

Visualización de Datos

Código: 104365
Créditos ECTS: 6

Titulación	Tipo	Curso	Semestre
2503758 Ingeniería de Datos	OB	3	2

La metodología docente y la evaluación propuestas en la guía pueden experimentar alguna modificación en función de las restricciones a la presencialidad que impongan las autoridades sanitarias.

Contacto

Nombre: Enric Martí Godia
Correo electrónico: Enric.Marti@uab.cat

Uso de idiomas

Lengua vehicular mayoritaria: catalán (cat)
Algún grupo íntegramente en inglés: No
Algún grupo íntegramente en catalán: Sí
Algún grupo íntegramente en español: No

Equipo docente

Guillermo Marin Getino
Judit Chamorro Servent

Prerequisitos

Se recomienda que el alumnado tenga o repase conocimientos y conceptos de:

- Análisis de grafos y búsqueda de información.
- Ajuste de modelos y optimización.
- Aprendizaje computacional.
- Análisis estadístico.

Objetivos y contextualización

L'assignatura pretén donar al alumne elements per calcular la millor tècnica de visualització de dades que li permet un millor anàlisi de les dades, o li permet un tipus d'anàlisi específic.

CONOCIMIENTOS: Al finalitzar el curs el alumnado ha de ser capaç de:

- Conocer y comprender los sistemas básicos de visualización y de análisis exploratorio de datos.
- Conocer y comprender los principales modelos de visualización.
- Conocer y comprender las librerías actuales para la visualización de datos (Google Charts, JQuery plug-ins, D3.js)

HABILIDADES: Se pretende que el alumnado adquieran las siguientes habilidades:

- Dado un conjunto de datos, saber utilizar el sistema y modelo de visualización más conveniente para la visualización de datos que permita el análisis que se pretende hacer.
- Dado un conjunto de datos, saber utilizar una de las librerías de visualización de datos para obtener los mejores resultados gráficos.

RESULTADOS DE APRENDIZAJE:

1. Conocer y comprender los sistemas básicos de visualización y de análisis exploratorio de datos.
2. Conocer y comprender los principales modelos de visualización.
3. Conocer y comprender las librerías actuales para la visualización de datos (Google Charts, JQuery plug-ins, D3.js)
4. Saber aplicar las técnicas de visualización de datos más óptimas para el análisis de datos a partir de un caso concreto.
5. Trabajar cooperativamente.

Competencias

- Buscar, seleccionar y gestionar de manera responsable la información y el conocimiento.
- Comprender las técnicas de visualización de datos masivos y ser capaces de seleccionar la más adecuada para su análisis.
- Que los estudiantes tengan la capacidad de reunir e interpretar datos relevantes (normalmente dentro de su área de estudio) para emitir juicios que incluyan una reflexión sobre temas relevantes de índole social, científica o ética.

Resultados de aprendizaje

1. Buscar, seleccionar y gestionar de manera responsable la información y el conocimiento.
2. Identificar las necesidades de usabilidad e interactividad de un método de visualización de datos y ser capaz de elaborar una nueva versión de la visualización que mejore dichos aspectos.
3. Que los estudiantes tengan la capacidad de reunir e interpretar datos relevantes (normalmente dentro de su área de estudio) para emitir juicios que incluyan una reflexión sobre temas relevantes de índole social, científica o ética.
4. Relacionar las técnicas estadísticas y de inteligencia artificial con las visualizaciones que mejor expresan la información subyacente.

Contenido

1.

1. Introducción a la Visualización (2 horas)	1. C
	1. F
	2. A
2. Codificación Visual (2 horas)
 1. Tipos de datos y tipos de gráficos
 2. Canales visuales
 3. Escoger gráficos según su función
3. Percepción visual y diseño (2 horas)
 1. Percepción visual
 2. Fundamentos de diseño
 3. Tipografía
 4. Color
4. Precisión visual (2 horas)
 1. Errores comunes en la visualización
 2. Honestidad visual
 3. Buenas y malas prácticas en estadística
5. Tratamiento de datos para visualización (4 horas)

1. Uncertainty y Error
2. Reducción de dimensionalidad
3. Transformaciones
4. Data massatge
5. Cálculo y selección de métricas relevantes
6. Sistemas avanzados de visualización (4 horas)
 1. Múltiples variables y múltiples dimensiones
 2. Redes
 3. Datos 3D
 4. Campos de vectores
 5. Visualización científica
 6. Mapas
7. Interacción. (4 horas)
 1. Animación
 2. Dashboards
8. Metodologías de trabajo (2 horas)
 1. Storytelling con datos
 2. Estrategias de planificación
 3. Diseño centrado en l'usuario
9. Validación (2 horas)
 1. Usabilidad

Metodología

PLATAFORMA VIRTUAL: Caronte (<http://caronte.uab.cat>, grado de Enginyeria de Dades (Ingeniería de Datos), asignatura Visualització de Dades (Visualización de Datos), curso VD *Curs 2020-21*).

El proceso de aprendizaje del alumnado se fundamenta en dos tipos de actividades presenciales (Teoría y Seminarios). La asistencia a las clases de teoría y de seminarios NO ES OBLIGATORIA, pero SI MUY RECOMENDABLE.

