



UNIVERSIDAD DE SAN CARLOS DE GUATEMALA
FACULTAD DE CIENCIAS QUÍMICAS Y FARMACIA

PROGRAMA DE CURSO

1. Descripción general del curso		
1.1	Nombre	Ciencia de Datos aplicada
1.2	Código	Pendiente
1.3	Créditos	4
1.4	Carrera a la que se le sirve	Escuela de Biología
1.5	Requisitos	Investigación Aplicada II Bioestadística II
1.6	Año y ciclo lectivo en que se ofrece	Año 2021; 10.º Ciclo
1.7	Fecha de inicio y finalización	14 de julio de 2021 30 de noviembre de 2021
1.8	Horario	Teoría: Viernes de 16:00 a 19:00 Laboratorio: Lunes 9:00:00 a 12:00:00 horas
1.9	Salón, laboratorio y otros espacios en los que se realizará	Curso en Línea de Coursera
1.10	Página web o blog	https://www.coursera.org/specializations/applied-data-science

2. Personal académico		
2.1	Departamento al que pertenece el curso	Departamento de Zoología, Genética y Vida Silvestre
2.2	Escuela o Programa	Escuela de Biología
2.3	Profesor/es	Revisor de avance del curso: M.Sc. Edgar Gustavo Ruano Fajardo
2.4	Correo electrónico	ruano.gustavo@profesor.usac.edu.gt
2.5	Auxiliar de cátedra	No aplica
2.6	Atención al estudiante	Horario: 11:00 a 19:00 horas

3. Descripción general del curso

3.1	Descriptor	Esta especialización cubre los conceptos y metodologías para aplicar las herramientas que necesita para analizar los datos, manejo de bases de datos y modelación científica; aprovechando la ciencia de la computación y el análisis estadístico.
3.2	Introducción	<p>Este curso brinda herramientas necesarias para analizar datos y tomar decisiones sobre matrices de cualquier naturaleza aprovechando la informática y el análisis estadístico, especialmente en datos con potencial uso en investigación biológica. Por lo que un parte fundamental es que el estudiante aprenda a utilizar el lenguaje de programación Python; desde la programación básica, el análisis y visualización de datos.</p> <p>El objetivo principal del Curso es encontrar patrones en los datos a partir de la extracción, el tratamiento y el pre-procesamiento de los datos. El estudiante tiene la meta de hacer predicciones a partir de los datos. Al final, el estudiante debe ser capaz de proponer potenciales caminos para la toma de decisiones en distintos ámbitos de investigación biológica.</p>
3.3	Valores y principios	<p>Responsabilidad: valor que permite a los miembros de la comunidad universitaria a interactuar, comprometerse y aceptar las consecuencias de las acciones y decisiones. Los actos responden íntegramente a los compromisos, sin necesidad de tener supervisión, en cumplimiento del deber con eficiencia y eficacia.</p> <p>Honestidad: valor que distingue a los miembros de la comunidad de la Universidad de San Carlos de Guatemala por sus actos de probidad, rectitud, decoro y decencia.</p> <p>Servicio: es la disposición de la comunidad universitaria de atender con agilidad, cordialidad, eficiencia y diligencia a la comunidad universitaria y</p>

a las personas que hacen uso de los servicios de la Universidad, en función de la misión institucional.

4. Objetivos de aprendizaje del curso

4.1	Nivel cognitivo	<p>I. Objetivos generales</p> <p>Que el estudiante:</p> <ul style="list-style-type: none"> • Comprenda los fundamentos para desarrollar análisis de datos en el área de las ciencias biológicas. <p>Objetivos específicos</p> <p>Que el estudiante:</p> <ul style="list-style-type: none"> • Integre como el complementar del uso de la programación en la investigación biológica. • Desarrolle la capacidad de como sintetizar obtener datos en medios digitales y transfórmalos en información para estudios en ciencias biológicas
4.2	Nivel psicomotriz	<p>Que el estudiante:</p> <ul style="list-style-type: none"> • Organice información de una forma sistematizada. • Seleccione y aplique métodos analíticos adecuados para explorar datos. • Integre como el complementar del uso de la programación en la investigación biológica.
4.3	Nivel afectivo	<p>Que el estudiante:</p> <ul style="list-style-type: none"> • Complete valores éticos, en base a estándares y criterios específicos, que se utilizan en el uso de la ciencia de datos.

