

Información General

- _ Profesora: Patricia Ordóñez
- _ Salón: CN A-143
- _ Horario: martes y jueves de 4 – 5:30 pm
- _ Email: patricia.ordonez@upr.edu (favor de indicar CCOM 4995 en el sujeto)
- _ Web personal: ccom.uprrp.edu/~pordonez
- _ Oficina: A-147
- _ Horas de Oficina: martes 2-4pm y jueves de 3-4pm y 5:30-6:30pm, o por cita
- _ Teléfono: 764-0000 x 88352
- _ Créditos: 3 horas

Descripción del curso

Data Science es un curso de programación introductorio, aplicado, colaborativo y a ritmo propio que reúne a estudiantes de diferentes disciplinas, incluida la biología y otras ciencias naturales, estadísticas, informática e incluso humanidades, para trabajar en proyectos locales y culturalmente relevantes utilizando datos. ciencia. Los requisitos previos para el curso incluyen el curso Introducción a la estadística en R y el curso Introducción a la informática utilizando Python para que todos los alumnos comprendan los conceptos básicos del pensamiento computacional, la programación en R y Python, las estadísticas y el método científico. La estructura del curso es un aula invertida donde las conferencias se realizan en línea a través de DataCamp, un marco europeo de educación en ciencia de datos. Se espera que los estudiantes vayan a su propio ritmo y completen la pista de certificación de Data Analyst o la pista de certificación de Data Science en función de su nivel de habilidad de programación y experiencia en el dominio. En clase, los estudiantes trabajan en 4 proyectos grupales en equipos interdisciplinarios donde cada miembro del equipo tiene experiencias variadas en R, Python, Estadísticas y ciencias naturales. Se espera que los estudiantes aprendan cómo usar la línea de comandos, git, GitHub y Jupyter notebooks para publicar investigaciones reproducibles y estar preparados para internados de verano en ciencias de datos.

English:

Data Science is an introductory, applied, collaborative, self-paced programming course that brings students from different disciplines together, including biology and other natural sciences, statistics, computer science, and even the humanities, to work on local, culturally relevant projects using data science. Prerequisites for the course include the Introduction to Statistics in R course and the Introduction to Computer Science course using Python so that all students understand the basics of computational thinking, programming in R and Python, statistics, and the scientific method. The structure of the course is a flipped classroom where the lectures occur online through DataCamp, a European data science education framework. Students are expected to go at their own pace and complete either the Data Analyst certification track or the Data Science certification track depending on their programming skill level and domain expertise. In class, the students work on 4 group projects in interdisciplinary teams where each team member has varied experiences in R, Python, Statistics and a natural science. Students are expected to learn how to use the command line, git, GitHub, and Jupyter notebooks to publish reproducible research and be prepared for summer internships in data science.

Calendario Tentativo

El calendario actual para es curso se puede encontrar en el siguiente enlace:

<https://tinyurl.com/ccom4995-033spr2019>

Semana	Data Analyst Track	Data Scientist Track
Jan. 14	Introduction to Python	Intro to Python & Intermediate Python for DS
Jan 21	Intermediate Python for Data Science	Python Data Science Toolbox (Part I & II)
Jan. 28	Python Data Science Toolbox (Part I)	Importing Data in Python (Part I & II)
Feb. 4	Python Data Science Toolbox (Part II)	Cleaning Data in Python

Feb. 11	Importing Data in Python (Part I)	pandas Foundations
Feb. 18	Cleaning Data in Python	Manipulating DataFrames with pandas
Feb. 25	pandas Foundations	Merging DataFrames with pandas
Mar. 4	Manipulating DataFrames with pandas	Intro to SQL for Data Science
Mar. 11	Merging DataFrames with pandas	Introduction to Databases in Python
Mar. 18	Introduction to Databases in Python	Introduction to Data Visualization with Python & Interactive Data Visualization with Bokeh
Mar. 25	Introduction to Data Visualization with Python	Statistical Thinking in Python (Part I & II)
Apr. 1	Statistical Thinking in Python (Part I)	Joining Data in SQL
Apr. 8	Statistical Thinking in Python (Part II)	Supervised Learning with Scikit-learn
Apr. 15		Machine Learning with the Experts: School Budgets
Apr. 22		Unsupervised learning in Python
Apr. 29		Deep Learning in Python
May 6		Network Analysis in Python (Part I)
May 13	Final Project and Presentation	
May 21	Delivery of Final Grade	

Recursos de Aprendizaje Requeridos

Esta clase es apoyada por DataCamp, la plataforma de aprendizaje más intuitiva para la ciencia de datos. Aprenda R, Python y SQL de la manera que mejor aprende a través de una combinación de videos cortos de expertos y ejercicios prácticos con el teclado. Los estudiantes pueden tomar más de 100 cursos impartidos por instructores expertos en temas como la importación de datos, la visualización de datos o el aprendizaje automático y aprenda más rápido a través de comentarios inmediatos y personalizados en cada ejercicio por seis meses.

(<https://www.datacamp.com/>)

Destrezas en el uso de computadoras

Acceso a computadora con conexión rápida de Internet

Cuenta de correo electrónico institucional (upr.edu)

Cuenta en la plataforma de Moodle

Cuenta en la plataforma de Piazza

Programas de Microsoft Office o equivalentes (Word, Excel, PowerPoint)

Programa de Adobe Acrobat o equivalente

Textos recomendados:

Data Analysis for the Life Sciences by Rafael Irizarry and Michael I. Love

<https://leanpub.com/dataanalysisforthelifesciences>

Objetivos del Curso

Al finalizar del curso el estudiante podrá:

Definir el campo de la ciencia de datos

Trabajar en equipo interdisciplinario para resolver un problema

Aplicar análisis exploratorio de datos de alta dimensión o volumen

Utilizar librerías de Python para extraer data de bases de datos

Utilizar librerías de Python para analizar datos

Utilizar métodos de *machine learning* para extraer información de datos

Técnicas Instruccionales

Laboratorios

Escritura

Programas

Proyectos
Exámenes
Asignaciones

En la sala de clases: Conferencias, discusiones, trabajos grupales, tareas.

