Estimación de regresión por MCO</li> <li>• Exportar resultados de regresión</li> <li>• Metodologías de inferencia causal en R</li> </ul>
<b>Mayo 5</b>	

<b>Semana</b>	Semana 15 <b>TALLER 4</b>
<b>Objetivos de la semana</b>	<ul style="list-style-type: none"> <li>• Presentaciones trabajo final</li> </ul>
<b>Mayo 12</b>	

<b>Semana</b>	Semana 16
<b>Objetivos de la semana</b>	<ul style="list-style-type: none"> <li>• Presentaciones trabajo final</li> </ul>
<b>Mayo 19</b>	

## 6. Referencias

- Nussbaumer, C., 2015, *Storytelling with Data: A Data Visualization Guide for Business Professionals*
- Heiberger RM., Holland B., 2015, *Statistical Analysis and Data Display: An Intermediate Course with Examples in R*. Springer New York, New York, pp 377–426. [https://doi.org/10.1007/978-1-4939-21225\\_12](https://doi.org/10.1007/978-1-4939-21225_12)

- Christian Heumann., Michael Schomaker., 2022, *Introduction to Statistics and Data Analysis. With Exercises, Solutions and Applications in R*
- Matloff, N., 2011, *The Art of R Programming*
- Wickham, H., Çetinkaya-Rundel, M., Golemund, G., 2023, *R for Data Science: Import, Tidy, Transform, Visualize, and Model Data*, O'Reilly Media.
- Silge, J., Robinson, D., 2017, *Text Mining with R: A Tidy Approach*, O'Reilly Medi

## 7. Metodología

El curso se desarrolla a través de clases magistrales durante el horario de 6:00pm a 7:20pm. En este espacio se presentarán los temas la primera mitad de la clase y se destinará la segunda mitad de la clase a un desarrollo guiado de *scripts* que refuercen la aplicación de los mismos. La asistencia a clase no es obligatoria, pero es muy recomendable para la comprensión de todos los temas. Cuando hay inasistencias, los estudiantes son responsables de enterarse de todo lo que se diga durante la clase, aunque esto no se encuentre en las diapositivas o en el programa del curso. **Las clases serán grabadas** y estarán disponibles en línea dentro de 24 horas tras la finalización de la clase. Estas se consolidarán en un repositorio dentro de Bloque Neón.

Adicionalmente, los estudiantes cuentan con dos (2) horarios de atención para la resolución de dudas y preguntas referentes al curso.

## 8. Evaluaciones

- **Talleres prácticos (20% Cada Uno):**

Se asignarán 3 talleres prácticos en el semestre donde se evidencie que los estudiantes logran realizar las actividades asociadas a los resultados de aprendizaje correspondientes a los tópicos/semanas definidas. Así mismo, en aplicar los conocimientos y reforzar los elementos vistos. Un estudiante por pareja deberá entregar el taller a través de Bloque Neón a más tardar las 6:30pm del día correspondiente. **La entrega de talleres por fuera del horario establecido se penalizará de la siguiente forma:**

- Entregas entre 6:30 y 7:30p.m se califican sobre 4.5
- Entregas entre 7:30 y 8:30p.m se califican sobre 4.0
- Entregas entre 8:30 y 9:30p.m se califican sobre 3.0
- Entregas después de las 9:30p.m. tienen una calificación automática de 0.0

En cualquier caso que un horario de entrega sea anticipado, el esquema se aplica de la misma forma aplicando las diferencias horarias pertinentes.

- **Taller Inteligencia Artificial (10%):**

Se asignará 1 taller de uso de herramientas de inteligencia artificial, este se basa en la solución de un taller práctico para resolver un problema específico, pero que involucra el uso de herramientas de inteligencia artificial para su desarrollo. En particular, se busca comparar el desarrollo de los estudiantes contra el de herramientas de inteligencia artificial, con el fin de identificar formas más eficientes de solucionar el problema. Además, de enfrentarse a retos que no han sido abordados en el curso con ayuda de herramientas de IA. El esquema de calificación y sus penalizaciones serán la misma que los demás talleres prácticos

- **Documento del Trabajo Final (15%):**

Esta actividad se desarrollará en grupos de 5 estudiantes, este será desarrollado autónomamente desde la **primera semana** del curso. Los grupos deberán realizar una entrega **no evaluativa** correspondiente a la propuesta de trabajo final, al finalizar la semana 9 del curso. Esta propuesta será retroalimentada y permitirá a los estudiantes tener comentarios sobre la idea general y la conceptualización de la solución (los lineamientos están en Bloque Neón). Los grupos pueden elegir una (**y solo una**) de las siguientes opciones:

Trabajo práctico:

El trabajo corresponde al uso de análisis de datos para la solución de un problema de negocio real al que se enfrenta alguna empresa, corporación o entidad. Para esto, se exhorta a los estudiantes a investigar, preguntar y contactar empresas u organizaciones de carácter público con el fin de entender la estructura de negocio de estos y poder identificar y solucionar problemas que el cliente necesite. Este debe ser desarrollado haciendo uso de las herramientas aprendidas en el curso, usando únicamente el lenguaje de programación de R. Los lineamientos detallados con rúbrica de calificación se encuentran disponibles en Bloque Neón.

Trabajo de investigación:

Corresponde a un trabajo original de investigación donde se responda una pregunta de relevancia económica y se desarrolle un artículo con la estructura definida y de (posible) carácter publicable. Este debe ser desarrollado haciendo uso de las herramientas aprendidas en el curso, usando únicamente el lenguaje de programación de R. Los lineamientos detallados con rúbrica de calificación se encuentran disponibles en Bloque Neón.

