



UNIVERSIDAD DE COSTA RICA. FACULTAD DE CIENCIAS ECONÓMICAS ESCUELA DE ESTADÍSTICA		
<b>Sigla y nombre del curso:</b> XS-2230. Estadística Computacional II	<b>Créditos:</b> 4	<b>Ciclo:</b> II-2022
<b>Requisitos:</b> XS-2310 ó XS-0121 XS-2210 ó XS-0220 XS-2110 ó XS-0211	<b>Correquisitos:</b> Ninguno	<b>Plataforma:</b> Mediación virtual. <b>Modalidad del curso:</b> Alto Virtual
<b>Grupo 1</b> <b>Docente:</b> Carlos Solís Fonseca <b>Horario de clases:</b> jueves de las 17:00 a las 20:50 <b>Aula:</b> Clases Virtuales por Zoom <b>Horario de consulta:</b> lunes de las 17:00 a las 18:50 <b>Correo electrónico:</b> <a href="mailto:carlos.solisfonseca@ucr.ac.cr">carlos.solisfonseca@ucr.ac.cr</a>	<b>Grupo 2</b> <b>Docente:</b> César Gamboa Sanabria <b>Horario de clases:</b> jueves de las 17:00 a las 20:50 <b>Aula:</b> Clases Virtuales por Zoom <b>Horario de consulta:</b> martes de las 17:00 a las 18:50 <b>Correo electrónico:</b> <a href="mailto:cesar.gamboasanabria@ucr.ac.cr">cesar.gamboasanabria@ucr.ac.cr</a>	

### DESCRIPCIÓN DEL CURSO

El curso de Estadística Computacional II busca introducir al estudiantado en los conceptos fundamentales de la creación de aplicaciones de entrada de datos mediante CPro, el uso de PSPP para el análisis estadístico de datos, los principios de programación en el lenguaje Python, y los conceptos básicos del manejo de bases de datos.

### OBJETIVO GENERAL

Al final del curso cada estudiante deberá tener los conocimientos y habrá desarrollado habilidades en el procesamiento estadístico de datos desde los procesos de diseño de las investigaciones, pasando por la recolección y depuración de datos hasta el análisis utilizando paquetes estadísticos.

### OBJETIVOS ESPECÍFICOS

Conocer las consideraciones computacionales en el diseño y obtención de datos de investigaciones estadísticas y aprender a utilizar el paquete para la captura y depuración de datos CPro.

Aprender a utilizar PSPP para elaborar estadísticas descriptivas, visualizaciones y manipulación de conjuntos de datos.

Aprender a utilizar Python para construir y manipular archivos, definir diccionarios de datos, elaborar estadísticas descriptivas, simulaciones, y aplicar otros métodos y técnicas relacionadas al quehacer estadístico.

Conocer los fundamentos básicos de la manipulación de bases de datos mediante SQL.





## METODOLOGÍA

Los contenidos del curso se desarrollarán mediante lecciones sincrónicas y asincrónicas, de manera que cada estudiante pueda repasar las presentaciones y/o videos desde cualquier dispositivo con acceso a internet. Cada profesor realizará actividades formativas para apoyar el proceso de aprendizaje, las cuales incluyen pruebas cortas y trabajos prácticos.

Para el estudio de los contenidos del curso el personal docente facilitará espacios para la puesta en práctica de los conocimientos adquiridos a lo largo del semestre, por lo que se propone una metodología de un aprendizaje activo, donde cada estudiante se compromete a realizar las actividades que su profesor proponga, así como asistir y participar en los espacios de trabajo facilitados. El curso está dividido en cuatro partes:

1. Desarrollar una aplicación de entrada de datos mediante CSPro para la correcta tabulación de la información.
2. Aplicar técnicas de análisis exploratorio, manipulación, y visualización de datos mediante PSPP.
3. Aplicar técnicas de análisis exploratorio, manipulación, simulación, y visualización de datos mediante el uso básico e intermedio del lenguaje Python.
4. Aplicar métodos de manipulación de bases de datos mediante SQL.

## EVALUACIÓN

Se realizarán pruebas cortas individuales y trabajos prácticos grupales. El objetivo de las pruebas cortas es apoyar el proceso de aprendizaje, estos no se avisarán y pueden cubrir cualquier tema visto en clase (incluidas las lecturas asignadas). Por su naturaleza, las pruebas cortas no se repiten (al no tener fechas preestablecidas).

Según el artículo 15 del Reglamento Académico Estudiantil es posible realizar pruebas cortas sin anunciar al estudiante la fecha con antelación, en la medida que se especifique en el programa del curso. Únicamente las evaluaciones estipuladas en el artículo 18 y que son anunciadas al estudiante son sujeto de reposición (o sea exámenes parciales y de ampliación).