Métodos alternos, no presenciales: Video conferencias, foros de discusión grupales, módulos instruccionales, charlas virtuales, tareas, pruebas cortas en línea.

Técnicas de Evaluación

Asignaciones: 15%, Laboratorios: 15%

Examen Parcial: 15%, Examen Final: 15%, Proyecto de Circuitos: 15%,

Proyecto Programación: 15% Participación: 10%

A continuación, se incluye la asignación de notas de acuerdo al porcentaje obtenido en clase:

90-100 -> A

80-89 -> B

70-79 -> C

60-69 -> D

0-59 -> F

* En caso de ser necesario los exámenes podrán ser en formato de proyectos o tareas o podrán ser ofrecidos en línea utilizando la Plataforma Moodle.

** Se harán acomodos razonables para estudiantes que necesitan más tiempo en los exámenes. Es importante que el estudiante se notifique con la profesora antes del día del examen para coordinarlo.

Modo de Comunicación

Este semestre usaremos Piazza para la discusión en clase. El sistema está altamente adaptado para obtener ayuda rápida y eficiente de tus compañeros y de mí. En lugar de enviar preguntas por correo electrónico al personal docente, le animo a que publique sus preguntas en Piazza. Si tiene algún problema o comentario para los desarrolladores de Piazza, envíe un correo electrónico a team@piazza.com.

Encuentra nuestra página de clase en: <https://piazza.com/uprrp/term12019/ccom4995033/home>

POLÍTICAS DEL CURSO

Honestidad Académica

La Universidad de Puerto Rico promueve los más altos estándares de integridad académica y científica. El Artículo 6.2 del Reglamento General de Estudiantes de la UPR (Certificación Núm. 13, 2009-2010, de la Junta de Síndicos) establece que “la deshonestidad académica incluye, pero no se limita a: acciones fraudulentas, la obtención de notas o grados académicos valiéndose de falsas o fraudulentas simulaciones, copiar total o parcialmente la labor académica de otra persona, plagiar total o parcialmente las respuestas de otra persona a las preguntas de un examen, haciendo o consiguiendo que otro tome en su nombre cualquier prueba o examen oral o escrito, así como la ayuda o facilitación para que otra persona incurra en la referida conducta.” Cualquiera de estas acciones estará sujeta a sanciones disciplinarias en conformidad con el procedimiento disciplinario establecido en el Reglamento General de Estudiantes de la UPR vigente.

* Cualquier violación de la política de integridad académica resultará en una multa mínima de cero en esa tarea, proyecto, o examen.

Métodos Alternos de Enseñanza

La Certificación Núm. 112 (2014-15) de la Junta de Gobierno define un curso presencial como un curso en el cual 75% o más de las horas de instrucción requieren la presencia física del estudiante y el profesor en el salón de clases. Esto quiere decir que 25% de un curso presencial pudiera ofrecerse sin requerir la presencia física de los estudiantes y el profesor en el salón de clases. En caso de ser necesario, este curso podrá completar hasta 25% de las horas contacto (11.25) de forma no presencial por métodos alternos como, por ejemplo: videoconferencias, módulos instruccionales, foros de discusión y cibercharlas entre otros. De ser así, se modificará el calendario/temario para incluir los temas que serán cubiertos por métodos alternos.

Derechos de Estudiantes con Impedimentos

La Universidad de Puerto Rico cumple con todas las leyes federales y estatales, y reglamentos concernientes a discriminación, incluyendo "The American Disabilities Act" (Ley ADA) y la Ley 51 del Estado Libre Asociado de Puerto Rico. Los estudiantes que reciban servicios de rehabilitación vocacional deben comunicarse con el profesor al inicio del semestre para planificar el acomodo razonable y equipo de asistencia necesario conforme con las recomendaciones de la Oficina de Asuntos para las Personas con Impedimentos (OAPI) del Decanato de Estudiantes.

Política Anti-discrimen y Hostigamiento de la UPR

La Universidad de Puerto Rico prohíbe el discrimen por razón de sexo y género en todas sus modalidades, incluyendo el hostigamiento sexual. Según la Política institucional contra el Hostigamiento Sexual en la Universidad de Puerto Rico, Certificación Núm. 130, 2014-2015 de la Junta de Gobierno, si un estudiante está siendo o fue afectado por conductas relacionadas a hostigamiento sexual, puede acudir ante la Oficina de la Procuraduría Estudiantil, el Decanato de Estudiantes o la Coordinadora de Cumplimiento con Título IX para orientación y/o presentar una queja.

En el salón:

1. Por favor apaguen el timbre (sonido) a los celulares antes de entrar a la clase.
2. Si tiene alguna duda sobre el código de ética institucional con relación al plagio por favor informarlo al profesor.
3. En este curso se fomenta la diversidad y el choque de ideas por lo cual la libre expresión va a ser defendida.

Referencias

1. <https://www.edx.org/course/high-dimensional-data-analysis-harvardx-ph525-4x-1>
2. <https://www.csee.umbc.edu/~oates/classes/16/data-science/home.html>
3. <http://amplab.github.io/datascience-sp14/>