

PROGRAMA

I.- DATOS GENERALES			
Nombre del curso:		Introducción a Ciencia de Datos	
Código del curso:	80460	Clasificación Asignatura:	
Número de Unidades Crédito:	4	Horas de acompañamiento docente (Teoría):	1
Coordinación Académica:	Economía (ECON)	Horas de acompañamiento docente (Práctica):	1
Escuela:	Escuela de Economía (ECON)	Horas de acompañamiento docente (Laboratorio):	1
Facultad	Ciencias Económicas y Sociales	Horas de Preparaduría:	0
Tipo de Evaluación:	Evaluación Continua con reparación	Horas de trabajo independiente recomendado al estudiante:	4
Modalidad :	Presencial	Pre-requisitos:	Estadística II
Tipo de Asignatura:	Electiva	Régimen de Estudios:	Semestral
Ubicación de la asignatura:		Fecha de Aprobación Consejo de Facultad:	

II.- JUSTIFICACIÓN

La demanda de profesionales calificados en ciencia de datos en la industria, la academia y el gobierno crece rápidamente en un mundo en que cada vez se generan más datos y depende más de tecnologías de la información. Es necesario que los estudiantes aprendan a manejar herramientas aplicadas a problemas del mundo real para formarse como profesionales completos y competitivos.

El programa cubre conceptos y habilidades que permitan abordar desafíos de análisis de datos en un contexto aplicado. Se pretende abarcar conceptos de probabilidad, inferencia estadística, regresión lineal, y *machine learning* utilizando el lenguaje de programación R. También ayudará a desarrollar habilidades de organización y manejo de proyectos con herramientas tecnológicas avanzadas.

CONTRIBUCIÓN DE LA ASIGNATURA AL DESARROLLO DE LAS COMPETENCIAS

Competencias generales

CG1: Aprender a aprender con calidad:

Utiliza estrategias de forma autónoma para incorporar e incrementar conocimientos, habilidades y destrezas en el contexto de los avances científicos y culturales requeridos para un ejercicio profesional globalmente competitivo.

Unidad de Competencia

Criterios de desempeño

(CG1 – U1) Abstrae, analiza y sintetiza información

Descompone en partes, identifica factores comunes y resume lo realizado, en situaciones que requieran manejar información.

1. Identifica elementos comunes en diferentes situaciones o contextos.
2. Descompone, identifica, clasifica y jerarquiza elementos comunes.
3. Resume información de forma clara y ordenada.
4. Integra los elementos de forma coherente.
5. Valora críticamente la información

(CG1 – U2) Aplica los conocimientos en la práctica

Emplea conceptos, principios, procedimientos, actitudes y valores para plantear y resolver problemas en situaciones habituales, académicas, sociales y laborales

1. Selecciona la información que resulta relevante para resolver una situación
2. Elabora una síntesis para sí mismo o para comunicarla a otras personas
3. Establece y evalúa la eficacia y la eficiencia de los cursos de acción a seguir de acuerdo con la información disponible.
4. Implementa el proceso a seguir para alcanzar los objetivos mediante acciones, recursos y tiempo disponible
5. Evalúa los resultados obtenidos.

CG3: Aprender a trabajar con el otro:

Interactúa con otros en situaciones diversas y complejas para alcanzar objetivos comunes, en un entorno donde el equilibrio de los roles: colaborador o líder y la fluidez comunicativa procuran resultados beneficiosos para todos

(CG3 – U1) Participa y trabaja en equipo

Se integra en equipos asumiendo diversidad de roles y tareas, orientado hacia el logro de una meta común

1. Identifica roles y funciones de todos los miembros del equipo
2. Realiza las tareas establecidas por el equipo
3. Cumple diversos roles dentro del equipo
4. Utiliza formas de comunicación que favorecen las relaciones de interdependencia.
5. Coordina las acciones del equipo hacia el logro de la meta

común	
CG4: Aprende a interactuar en el contexto global: Actúa y se integra a los escenarios globales mediante el dominio de otros idiomas y de las tecnologías de la información y comunicación, esenciales para su interacción en el escenario global.	
(CG4 – U1) Maneja adecuadamente las tecnologías de información y comunicación Utiliza con destreza equipos y aplicaciones tecnológicas de comunicación e información para interactuar con otros en el contexto global	<ol style="list-style-type: none"> 1. Emplea recursos de internet como herramienta comunicacional 2. Gestiona adecuadamente los programas y aplicaciones de uso frecuente 3. Valida la información que consulta en Internet 4. Se actualiza permanentemente en las tecnologías de información y comunicación y en las medidas de seguridad y protección de estos sistemas 5. Interactúa en grupos de trabajo empleando las tecnologías de información y comunicación
Competencia Profesional: CP1: Analiza modelos que simplifican las relaciones económicas. Representa de manera simplificada la realidad económica.	
Unidad de Competencia	Criterios de desempeño
(CP1 – U1) Determina las relaciones entre las variables económicas. Evalúa el comportamiento histórico de las variables. Modela las relaciones entre las variables económicas mediante técnicas matemáticas y estadísticas.	<ol style="list-style-type: none"> 1. Analiza el impacto de una variable sobre el modelo
(CP1 – U2) Proyecta el comportamiento de las variables económicas Determina mediante el empleo de escenarios el comportamiento futuro de las variables.	<ol style="list-style-type: none"> 1. Recopila y analiza la información cuantitativa y cualitativa. 2. Identifica escenarios posibles para las variables económicas. 3. Simula estadística y matemáticamente el comportamiento de las variables en cada escenario.

