

Programa de Asignatura

1. Identificación Asignatura

Nombre:	Visualización de Datos		Código:	
Carrera:	Ingeniería Civil Informática	Unidad Académica:	Ciencias Naturales y Tecnología	
Ciclo Formativo:	Ciclo Licenciatura	Línea formativa:	Especializada	
Semestre	VII	Tipo de actividad:	Obligatoria	
N° SCT:	6	Horas Cronológicas Semanales		
		Presenciales:	4.5	Trabajo Autónomo:
Pre-requisitos	Sin pre-requisitos (propuesta Minería de Datos + Lenguajes de Programación)			

2. Propósito formativo

El propósito del curso de Visualización de Datos es sumergir a los estudiantes en el estudio y la aplicación práctica de técnicas avanzadas de visualización, enfatizando su rol crucial en el análisis de datos y la toma de decisiones estratégicas. A través de un enfoque teórico-práctico, los estudiantes adquirirán habilidades para manejar, analizar y visualizar datos complejos, aprendiendo a convertirlos en representaciones visuales que faciliten la comprensión y el análisis.

La asignatura se desarrolla alrededor del ciclo completo de los datos, destacando la visualización como fase final crítica, donde los datos se transforman en conocimiento. Se cubrirán aspectos fundamentales como la selección de las herramientas y técnicas de visualización más apropiadas para diferentes tipos de datos, la importancia de una visualización efectiva en la extracción de insights y la creación de narrativas de datos convincentes.

Este enfoque integral asegura que los estudiantes no solo aprendan a manejar herramientas de visualización de datos, sino también a comprender la teoría y principios que guían la práctica de visualización, preparándolos para aplicar estos conocimientos en situaciones reales y convertirse en profesionales capaces de liderar proyectos de visualización de datos.

Esta asignatura es un recurso de apoyo relevante para la línea de ciencia de datos y aunque no es pre-requisito, es un primer acercamiento para llegar con mejor base al curso Big Data del semestre VIII. Es importante contar con bases conceptuales de minería de datos, programación y estadística aplicada.

3. Contribución al perfil de egreso

Esta asignatura contribuye a los siguientes desempeños o resultados de aprendizaje globales declarados en el Perfil de Egreso de la carrera:

1. Genera información relevante, por medio de métodos analíticos y experimentales, a partir de fuentes de datos de gran volumen y diversa complejidad, utilizando técnicas pertinentes de minería de datos, con el fin de apoyar la toma de decisiones.
2. Gestiona componentes organizacionales que estén asociados a la captura, almacenamiento, transformación y provisión de información, de forma que se puedan satisfacer necesidades de información con estándares de calidad adecuados

4. Resultados de aprendizaje específicos

Resultado de Aprendizaje Específico	Criterios de evaluación	Evidencia
RA1. Aplica el proceso de visualización de datos para analizar grandes volúmenes de datos.	1.1. Aplica técnicas para pre-procesar, analizar y visualizar grandes volúmenes de datos desconocidos. 1.2. Las visualizaciones generadas dan cuenta del uso del proceso de visualización (pre-procesamiento de los datos, selección de vistas, mapeo visual, despliegue en pantalla).	Laboratorios, guías de ejercicio.
RA2. Conoce y aplica los fundamentos de visualización para crear visualizaciones eficaces.	2.1. Conoce la fisiología de la visión humana y las distintas etapas del procesamiento de la percepción visual. 2.2. Reconoce las variables visuales y fundamenta adecuadamente su uso en visualizaciones.	Laboratorios, guías de ejercicio.
RA3. Conoce y utiliza herramientas de visualización para distintos tipos de datos.	3.1. Conoce y aplica técnicas de visualización para datos espaciales (multidimensionales), jerárquicos, grafos, geoespaciales, temporales, textuales. 3.2. Las visualizaciones generadas resaltan y comunican aspectos interesantes de los datos.	