La nota obtenida en cada trabajo práctico se divide en dos rubros:

1. Un 70% del valor de cada trabajo práctico será asignado por el profesor, que se compone de:
  - a. Dominio del contenido evaluado (se evalúa individualmente).
  - b. Documentación del código (se evalúa grupalmente).
  - c. Lógica y validez de los resultados obtenidos (se evalúa grupalmente).
  - d. Fuentes consultadas para resolver el trabajo (se evalúa grupalmente).
  - e. Eficiencia del código generado (se evalúa grupalmente).



2. Un 30% del valor de cada trabajo práctico será asignado a cada estudiante por los demás miembros de su grupo.

Si un estudiante faltase a alguna presentación de un trabajo práctico, debe presentar la justificación al profesor de su grupo, indicando las razones de la ausencia, acompañada de los documentos justificantes. La misma debe remitirse al profesor citado (por medio de correo electrónico institucional) a más tardar en cinco días hábiles del reintegro a lecciones. La causa debe estar contemplada dentro del Reglamento de Régimen Académico Estudiantil, aprobado por el Consejo Universitario en la sesión 4632 del 3 de mayo de 2001 (sólo el profesor del curso recibirá dicha justificación, no será recibida en la secretaría de la Escuela de Estadística). Se entenderá por causa justificada:

1. Enfermedad comprobada mediante dictamen médico y que imposibilite al estudiante participar en la presentación.
2. Choque en día y hora con otro examen dentro de la Universidad de Costa Rica siempre y cuando sea un examen de cátedra. El estudiante deberá presentar una constancia con la firma del profesor y sello de la Unidad Académica respectiva, donde se indique el horario donde el estudiante realizó el examen.
3. Otra causa grave (muerte de pariente en primer o segundo grado y causas fortuitas). Según el oficio OJ-1306-2008 de la Oficina Jurídica, se entenderá como causa fortuita: “acontecimientos que no han podido preverse y cuyas circunstancias deben ser irresistibles o inevitables, siendo impotente el hombre para impedir su ocurrencia”.

Al final del curso se calculará el promedio simple de los trabajos prácticos (eliminando la menor nota), a efectos de aplicar el porcentaje correspondiente de la nota final del curso. La fecha de entrega de los trabajos prácticos se indica en este programa. **Después de la fecha de entrega de los trabajos prácticos, los trabajos enviados por los estudiantes tendrán nota de cero.**

EVALUACIÓN	PONDERACIÓN
Trabajos prácticos	70%
Pruebas cortas	30%

El examen de **AMPLIACIÓN** que incluye toda la materia del curso se realizará el 2 de diciembre del 2022 a las 17:00 horas, se aplicará a todos aquellos estudiantes cuya nota final sea 6,0 o 6,5. El estudiante que obtenga 7,0 o más en este examen aprobará el curso con nota de 7,0.

Entrega de resultados finales: jueves 24 de noviembre del 2022.

NOTA: El examen de ampliación se realizará el 2 de diciembre del 2022 a las 17:00 horas en la Plataforma Mediación Virtual de la Universidad de Costa Rica.





## CONTENIDOS DEL CURSO

### 1. CPro

- 1.1. Diccionario de datos.
- 1.2. Módulos.
- 1.3. Preguntas en los módulos.
- 1.4. Aplicación de entrada de datos.
- 1.5. Formularios.
- 1.6. Digitación y exportación de datos.
- 1.7. Estadísticos descriptivos.

### 2. PSPP

- 2.1. Análisis de datos exploratorio.
  - 2.1.1. Medidas de posición y variabilidad.
  - 2.1.2. Frecuencias.
  - 2.1.3. Tablas cruzadas.
  - 2.1.4. Visualizaciones.
- 2.2. Funciones intermedias.
  - 2.2.1. Crear, modificar y recodificar variables.
  - 2.2.2. Agregaciones.
  - 2.2.3. Uso de ponderadores.

### 3. Python.

- 3.1. Introducción a Python.
  - 3.1.1. Usos básicos.
  - 3.1.2. Variables y sus tipos.
  - 3.1.3. Listas.
  - 3.1.4. Funciones y paquetes de Python.
  - 3.1.5. Numpy.
- 3.2. Python intermedio.
  - 3.2.1. Diccionarios de datos.
  - 3.2.2. Pandas.
  - 3.2.3. Estructuras de control.
  - 3.2.4. Ciclos.
  - 3.2.5. Creación de funciones.
- 3.3. Visualización.
  - 3.3.1. Matplotlib.
  - 3.3.2. Seaborn.
  - 3.3.3. Mapas.
- 3.4. Bases de datos.
  - 3.4.1. Importación y exportación.
  - 3.4.2. Limpieza de datos.
- 3.5. Análisis de datos.
  - 3.5.1. Medidas de posición y variabilidad.
  - 3.5.2. Medidas de asociación y correlación.
  - 3.5.3. Aplicaciones de modelos de regresión.



