

CARRERA: Licenciatura en Ciencias de Datos	CURSO LECTIVO: 2024
CÁTEDRA: Programación Orientada a Datos	CURSO: 1º año - 2º semestre
DURACIÓN: Semestral	Hs. TOTALES: 80 Hs. Reloj Totales
SEMANAS: 16	Hs. TEÓRICAS: 32 Hs. Reloj Totales Hs. PRÁCTICAS: 48 Hs. Reloj Totales

PROFESOR ASISTENTE: Saavedra,
Lucas

1. OBJETIVOS DE LA ASIGNATURA

Que los alumnos logren:

- Familiarizarse con las técnicas de programación que les permitirán manipular datos a lo largo de toda la carrera.
- Introducirse a la Ingeniería de Software y las metodologías de desarrollo de software. Reconocer la importancia de la Ingeniería de Software en los procesos de la Ingeniería de Datos en particular y de las Ciencias de Datos en general.
- Reconocer la importancia del paradigma de programación orientada a objetos.
- Comprender el contexto en donde debe aplicarse cada una de las metodologías de programación: dónde aplica la orientación a objetos y dónde aplica la orientación a datos.
- Resolver programáticamente problemas de manipulación y análisis de datos.

2. UNIDADES TEMÁTICAS

1. Introducción a la Ingeniería de Software.

Definición. Historia. Mantenibilidad. Principios fundamentales de la Ingeniería de Software. Casos de Uso. Modelado UML. Diagramas de Secuencia. Diagramas de Estados. Nociones de Diseño y Arquitectura de Software. Calidad del Software. Testing Unitario. Testing integral. Software Configuration Management. Métrica. Metodologías Ágiles. El proceso del Software. Métricas del proceso y del proyecto. Metodologías clásicas. Metodologías Ágiles.

2. Programación Orientada a Objetos.

Estado y Comportamiento. Análisis Orientado a Objetos. Diagramas de clase UML. Breve introducción a los patrones de Diseño. Modelo de Dominio. Relación del modelo Orientado a Objetos con las bases de datos. Nociones de Programación Orientada a Aspectos. Encapsulación de Datos.

3. El enfoque de la Programación Orientada a Datos.

La evolución del modelado de datos en los programas. Fuentes de complejidad del paradigma de objetos. Visión sobre la encapsulación de datos.

4. Fuentes de Datos

Archivos de texto plano. Archivos de texto con separador, CSV. Planillas de Cálculo. Colas de mensajería, Streaming. Bases de Datos Relacionales. Bases de Datos clave-valor. Bases de Datos documentales. API Rest.

5. Series y Dataframes

Pandas. Estructuras útiles para la manipulación de datos. Series. Dataframes. NumPy. Similaridades entre Series y Dataframes. Orden. Índice. Row Index. Column Index. Selección de Filas y Columnas en un Dataframe. Filtros. Optimización para uso en memoria. Manejo de Duplicados. Manejo de Colecciones. Operación Reduce.

6. Concurrencia

Inmutabilidad. Variable vs Valores. Concurrencia. Concurrencia optimista. Algoritmo de Diferencia Estructural. Locks. Deadlocks. Thread Safety.

7. Validación de Datos

Data Validation. Data Schema. JSON Schema. Modo estricto. Modo Flexible.

8. Introducción a Spark

Introducción al cómputo distribuido en Datos. Spark. PySpark. Dataframe distribuido.

3. BIBLIOGRAFÍA

3.1 BIBLIOGRAFÍA GENERAL OBLIGATORIA

- “Data Oriented Programming: Reduce software complexity” - Y. Sharvit - Ed. Manning - 2022 - ISBN 9781617298578 .
- “Software Engineering: A Practitioner's Approach” - R. Pressman & B.R. Maxim - 8va Edición - McGraw-Hill Education - 2015 - ISBN 1259253155, 9781259253157.
- “Pandas in Action” - Boris Paskhaver - Ed. Manning - 2021 - ISBN 9781617297434.

4. METODOLOGÍA

El curso está organizado en 8 unidades temáticas divididas en encuentros teórico-prácticos de 5hs semanales, a realizarse en formato presencial. La modalidad adoptada para el dictado será **teórico-práctica**. En las clases se presentarán los temas de cada unidad, proponiendo espacios de intercambio con el docente y entre los estudiantes a partir de consignas específicas. Se facilitará material de lectura obligatoria y complementaria para complementar la comprensión de las unidades.

5. EVALUACIONES Y CRITERIOS PARA LA APROBACIÓN

La aprobación de la materia estará supeditada al cumplimiento de la condición

de asistencia exigida por la Universidad, la aprobación de todas las actividades prácticas (TP) y la aprobación de los dos exámenes parciales.

Los trabajos prácticos podrán ser individuales o grupales, debiéndose cargar a través de la plataforma de Entornos Virtuales de Aprendizaje en tiempo y forma, otorgándole una única instancia de revisión y recuperación. Las actividades prácticas deberán contar con su aprobación para acceder a la instancia de evaluación final.

Para los trabajos prácticos y la evaluación final se realizarán sesiones de consultas individuales y grupales, haciendo además puesta en común general si el caso lo requiera. A los estudiantes que presenten dificultades se les observará y se los guiará para resolver el conflicto.

Las instancias de recuperación están previstas para estudiantes que no hayan aprobado los exámenes parciales o que hayan estado ausentes al momento de realizarse.

Criterios de Evaluación:

- Respeto de las consignas presentadas.
- Resolución correcta de los problemas planteados.
- Adecuada respuesta a los contenidos teóricos.

6. CRITERIOS y MODALIDAD PARA LA EVALUACIÓN DEL EXAMEN FINAL

El examen final consiste en una evaluación oral y escrita, presencial e individual, donde el alumno deberá demostrar conocimientos teóricos y prácticos. El examen final se diferencia en que abarca todos los temas del programa. Los ejercicios prácticos tendrán un carácter integrador, articulando los distintos contenidos vistos en la materia. Finalmente, en las preguntas teóricas se pretende que el alumno demuestre un conocimiento profundo de los temas, relacionando conceptos entre sí.

Criterios de Evaluación:

- Respeto de las consignas presentadas.
- Adecuada respuesta a los contenidos teóricos.
- Relación de conceptos pertinente.
- Resolución correcta de los problemas planteados.
- Fundamentación bibliográfica de los temas.