

PROGRAMA DE UNIDAD CURRICULAR

I.- DATOS GENERALES

Nombre de la Unidad Curricular: Introducción a la Ciencia de Datos / Introduction to Data Science

Código de la Unidad Curricular: 80460

Unidad de Gestión Académica: Escuela de Economía (ECON)

Nivel: Pregrado

Régimen: Semestral

Número de Unidades Crédito: 4

Taxonomía: TA-8

Tipo de Evaluación:

Evaluación Continua

Evaluación Continua con Reparación

X

Evaluación Final y Reparación

N° horas semanales trabajo independiente (HTI)

4

N° horas semanales de acompañamiento docente (HAD)

Teóricas

1

Seminario

Prácticas

1

Laboratorio

1

Investigación

Categoría de la Unidad Curricular:

Intra-Facultad

X

Inter-Facultad

Institucional

Escuela / Programa

Modalidad:

Presencial

Semipresencial

Virtual

En línea

X

Instancia aprobatoria:

Fecha de aprobación :

II.-Resumen/Abstract (en español e inglés)

La demanda de profesionales capacitados en ciencia de datos en la industria, la academia y el gobierno está creciendo rápidamente en un mundo cada vez más impulsado por los datos y más dependiente de las tecnologías de la información. Es esencial que los estudiantes aprendan a utilizar herramientas aplicadas a problemas del mundo real para convertirse en profesionales integrales y competitivos.

El programa cubre conceptos y habilidades que permiten a los estudiantes abordar los desafíos del análisis de datos en un contexto aplicado. Se pretende cubrir conceptos de probabilidad, inferencia estadística, regresión lineal y aprendizaje automático utilizando el lenguaje de programación R. También ayudará a

desarrollar habilidades organizacionales y de gestión de proyectos con herramientas tecnológicas avanzadas.

The demand for skilled data science professionals in industry, academy and government is growing quickly in a world that is increasingly data-driven and more reliant on information technologies. It is essential for students to learn to use tools applied to real-world problems to become integral and competitive professionals.

The program covers concepts and skills that enable students to address data analysis challenges in an applied context. It is intended to cover concepts of probability, statistical inference, linear regression, and machine learning using the R programming language. It will also help develop organizational and project management skills with advanced technological tools.

III.- CONTRIBUCION DE LA ASIGNATURA AL DESARROLLO DE LAS COMPETENCIAS (*enumerar y enunciar las competencias generales (competencias comunes, en el caso de postgrado) con sus respectivas unidades y criterios de desempeño, así como las competencias profesionales (básicas y específicas, según el plan de estudio) o competencias específicas (en el caso de postgrado) con sus unidades y criterios de desempeño.*)

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.

(CG1 – U1) Abstrae, analiza y sintetiza información

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

Compara los modelos macro y microeconómicos que explican la realidad económica.

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, equipo 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

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.

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.

(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.

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.

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 *(Colocar las unidades temáticas y temas de la asignatura. Recuerde que las unidades temáticas compilan los temas de la misma)*

UNIT I: I Introduction to R

R as a programming language: advantages, scope and limitations

R: Structure, RStudio interface, packages.

Basic programming concepts.

Import data.

Functions aimed at data science: *the tidyverse*.

UNIT II: Data management with R

Handling different types of data: strings, dates and times.

Manipulation of databases and tables.

Organize and summarize information contained in databases.

Change database format: data panel, wide format, long format.

Database joining.

UNIT III: Data Visualization

The ggplot2 package

Components of a graph: ggplot type objects, scales, labels

Visualize variable distributions

Principles of data visualization: presentation of information, comparison, not distorting values,

Missing values

UNIT IV: statistics with R

Missing values

Probability

Random variables

Statistical inference

Statistical models

Regressions

Linear models

UNIT V: Machine Learning

Introduction to machine learning

Machine learning algorithm

Cross-validation

Machine learning tools: introduction to H2O platform

Machine learning in practice

UNIT VI: Productivity tools

Project management and organization with RStudio

GitHub: version control and project organization collaborative project

V.- ESTRATEGIAS DE ENSEÑANZA Y DE APRENDIZAJE

- The following are general suggested strategies. The teacher can propose and develop

different strategies in the classroom in order to develop the competencies related to this subject. The suggested strategies are based on the recommendations of the UNIDEA: Innovation and Academic Development Unit, which aim to contribute to the development of both general and professional competencies related to the subject Introduction to Data Science.

- These strategies are based on the number of students per course and other factors that may affect their effectiveness, and will be constantly reviewed and updated according to the requirements of the course and its success. Teacher's presentation: Oral explanation of concepts, theories or principles related to a topic. Power Point Presentations: Material exposed through the video-beam in the classroom that will allow students to get a structured review of the topics covered, as it enables them to visualize key words and content in the development of the topic in question, also reinforces the oratory of the teacher with content presented in a visual way. Mind maps: A tool that allows developing a topic hierarchically starting from a macro-concept or main concept, it helps to memorize key information in a logical way. It can be developed by taking notes on a topic, expressing the ideas raised in a cartographic form.
- Digital sources: Search for information on a topic developed in class, with the intention of reinforcing what has been seen, establishing certain parameters of the teacher. Learning with mobile devices:
- Taking advantage of the current use of these mobile media (tablets, smartphones, laptops), we intend to use this type of media as a teaching-learning strategy, as a modality of m-learning (mobile-learning or mobile learning). Videos: Sometimes, the use of videos can reinforce the content seen in class in a more interactive way, controlled in time and space by the student, being able to access the information at the most convenient place and time of the day or night.

VI.- ESTRATEGIA DE EVALUACIÓN

- Formative Evaluation: Resolution of practical problems and exercise guides proposed by the teacher and reviewed in the classroom. Short tests at the end of a topic. Each class will consist of evaluated activities of the contents seen per session, balancing the theory presented with its immediate application. Final project in teams where the student must apply all the knowledge learned during the course.

VII.- REFERENCIAS PRINCIPALES

- 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”.

REFERENCIAS COMPLEMENTARIAS