5. Metodología

Para alcanzar los objetivos del curso se desarrollarán: clases magistrales impartidos por los profesores de Coursera por un medio digital. Este mecanismo tiene discusiones grupales por medio de plataforma digital, videos educativos, prácticas de laboratorio por medio digitales. Además, se tendrá presentaciones a distancia del avance del curso al revisor de la escuela de biología por parte de los estudiantes cada semana.

6. Programación de las actividades académicas del curso: Sección pedagógica

Unidad	Fechas	Objetivos	Contenidos	Actividades de aprendizaje	Materiales y recursos	Instrumento de evaluación
Módulo I Conceptos básicos de Python	14/07/2021	<ul style="list-style-type: none"> Aprendizaje de Python Soluciones en toma de datos y su análisis. 	Conceptos básicos de Python Estructuras de datos de Python Fundamentos de programación Trabajar con datos	Ejercicios prácticos por módulos y creación de proyecto final.	Videos Lecturas Ejercicios	Examen practico Examen graduado
Módulo II Estructuras de datos de Python	16/07/2021	Aprendizaje de expresiones y operaciones matemáticas	<ul style="list-style-type: none"> Almacenar valores en variables y diferentes formas de manipular cadenas 	Ejercicios prácticos	Videos Lecturas Ejercicios	Examen graduado
Módulo II Estructuras de datos de Python	22/07/2021	Uso de listas y tuplas para almacenar colecciones de datos	Diccionarios y funcionamiento de pares de claves, valores y conjuntos.	Ejercicios prácticos	Videos Lecturas Ejercicios	Examen graduado

Módulo III Fundamentos de programación en Python	23/07/2021	Aprender los conceptos de condiciones y ramificaciones	<ul style="list-style-type: none"> • Creación de funciones de clases. 	Ejercicios prácticos	Videos Lecturas Ejercicios	Examen graduado
Módulo IV Trabajar con datos en Python	28/07/2021	Aprender a leer y escribir archivos desde Python	<p>Uso de las</p> <ul style="list-style-type: none"> • bibliotecas de manipulación de datos y operaciones matemáticas 	Ejercicios prácticos	Videos Lecturas Ejercicios	Examen Graduado
Curso 2 Python Project for Data Science	04/08/2021	Aplicar fundamentos y estructuras para construir modelos. Construir panel de control mediante Jupyter notebook.	Cuadros de mandos de Crowdsourcing HTML Webscraping	Ejercicios prácticos	Videos Lecturas Ejercicios	Examen graduado
Curso 3 - Data Analysis with Python	12/08/2021	Preparación de datos y métodos estadísticos Uso de herramientas y elaboración de pipelines	Importación de conjunto de datos Gestión de datos Análisis exploratorio de datos Desarrollo de modelos Evaluación de modelos	Ejercicio prácticos por módulo y proyecto final	Videos Lecturas Ejercicios	Pruebas de practica Examen practico Examen graduado

Módulo I Importación de conjuntos de datos	18/08/2021	Reconocer diferentes bibliotecas para importación de datos	Bibliotecas de Python	Ejercicios prácticos	Videos Lecturas Ejercicios	Examen graduado
Módulo II Gestión de datos	19/08/2021	Reconocer diferentes bibliotecas para depuración de bases de datos	Bibliotecas de Python	Ejercicios prácticos	Videos Lecturas Ejercicios	Examen graduado
Módulo III Análisis exploratorio de datos	25/08/2021	Utilizar funciones y bibliotecas para resumir e identificar conjuntos de datos	Bibliotecas de Python	Ejercicios prácticos	Videos Lecturas Ejercicios	Examen graduado
Módulo IV Desarrollo de modelos	01/09/2021	Desarrollar modelos multivariados.	Regresión Lineal Regresión Lineal Múltiple Visualización y evaluación del modelo Pipelines y regresión polinomial	Ejercicios prácticos	Videos Lecturas Ejercicios	Examen graduado