5. Unidades de Aprendizaje

<p>1. Introducción</p> <p>1.1. Fundamentos de la Ingeniería de Datos</p> <ul style="list-style-type: none"> - Introducción al CRISP-DM - Concepto de Data Frame - Repositorios Git para gestión de proyectos de datos <p>1.2. Manejo de Datos</p> <ul style="list-style-type: none"> - Datos abiertos y fuentes de datos (kaggle.com, pydataset) - Lectura, limpieza y filtrado de datos - Escritura y almacenamiento de datos <p>1.3. Visualización de Datos</p> <ul style="list-style-type: none"> - Principios y fundamentos de visualización de datos - Herramientas de visualización: matplotlib, seaborn - Fundamentos a técnicas de visualización de datos <p>1.4. Proyectos Prácticos y Herramientas</p> <ul style="list-style-type: none"> - Implementación de proyectos básicos - Uso de Jupyter Notebooks para análisis de datos
<p>2. Conceptos Básicos de Estadística y Análisis Exploratorio</p> <p>2.1. Estadística Descriptiva</p> <ul style="list-style-type: none"> - Estadísticos básicos y gráficas

- Muestreo aleatorio y teorema central del límite
- 2.2. Inferencia Estadística**
 - Contrastes de hipótesis
 - Correlación y causalidad
- 2.3. Selección de Atributos**
 - Métodos y técnicas de selección de atributos
 - Importancia de variables en modelos predictivos
- 2.4. Proyecto Intermedio**
 - Aplicación de estadística en proyectos de datos
 - Avance de proyecto bajo metodología CRISP-DM

3. Análisis de Regresión y Modelos Predictivos

- 3.1. Regresión Lineal y Logística**
 - Regresión lineal simple y múltiple
 - Regresión logística para clasificación
- 3.2. Modelos Avanzados de Regresión**
 - Regresión polinomial
 - Árboles de regresión y Random Forest
- 3.3. Evaluación de Modelos**
 - Métricas de rendimiento
 - Validación cruzada
- 3.4. Proyectos de Regresión**
 - Desarrollo de modelos predictivos
 - Análisis de casos de estudio

4. Clustering, Clasificación y Big Data

- 4.1. Fundamentos de Clustering y Clasificación**
 - Concepto de distancia y métodos de enlace
 - K-medoides, clustering espectral, K-means
- 4.2. Modelos de Clasificación Avanzados**
 - K Nearest Neighbors (kNN)
 - Árboles de decisión y Random Forest
- 4.3. Introducción al Big Data**
 - Almacenamiento de grandes cantidades de datos
 - Ecosistema Hadoop: conceptos básicos y herramientas
- 4.4. Visualización Avanzada de Datos**
 - Análisis y visualización de datos con Tableau/Power BI
 - Conexión y manipulación de datos para visualización avanzada
 - Técnicas de visualización de datos geoespaciales, jerárquicos, temporales, textuales.

6. Recursos de Aprendizaje

Bibliografía:

B1: Matthew O. Ward, Georges Grinstein, Daniel Keim. Interactive Data Visualization: Foundations, Techniques, and Applications. Second Edition. 2015.

B2: Andy Kirk "Data Visualisation: A Handbook for Data Driven Design" 2016

B3: Kyrian Dale "Data Visualization with Python and JavaScript: Scrape, Clean, Explore & Transform Your Data" 2016

Recursos materiales e infraestructura:

- Acceso a Ucampus.

Computadores debidamente equipados para utilizar lenguajes de alto nivel (por ej.: Python, Tableau, Power BI, otros).

7. Comportamiento y ética académica:

Se espera que, en el marco de sus actividades académicas y estudiantiles, los estudiantes demuestren un compromiso inquebrantable con los estándares de ética y honestidad académica que distinguen a nuestra comunidad universitaria. Este compromiso debe reflejarse en todas las acciones y trabajos realizados dentro del ámbito de la Universidad de Aysén, siguiendo rigurosamente las directrices establecidas en el Reglamento de Estudiantes, particularmente lo indicado en los artículos 23°, 24°, y 26°.

