

**PLANIFICACIÓN DE CURSO**  
Primer Semestre académico 2024

**I. ACTIVIDAD CURRICULAR Y CARGA HORARIA**

Asignatura:	Exploración, Visualización y Mantenimiento de Datos	Código:	MMD3201
Semestre de la Carrera:	V semestre		
Carrera:	Ing. Civil en Modelamiento Matemático de Datos		
Escuela:	Escuela de Ingeniería		
Docente(s):	Raúl Valenzuela		
Horario:	Lunes y Miércoles 12:00 a 13:30		

Créditos SCT:	6
Carga horaria semestral <sup>1</sup> :	180 horas
Carga horaria semanal:	8,5 horas

Tiempo de trabajo directo semanal:	3 horas
Tiempo de trabajo del estudiante semanal:	5,5 horas

**II. RESULTADOS U OBJETIVOS DE APRENDIZAJE ESPERADOS ESTE SEMESTRE**

1)	Conocer distintas herramientas de la estadística para explorar datos
2)	Aplicar técnicas de exploración visual de datos utilizando un lenguaje de programación
3)	Entender y aplicar los principios de visualización de datos para comunicar un análisis exploratorio y explicativo a una audiencia amplia y una audiencia técnica
4)	Determinar y describir las características principales y limitaciones de una base de datos real utilizando métodos exploratorios visuales y estadísticos
5)	Estar familiarizado con las tareas de mantenimiento de datos e investigación reproducible

<sup>1</sup> Considere que 1 crédito SCT equivale a 30 horas de trabajo total (directo y autónomo) en el semestre.

### III. UNIDADES, CONTENIDOS Y ACTIVIDADES

UNIDAD: (REPLICAR TANTAS VECES COMO UNIDADES SE DISPONGAN EN EL CURSO)				
Semana	Contenidos	Actividades de enseñanza y aprendizaje		Actividades de evaluación diagnóstica, formativa y/o sumativa
		Tiempo directo	Tiempo trabajo autónomo del o la estudiante	
1	- <i>Introducción al curso</i> - <i>Medidas de tendencia central y dispersión</i>	1,5	7,0	Lectura Peck & Devore 4.1-4.2
2	- <i>Medidas de tendencia central y dispersión</i> - <i>Regresión lineal</i>	3	5,5	Lectura Peck & Devore 5.1-5.3
3	- <i>Pandas y Regresión Int.a Matplotlib y Proplot</i>	3	5,5	Quiz 1
4	- <i>Manipulación e imputación de datos</i> - <i>Lab: Pandas y Regresión</i>	3	5,5	
5	- <i>Origen y técnicas de Análisis Exploratorio de Datos</i> - <i>Lab: Plot líneas, scatter, barras, área, histograma</i>	3	5,5	Quiz 2
6	<i>Consultas sobre proyectos</i> - <i>Lab: Plot formato, líneas auxiliares, paneles, ejes, anotaciones</i>	3	5,5	
7	- <i>Narrativa con datos</i>	3	5,5	

8	- <i>Series de tiempo con pandas</i> - <i>Lab: series de tiempo</i>	3	5,5	Quiz 3
9	<i>Presentación de proyectos I</i>	3	5,5	Presentación proyectos
10	- <i>Teoría de colores para visualización</i> <i>Lab: uso de colores, mapa de colores</i>	3	5,5	
11	- <i>Visualización exploratoria vs explicativa</i> - <i>Lab: heatmap, histograma 2D</i>	3	5,5	Quiz 4
12	- <i>Introducción a geodatos con Python</i> - <i>Lab: geodatos</i>	3	5,5	
13	- <i>Ciclo de vida y planificación de datos</i> - <i>Control de cátedra</i>	3	5,5	Control de cátedra
14	- <i>Presentaciones Proyectos II</i>	3	5,5	Presentaciones proyecto

#### IV. CONDICIONES Y POLÍTICAS DE EVALUACIÓN

El curso tiene distintas instancias de evaluación las cuales se detallan a continuación:

- Control de cátedra (40%)
- Presentación de proyecto (30%): propuesta (30%) y proyecto final (70%)
- Quizzes de 10 minutos (4) - individual (30%)

La nota final del curso corresponde a la suma ponderada de las notas de evaluaciones parciales con los porcentajes indicados arriba. El curso se aprueba con nota final  $\geq 4,0$ . Estudiantes que estén con nota final entre 3.8 y menor a 4.0 podrán rendir una prueba oral para determinar si aprueban el curso. **Este curso no considera la realización de examen final.**

Cualquier sospecha de plagio será sancionada de acuerdo a al reglamento de Escuela de Ingeniería de la Universidad de O'Higgins.

Se espera asistencia permanente a clases. La ausencia a cualquier actividad de evaluación debe ser debidamente justificada mediante la Dirección de Asuntos Estudiantiles.

Todas las evaluaciones tienen 60% de exigencia.

#### V. BIBLIOGRAFÍA Y RECURSOS OBLIGATORIOS

1. McKinney, W. (2017). Python for Data Analysis. O'Reilly Media.
2. Nussbaumer, C. (2015). Storytelling with Data. John Wiley & Sons.
3. Peck, R., & Devore, J. L. (2011). Statistics: The Exploration & Analysis of Data. Cengage Learning.

#### VI. BIBLIOGRAFÍA Y RECURSOS COMPLEMENTARIOS

1. Briney, K. (2015). Data Management for Researchers: Organize, Maintain and Share Your Data for Research Success. Pelagic Publishing.
2. Devore, J. L. (2011). Probability and Statistics for Engineering and the Sciences. Cengage Learning.
3. Enders, C. K. (2010). Applied Missing Data Analysis. Guilford Press.
4. Manski, Charles F. Identification for prediction and decision. Harvard University Press, 2009.