- 3.6. Simulaciones.
  - 3.6.1. Números aleatorios.
  - 3.6.2. Métodos de remuestreo.
  - 3.6.3. Método Montecarlo.
- 3.7. Web Scrapping.
- 4. SQL.**
  - 4.1. Conceptos básicos.
    - 4.1.1. Coincidencias con dplyr (R).
    - 4.1.2. Selección de columnas.
    - 4.1.3. Filtrado de casos.
    - 4.1.4. Agregado de datos.
    - 4.1.5. Agrupamientos.
  - 4.2. Unión de datos.
    - 4.2.1. Coincidencias con dplyr (R).
    - 4.2.2. Uso de los joins.
  - 4.3. Consultas.
  - 4.4. PostgreSQL.
  - 4.5. Manipulación de datos con PostgreSQL.
  - 4.6. Conexión con R y Python.

### CRONOGRAMA DE ACTIVIDADES

FECHA	MODALIDAD		ACTIVIDAD
	Grupo 1	Grupo 2	
18/08/2022	Virtual	Virtual	Temas 1.1 al 1.9
25/08/2022	Virtual	Virtual	Avance del trabajo práctico 1
01/09/2022	Virtual	Presencial: 140 CE	Tema 2.1 Entrega y presentación del trabajo práctico 1
08/09/2022	Virtual	Virtual	Avance del trabajo práctico 2
15/09/2022	Presencial: 140 CE	Virtual	Tema 2.2 Entrega y presentación del trabajo práctico 2
22/09/2022	Virtual	Virtual	Avance del trabajo práctico 3
29/09/2022	Virtual	Presencial: 140 CE	Temas 3.1 al 3.3 Entrega y presentación del trabajo práctico 3
06/10/2022	Virtual	Virtual	Avance del trabajo práctico 4
13/10/2022	Presencial: 140 CE		Temas 3.4 al 3.7 Entrega y presentación del trabajo práctico 4
20/10/2022	Virtual	Virtual	Avance del trabajo práctico 5
27/10/2022	Virtual	Presencial: 140 CE	Temas 4.1 al 4.2 Entrega y presentación del trabajo práctico 5

TEL. 2511-6483 / 2511-6500 / FAX 2511-6510





03/11/2022	Virtual	Virtual	Avance del trabajo práctico 6
10/11/2022	Presencial: 140 CE	Virtual	Temas 4.3 al 4.6 Entrega y presentación del trabajo práctico 6
17/11/2022	Presencial: 140 CE	Virtual	Avance del trabajo práctico 7
24/11/2022	Virtual	Presencial: 140 CE	Entrega y presentación del trabajo práctico 7 Entrega de notas finales del curso
02/12/2022			Ampliación

### BIBLIOGRAFÍA

SIGNATURA/ENLACE	TÍTULO
<a href="https://www.census.gov/data/software/cspro.documentation.html">https://www.census.gov/data/software/cspro.documentation.html</a>	United States Census Bureau. (2022). Census and Survey Processing System (CSPRO).
005.36 P838I	Introducción al análisis estadístico descriptivo con PSPP: metodología y casos prácticos.
<a href="https://tinyurl.com/2c3o5bng">https://tinyurl.com/2c3o5bng</a>	PSPP Users' Guide
<a href="https://tinyurl.com/2mkflkj9">https://tinyurl.com/2mkflkj9</a>	An introduction to statistics with python: With applications in the life sciences.
<a href="https://tinyurl.com/2mbvycd4">https://tinyurl.com/2mbvycd4</a>	Miller, C. (2018). Hands-On Data Analysis with NumPy and Pandas. Birmingham: Packt Publishing, Limited.
<a href="https://tinyurl.com/2fj7brzt">https://tinyurl.com/2fj7brzt</a>	Joe Celko. (2005). Chapter 23 - Statistics in SQL. In Joe Celko's SQL for Smarties (Third ed., pp. 509-548). Elsevier.
<b>MATERIAL COMPLEMENTARIO</b>	
<a href="https://www.postgresql.org/download/">https://www.postgresql.org/download/</a>	Instaladores PostgreSQL
<a href="https://www.postgresql.org/docs/9.6/static/index.html">https://www.postgresql.org/docs/9.6/static/index.html</a>	Manual de usuario PostgreSQL