Además, es imperativo que todas las acciones se alineen con los reglamentos institucionales aplicables y las normativas internas vigentes, incluido el código de ética de la universidad. Estos documentos no solo rigen nuestros comportamientos cotidianos sino que también establecen las expectativas para nuestra convivencia y trabajo académico.

Cualquier infracción a estos principios de honestidad académica, detectada durante la realización, presentación, o entrega de cualquier actividad evaluativa del curso, acarreará consecuencias serias. Estas incluyen la suspensión inmediata de la actividad en cuestión y la asignación de la nota mínima posible (1.0) como reflejo de la gravedad del acto. Esta política se aplica de manera estricta y sin excepciones, reafirmando nuestro compromiso con una formación académica íntegra y responsable.

Planificación del curso



8. Responsables

Académico (s) Responsable (s) y equipo docente	Profesor: Johnny Valencia C		
Contacto	Correo: johnny.valencia@uaysen.cl		
Año	2024	Periodo Académico	Primer semestre
Horario clases	Cátedra: Lu 14.30 – 16.00 hrs*. Mi 14.30 – 16.00 hrs*. Sesión Prácticas: Mi 16.15 – 17.45 hrs*. *Presencial obligatoria	Horario de atención estudiantes	Contactar previamente al profesor mediante email. Disponible Lu 16.00 – 18.00 Vi 14.30 – 16.00
Sala / Campus	Sala Virtual Ucampus – Google Sites		

9. Metodología de Trabajo:

En el curso se contemplan cuatro tipos de actividades docentes, las cuales se asocian a requerimientos de sala y al nivel de intervención del profesor/ayudante:

Actividad docente	Descripción	Intervención del profesor/ayudante	Requerimiento de sala
Exposición conceptual	El profesor introduce los fundamentos preliminares de visualización de datos y conceptos esenciales para el desarrollo de habilidades prácticas, con una participación activa en el aula tradicional o entornos virtuales.	Alta	Sala de clases UCampus Online UCampus Offline
Programación expositiva	Mediante la resolución de problemas específicos de ingeniería de datos, el profesor guía a los estudiantes a profundizar en el entendimiento de conceptos clave, empleando ejemplos prácticos en un entorno de sala de clases o plataformas online.	Alta	Sala de clases UCampus Online UCampus Offline
Programación tutorial	Esta modalidad incorpora pausas estratégicas en la exposición, permitiendo a los estudiantes completar tareas específicas bajo la supervisión del docente, favoreciendo el aprendizaje interactivo en laboratorios de computación o con uso de equipos personales	Media	Laboratorio de computación Computador persona
Actividad práctica / Programación autónoma	Fomenta la independencia de los estudiantes al enfrentar y resolver desafíos de visualización de datos, en grupos o individualmente, con mínima intervención docente, ideal para el trabajo en laboratorios de computación o mediante dispositivos personales.	Baja	Laboratorio de computación Computador persona

Durante el semestre, **se programarán evaluaciones menores** en cualquier momento, con el fin de revisar y reforzar los conocimientos adquiridos hasta la fecha. Estas evaluaciones podrán adoptar diversas formas:

- **Control (Quiz):** Evaluación, ya sea individual o en grupo, realizada en papel durante los primeros 45 minutos de la clase.
- **Trabajo en Clase:** Actividad evaluativa individual o grupal, llevada a cabo en computadora. Podrá ocupar los primeros 45 minutos o extenderse durante toda la sesión.

- **Entrega de Tareas:** Evaluación grupal realizada en computadora y fuera del horario de clase, promoviendo el trabajo colaborativo y la gestión del tiempo.