IV.- UNIDADES TEMÁTICAS	
UNIDAD I Introducción a R	Contenidos <ul style="list-style-type: none"> • R como lenguaje de programación: ventajas, alcance y limitaciones • R: estructura, interfaz RStudio, paquetes • Conceptos básicos de programación • Importar datos • Funciones destinadas a ciencia de datos: <i>the tidyverse</i>
UNIDAD II Manejo de datos con R	Contenidos <ul style="list-style-type: none"> • Manejo de diferentes tipos de datos: cadenas de caracteres (<i>strings</i>), fechas y horas • Manipulación de bases de datos y tablas • Organizar y resumir información contenida en bases de datos • Cambiar formato de bases de datos: <i>data panel</i>, formato <i>wide</i>, formato <i>long</i> • Unión de bases de datos
UNIDAD III Visualización de Datos	Contenidos <ul style="list-style-type: none"> • El paquete <i>ggplot2</i> • Componentes de un gráfico: objetos tipo <i>ggplot</i>, escalas, etiquetas • Visualizar distribuciones de variables • Principios de visualización de datos: presentación de la información, comparación, no distorsionar valores, valores faltantes • Visualización de datos en la práctica • Resúmenes de robustos y completos de la información: valores atípicos, criterios para determinar valores

	atípicos, medidas de tendencia central.
UNIDAD IV Estadísticas con R	Contenidos <ul style="list-style-type: none"> • Valores faltantes • Probabilidad • Variables aleatorias • Inferencia estadística • Modelos estadísticos • Regresiones • Modelos lineales
UNIDAD V Machine Learning	Contenidos <ul style="list-style-type: none"> • Introducción al <i>machine learning</i> • Algoritmos de <i>machine learning</i> • Validación cruzada • Herramientas de <i>machine learning</i>: introducción a plataforma H2O • <i>Machine learning</i> en la práctica
UNIDAD VI Herramientas de productividad	Contenidos <ul style="list-style-type: none"> • Manejo y organización de proyectos con RStudio • GitHub: control de versiones y organización de proyectos colaborativos

V.- ESTRATEGIAS DE ENSEÑANZA Y APRENDIZAJE

A continuación, se presentan estrategias generales sugeridas. El profesor de la cátedra puede proponer y desarrollar diferentes estrategias en el aula siempre en procura al desarrollo de las competencias relacionadas con esta materia.

Las estrategias sugeridas están basadas en las recomendaciones de la Unidad de Innovación y Desarrollo Académico (UNIDEA), las cuales pretenden contribuir al desarrollo de las competencias tanto generales, como profesionales, relacionadas con la asignatura Introducción a Ciencia de Datos.

Estas estrategias van sujetas al número de alumnos por curso y otros factores que puedan incidir en su efectividad, además estarán en permanente revisión y actualización según los requerimientos de la cátedra y el éxito de las mismas.

Exposición del profesor: Explicación oral de conceptos, teorías o principios relacionados con un tema.

Presentaciones en Power Point: Material que con ayuda del video-beam en el aula; permite a los alumnos una revisión estructurada de los temas tratados, ya que permite visualizar palabras y contenidos clave en el desarrollo del tema que se trate, además refuerza la oratoria del profesor con un contenido presentado de una forma visual.

Mapas mentales: Herramienta que permite desarrollar un tema jerárquicamente partiendo de un macro-concepto o concepto principal, ayuda a memorizar información clave de forma lógica. Puede desarrollarse tomando notas sobre un tema, expresando las ideas planteadas de forma cartográfica.

Consultas en fuentes digitales: Búsqueda de información sobre algún tema desarrollado en clase, con la intención de reforzar lo visto, estableciendo ciertos parámetros del profesor.

Aprendizaje con dispositivos móviles: Provechando el actual uso de estos medios móviles (tablets, smartphones, laptops) se pretende usar este tipo de medios como estrategia para la enseñanza – aprendizaje, como una modalidad de m-learning (móvil-learning o aprendizaje móvil).

Videos: En ocasiones, el uso de videos puede reforzar el contenido visto en clase de una forma más interactiva, controlada en tiempo y espacio por el alumno, pudiendo acceder a la información en el lugar y momento más conveniente del día o la noche.

VI.- ESTRATEGIAS DE EVALUACIÓN

Evaluación Formativa:

Resolución de problemas prácticos y guías de ejercicios propuestas por el profesor y revisadas en el aula.

Pruebas cortas al finalizar un tema.

Cada clase constará de actividades evaluadas de los contenidos vistos por sesión, balanceando la teoría expuesta con su aplicación inmediata.

Proyecto final en equipos en donde el estudiante deba aplicar todos los conocimientos aprendidos durante el curso.

VII.- REFERENCIAS BIBLIOGRÁFICAS

(Irizarry, 2019) Rafael A. Irizarry, "Introduction to Data Science: Data Analysis and Prediction Algorithms with R".
(O'Neil 2017) Catherine O'Neil, "Weapons of Math Destruction". 2017.
(Peng, 2020) Roger D. Peng, "R Programming for Data Science".