- **Presentación del Trabajo Final (15%)**

Se desarrollarán presentaciones donde los grupos expongan sus trabajos finales y puedan recibir retroalimentación grupal.

La calificación de esta actividad se centra en el contenido, aunque la forma de comunicar los resultados es relevante. Igualmente, se busca identificar en esta presentación la capacidad de los grupos de:

1. Manipular bases de datos
2. Ofrecer estadísticas descriptivas que den cuenta del análisis y exploración de los datos.
3. Representar a través de visualizaciones tendencias o patrones los datos.
4. Responder a la pregunta problema a través del análisis de estos datos.
5. Utilizar alguno de los temas vistos avanzados (econometría, nlp, sistemas georeferenciales, webscrapping, u otro tema que el grupo decida)

La nota de la presentación del trabajo se divide en dos partes:

- A. Presentación en Plenaria (10%): De acuerdo con lineamientos que se les brindarán a lo largo del semestre, cada grupo deberá presentar sus hallazgos encontrados en su proyecto.
- B. Retroalimentación a grupo asignado (5%): Cada grupo le será asignada la evaluación de otro de sus compañeros. Dicha evaluación será cualitativa acerca de la calidad, hallazgos, metodología y capacidad de comunicación del trabajo.

El propósito pedagógico de este ejercicio recae en la introducción del sistema de *peer- review* de resultados. Se les evaluará su capacidad de evaluación desde la crítica constructiva y su capacidad de síntesis de los proyectos de los demás.

Las evaluaciones presentadas anteriormente contarán con el siguiente cronograma:

Actividad	Fecha de entrega	% nota	% nota acumulada
Taller 1	Domingo 23 de febrero	20%	20%
Taller 2	Domingo 16 de marzo	20%	40%
Propuesta Trabajo Final	Domingo 30 de marzo	N. A	N. A
Taller 3	Domingo 27 de abril	20%	60%
Presentaciones TF	Semanas asignadas	15%	75%
Taller 4 (IA)	Domingo 18 de Mayo	10%	85%
Entrega Trabajo Final (TF)	Jueves 30 de Mayo	15%	100%

### **Reclamos**

De acuerdo con los Artículos 62 y 63 del reglamento general de estudiantes de maestría, el estudiante tendrá **cuatro (4) días hábiles** tras conocer las calificaciones en cuestión para presentar un reclamo de forma escrita. El estudiante debe incluir un documento en el cual anexe una descripción del reclamo debidamente sustentado. El profesor magistral responderá al reclamo en los **cinco (5) días hábiles** siguientes. Si el estudiante considera que la respuesta no concuerda con los criterios de evaluación podrá solicitar un segundo calificador al Consejo de la Facultad en **los cuatro (4) días hábiles posteriores a la recepción de la decisión del profesor**.

### **Notas Finales**

Las calificaciones definitivas de las materias serán numéricas de uno punto cinco (1.5) a cinco punto cero (5.0), en unidades, décimas y centésimas. La nota final del curso será el promedio ponderado de las evaluaciones parciales según los pesos descritos anteriormente. No se hará ningún tipo de aproximación y la nota final se entregará en unidades, décimas y centésimas.

## **9. Asistencia**

No hay control de asistencia durante las clases, pero es sugerida la asistencia para poder discutir, entender, preguntar y debatir los conocimientos y conceptos enseñados. No obstante, la asistencia a actividades evaluativas es obligatoria. De acuerdo con el Artículo 45 del reglamento general de estudiantes de maestría, en caso de no asistir a las actividades evaluativas los estudiantes tendrán tres (3) días calendario para presentar una excusa válida para poder recibir calificación por estas actividades. Para revisar que se configura como una excusa válida, por favor remitirse al artículo 45.

## 10. Políticas generales de los cursos de Economía y fechas importantes

### Ajustes Razonables y Momentos Difíciles

Recientemente la Universidad desarrolló su política de **Ajustes Razonables y Momentos Difíciles**. La política de momentos difíciles es una propuesta de varios profesores/as que reconocen que hay diferentes **situaciones que impactan el proyecto académico** de las y los estudiantes, y que consideran que con su ayuda y flexibilidad se puede mitigar ese impacto. Análogamente, el concepto de ajustes razonables surge de la Convención sobre los Derechos de las Personas con Discapacidad de las Naciones Unidas y se estructura a partir del enfoque social de entender la **discapacidad**. En el siguiente [enlace](#) se encuentra más información sobre la política. En caso de identificar la necesidad de esta, es responsabilidad del estudiante contactar al profesor oportunamente.

### Cláusula de respeto por la diversidad

Todos debemos respetar los derechos de quienes integran esta comunidad académica. Consideramos **inaceptable cualquier situación de acoso, acoso sexual, discriminación, matoneo, o amenaza**. Cualquier persona que se sienta víctima de estas conductas puede denunciar su ocurrencia y buscar orientación o apoyo ante alguna de las siguientes instancias: el equipo pedagógico del curso, la Coordinación o la Dirección del programa, la Decanatura de Estudiantes, la Ombudsperson o el Comité MAAD.

### Temas Varios

Por otra parte, en [este enlace](#) se encuentran temas adicionales de conocimiento común como las reglas sobre asistencia a clase, excusas válidas, fraude académico y faltas disciplinarias, reclamos, políticas de bienestar y fechas importantes del semestre.