Adicionalmente, como componente de la calificación final, los estudiantes deberán presentar, en parejas, un proyecto que abarque el análisis, diseño e implementación de una solución tecnológica. Para esta tarea, se proporcionará un instructivo detallado que guiará a los estudiantes en su preparación y entrega, el desarrollo del proyecto se gestará durante todo el semestre, evidenciando una metodología clara de forma progresiva.

10. Evaluaciones:

Evaluación	Ponderaciones específicas	Ponderación nota presentación
Pruebas de cátedra	<ul style="list-style-type: none"> • Prueba 1 (P1): 40% • Quiz 1 (Q1): 10% • Prueba 2 (P2): 40% • Quiz 2 (Q2): 10% 	60%
Evaluaciones menores	El promedio simple entre las notas consideradas corresponde al 100%.	40%

Calificación final:

- Nota de presentación: 70%
- Examen Final: 30 % (Exposición de proyecto de aplicación)

Requisito de Aprobación

- Asistencia: 65%
- Nota Final: 4,0
- La entrega de cualquiera de las evaluaciones indicadas en el programa de la asignatura por fuera del plazo definido serán evaluadas con nota mínima (1,0).
- Sólo para el caso de las evaluaciones presenciales, se permitirá comenzar con un atraso máximo de 15 minutos después de comenzada la misma. El tiempo de retraso no se podrá recuperar. Aquel/la estudiante que se presente a rendir la evaluación después de pasados los 15 minutos, se evaluará con nota mínima (1,0).
- En caso de inasistencia o no entrega en plazo de alguna evaluación, se podrán justificar mediante el mismo procedimiento antes descrito para la justificación de inasistencias actividades lectivas.
- Las clases teóricas tienen un 65% de asistencia mínima obligatoria, siguiendo el mínimo requerido por el Reglamento General de Estudios de Pregrado. La asistencia a las clases prácticas (laboratorios y terrenos) es del 100%. El no cumplimiento de estos porcentajes de asistencia será causal de reprobación de la asignatura.
- En casos debidamente justificados ante el Registro Académico, el/ la estudiante que no haya asistido a una salida a terreno o laboratorio tendrá derecho a rendir examen. Se considerarán debidamente justificadas las inasistencias ante el Registro Académico aquéllas que estén respaldadas con certificados médicos, laborales o algún documento validado por la Unidad de Acceso y Desarrollo Estudiantil. Las actividades de terrenos y laboratorios no podrán ser recuperadas.
- Se permitirá el ingreso posterior a la hora de inicio con un máximo de 15 minutos, siempre y cuando no sea una acción repetida por la/el estudiante (se aceptará máximo de 3 veces). Para las salidas a terreno, el tiempo de espera máximo será de 5 minutos. Para casos donde las actividades lectivas contemplen más de un módulo, el/la estudiante que no haya asistido al módulo anterior, podrá ingresar al comienzo del nuevo módulo.
- En caso de que ningún estudiante se presente a la actividad lectiva después de 15 minutos de comenzada, ésta se suspenderá. Los contenidos programados para dicha actividad se darán por dictados, será responsabilidad del estudiante ponerse al día con los contenidos de dicha clase. Los contenidos de dicha clase, y ejercicios, si así lo hubiera, serán enviados para ser realizados como trabajo autónomo.

Otros aspectos asociados al funcionamiento del curso:

- Toda la coordinación del curso (comunicaciones, actualización de notas, material, etc.) será realizada a través de UCampus.
- **Es deber del estudiante mantenerse informado de las noticias, avisos y material entregado por el profesor a través de estos medios, se sugiere instalar en su dispositivo móvil la aplicación de UCampus.**
- Recordar que los correos electrónicos serán respondidos en horario laboral (lunes a viernes de 9:00 a 18:00hs), no se responderán correos fuera de ese horario.
- Durante el desarrollo de las actividades lectivas, los teléfonos celulares deberán estar en silencio y guardados, a menos que el/la profesor/a específicamente requiera de estos equipos para la realización de su clase o durante algunos casos excepcionales conversados previamente con el/la docente a cargo.
- Las actividades lectivas se dictarán de forma presencial, salvo excepciones sujeto a contingencias presentes durante el transcurso de la asignatura.