			Evaluación de muestra Predicciones			
Módulo V - Evaluación de modelos	09/09/2021	Evaluar modelos, Seleccionar modelos	Evaluación y refinamiento de modelos Sobreajuste, infraajuste y selección de modelos Regresión Ridge Grid Search	Ejercicios prácticos	Videos Lecturas Ejercicios	Examen graduado
Módulo VI Proyecto	22/09/2021	Proveer un escenario de proyecto	Implementación de los módulos previos	Ejercicio práctico	Videos Lecturas Ejercicio	Examen final
Curso 4 - Data Visualization with Python	30/10/2021	Aprender a visualizar datos, extraer información y tomar decisiones eficaces.	Introducción a las herramientas de visualización de datos Herramientas de visualización básicas y especializadas Visualizaciones avanzadas y datos	Ejercicios prácticos	Videos Lecturas Ejercicios	Test de practica Pruebas de practica Examen practico Examen graduado

			geoespaciales Creación de cuadros de mando con Plotly y Dash			
Módulo I - Introducción a las herramientas de visualización de datos	06/10/2021	Utilizar trazados básicos con Matplotlib	Manejo de conjunto de datos en csv en marco de datos de Pandas Procesar y manipular en Matplotlib	Ejercicios prácticos	Videos Lecturas Ejercicios	Examen graduado
Módulo II Herramientas de visualización básica y especializada	13/10/2021	Crear gráficos de diferentes bibliotecas	Gráficos de área, barras, histogramas, circulares, caja, dispersión	Ejercicios prácticos	Videos Lecturas Ejercicios	Examen graduado
Módulo III Visualizaciones avanzadas y datos geoespaciales	22/10/2021	Utilizar herramientas de visualización avanzada.	Gráficos waffle, word clouds, seaborn. Biblioteca Folium, marcadores de mapas coropléticos.	Ejercicios prácticos	Videos Lecturas Ejercicios	Examen graduado

Módulo IV Cuadros de mando	29/10/2021	Creación de cuadros de mando con Plotly y Dash	Listas de cuadros Cuadros de mando Biblioteca Plotly Biblioteca Dash	Ejercicios prácticos	Videos Lecturas Ejercicios	Examen graduado
Curso 5 - Applied Data Science Capstone	5/11/2021	Manipular biblioteca Folium Comunicar resultados y conclusiones	Datos de localización y modelación Foursquare API Neighborhood Sementation and Clustering The Battle of Neighborhoods	Ejercicios prácticos por módulo y proyecto final	Videos Lecturas Ejercicios	Test de practica Pruebas de practica Examen practico Examen final
Módulo I - Introducción	12/11/2021	Utilizar diferentes proveedores de datos	Repositorios de Github	Ejercicios prácticos	Videos Lecturas Ejercicios	Examen graduado
Módulo II Foursquare API	12/11/2021	Crear cuenta de desarrollador Utilizar credenciales para búsqueda de datos	Foursquare API	Ejercicios prácticos	Videos Lecturas Ejercicios	Examen graduado

Módulo III Neighborhood Segmentation and Clustering	19/11/2021	Utilizar clustering de k-means	Clustering Foursquare API Segmentación y agrupación Raspado de código HTML Python Beautifulsoup	Ejercicios prácticos	Videos Lecturas Ejercicios	Examen graduado
---	------------	-----------------------------------	--	-------------------------	----------------------------------	--------------------

7. Evaluación del aprendizaje

Con el fin de que los estudiantes continúen con un desarrollo de sus habilidades de programación se seguirá un proceso de educación de avance supervisado cada semana por medio de una presentación. Además, se revisarán los certificados de prueba aprobados de cada uno de los módulos.

Actividad de aprendizaje	Punteo	Porcentaje
Presentaciones de sección	25	25 %
Comprobación de prueba de sección	25	50 %
Presentación de proyecto	15	65 %
Laboratorio práctico	15	80 %
Zona	80	80 %
Evaluación final del curso	20	20 %
Total	100	100 %

8. Referencias

	Descripción	Disponible en
7.1	<p>Bibliografía principal u obligatoria</p> <p>Anggri Yulio. (2018). Web Scraping dengan Python dan BeautifulSoup. <i>DevTriK</i>.</p> <p>Hilpisch, Y. J. (2016). <i>Python for finance : mastering data-driven finance</i>.</p> <p>Lubanovic, B. (2019). <i>Introducing Python, 2nd Edition</i>.</p> <p>VanderPlas, J. (2016). Python Data Science Handbook: Essential Tools for Working with Data. <i>O'Reilly</i>, 541.</p> <p>McKinney, W. (2018). Python for Data Analysis : Data Wrangling with Pandas, NumPy, and IPython. <i>Transplantation</i>, 71(10), 1385–1389.</p>	Recursos y clases de Especialización de Coursera
7.2	<p>Recursos en línea</p> <ul style="list-style-type: none"> • https://cloud.ibm.com/ • https://www.python.org/ • https://www.ibm.com/ • https://cloud.ibm.com/catalog 	