11. Planificación de las actividades de enseñanza- aprendizaje y de evaluación

Semana	Resultado(s) de Aprendizaje (*)	Tema (Unidades de aprendizaje) y actividades	Recursos utilizados o lecturas (*)	Actividad(es)
1	RA1	1.1. Fundamentos de la Ingeniería de Datos - Introducción al CRISP-DM - Concepto de Data Frame - Repositorios Git para gestión de proyectos de datos	B1, B2, B3	
2	RA1	1.2. Manejo de Datos - Datos abiertos y fuentes de datos (kaggle.com, pydataset) - Lectura, limpieza y filtrado de datos - Escritura y almacenamiento de datos	B1, B2, B3	
3	RA1	1.3. Visualización de Datos - Principios y fundamentos de visualización de datos - Herramientas de visualización: matplotlib, seaborn - Fundamentos a técnicas de visualización de datos	B1, B2, B3	
4	RA1, RA2	1.4. Proyectos Prácticos y Herramientas - Implementación de proyectos básicos - Uso de Jupyter Notebooks para análisis de datos	B1, B2, B3	Prueba (P1) (3 de abril del 2024)
5	RA1, RA2	2.1. Estadística Descriptiva - Estadísticos básicos y gráficas	B1, B2, B3	

		- Muestreo aleatorio y teorema central del límite		
6	RA1, RA2	2.2. Inferencia Estadística - Contrastes de hipótesis - Correlación y causalidad	B1, B2, B3	
7	RA1, RA2	2.3. Selección de Atributos - Métodos y técnicas de selección de atributos - Importancia de variables en modelos predictivos	B1, B2, B3	Quiz1 (Q1) (24 de abril del 2024)
8	RA1, RA2	2.4. Proyecto Intermedio - Aplicación de estadística en proyectos de datos - Avance de proyecto bajo metodología CRISP-DM	B1, B2, B3	
9	RA2, RA3	3.1. Regresión Lineal y Logística - Regresión lineal simple y múltiple - Regresión logística para clasificación	B1, B2, B3	
10	RA2, RA3	3.2. Modelos Avanzados de Regresión - Regresión polinomial - Árboles de regresión y Random Forest	B1, B2, B3	
11	RA2, RA3	3.3. Evaluación de Modelos - Métricas de rendimiento - Validación cruzada	B1, B2, B3	
12	RA2, RA3	3.4. Proyectos de Regresión - Desarrollo de modelos predictivos - Análisis de casos de estudio		Prueba P2(P2) (29 de mayo del 2024)
13	RA2, RA3; Error! No se encuentra el origen de la referencia.	4.1. Fundamentos de Clustering y Clasificación - Concepto de distancia y métodos de enlace - K-medoides, clustering espectral, K-means	B1, B2, B3	
14	RA2, RA3	4.2. Modelos de Clasificación Avanzados - K Nearest Neighbors (kNN) - Árboles de decisión y Random Forest	B1, B2, B3	
15	RA3	4.3. Introducción al Big Data - Almacenamiento de grandes cantidades de datos - Ecosistema Hadoop: conceptos básicos y herramientas	B1, B2, B3	Quiz 2 (Q2) (19 de junio del 2024)

16	RA3	4.4. Visualización Avanzada de Datos <ul style="list-style-type: none"> - Análisis y visualización de datos con Tableau/Power BI - Conexión y manipulación de datos para visualización avanzada - Técnicas de visualización de datos geoespaciales, jerárquicos, temporales, textuales. 	B1, B2, B3	
17		Periodo de exámenes		Exposición de un proyecto de aplicación (3 de Julio del 2024)
18		Periodo de exámenes		Exposición de un proyecto de aplicación

(*) Para referencias de Resultados de Aprendizaje y Recursos, consultar el programa del curso.