TEORIA

Las clases de teoría se imparten mediante clases magistrales con transparencias y pizarra. Todo el material del curso será accesible a través de la Plataforma Virtual. Algunas de las sesiones de teoría se dedicarán a resolver ejercicios planteados en la misma aula con el objetivo de profundizar en los conceptos más relevantes y enseñar al alumnado a identificar y analizar los errores típicos.

SEMINARIOS

En las sesiones de seminarios alternaremos dos actividades: problemas (7 sesiones) i prácticas (6 sesiones).

En las sesiones de PROBLEMAS propondremos ejercicios a resolver sobre las siguientes temáticas:

- Gráficas simples (barras, líneas, boxplots, scatter, etc.) - Herramienta: *R*
- Gráficas complejas (coordenadas paralelas, small múltiples, xarxes, etc) - Herramienta: *R*
- Otras herramientas (comerciales y libres) - Herramienta: *Tableau Public, DataWrapper, Gephi*
- Visualización de redes
- Visualización de mapas
- Gráficas de datos multivariable y multidimensionales
- Gráficas animadas e interactivas - Herramientas: *R, visor 3D*

- Presentación (Títulos, escalas, leyendas y anotaciones) - Herramientas: *R* i *software de diseño como Inkscape*
- Dashboards - Eines: *Tableau* vs. *R*

En estas sesiones de problemas se proponen los ejercicios a resolver, que el alumnado realizará con la tutorización del professor. Toda la documentación (enunciados scripts, resultados de los problemas) de los seminarios está disponible en la Plataforma Virtual.

En las sesiones de PRÁCTICAS se pretende capacitar al alumnado a estudiar i definir las diferentes visualizaciones de conjuntos de datos y valorar las ventajas e inconvenientes. Propondremos dos trabajos prácticos:

1. Tratamiento con un data set predeterminado. Dado un data set proporcionado por los profesores: limpiar, preparar y visualizar con gráficas exploratorias, justificando cuál es la mejor.
2. Tratamiento con un data set libre. A partir de un data set escogido o generado por el alumnado: limpiar, preparar y visualitzar con gráficas exploratorias, justificando cuál es la mejor.

Se realizarán en grupos de TRES personas i para cada alumne o alumna se requerirá una dedicación de 12 horas en 6 sesiones tuteladas por professor y 33 horas de trabajo individual o grupal de preparación de las prácticas, trabajo sin tutela del professor y elaboración de informes.

La asistencia a las sesiones de prácticas tuteladas por profesor ÉS OBLIGATÒRIA. Se estudiarán casos de ausencia a las sesiones de prácticas debidamente justificadas.

El alumnado prepara el treball de pràctiques explicat en l'enunciat de pràctiques. Esta preparación ha de ser previa a la asistencia a la sessió de prácticas correspondiente. El trabajo se realiza asistiendo a las sesiones de prácticas en los horarios establecidos.

La normativa de entrega y evaluació de prácticas la podéis encontrar en el documento de normativa disponible en la Plataforma Virtual. Se evaluará el trabajo de prácticas realizado y se hará un examen individual en la entrega de cada práctica.

Nota: se reservarán 15 minutos de una clase dentro del calendario establecido por el centro o por la titulación para que el alumnado rellene las encuestas de evaluación de la actuación del profesorado y de evaluación de la asignatura o módulo.

Actividades

Título	Horas	ECTS	Resultados de aprendizaje
Tipo: Dirigidas			
Clases magistrales	26	1,04	2, 3, 4
Explicación y resolución de problemas	14	0,56	1, 2, 4
Resolución de ejercicios prácticos	12	0,48	1, 2, 4
Tipo: Supervisadas			
Casos prácticos	4	0,16	1, 2, 4
Tipo: Autónomas			
Estudio individual	26	1,04	2, 3, 4
Explicación y resolución de problemas	30	1,2	1, 2, 4
Resolución de ejercicios prácticos	30	1,2	1, 2, 4

Evaluación

TEORIA

Habrà dos exàmenes individuales por escritot, con un peso del 50% cada uno sobre la Nota de Teoria. La primera prueba (Par1) se harà aproximadamente a la mitad del semestre y evaluarà los conceptos teóricos explicados hasta la fecha y la capacidad de resolución de problemas de visualización parecidos a los trabajados en las sesiones de seminarios:problemas. La segunda prueba (Par2) se realizará al final del semestre y avaluarà los conceptos explicados después del primer parcial i hasta el momento y la capacida de resolución de problemas parecidos a los tranajados en las sesiones de seminarios:problemas.

Examen de recuperación: En caso de que la nota de teoria no llegue al aprobado, el alumnado se podrá presentar a un examen de recuperación sobre los contenidos del examen parcial suspendido.

PRÀCTICA

La nota del trabajo pràctic (Pract) realizado en una sesión de seminarios tiene una puntuació de un 1 punto que es sumará al examen del primer parcial. Tiene dos indicadores:

- Nota Grupal: (nota, PG) Entrega de un informe en grupo. Puntuación: 0,75 punts
- Nota Individual (nota PI). Examen el dia de l'examen del primer parcial: Puntuación: 0,25 puntos.

PROYECTO

La nota del proyecto realizado en sesiones de seminarios tiene una puntuació de 10 puntos. Tiene dos indicadores ponderados:

- Nota Primera Entrega: (nota, Prj_1) Entrega de un informe en grupo. Ponderación: 40%
- Nota Segunda Entrega (nota Prj_2). Memoria y presentación oral del proyecto. Ponderación: 60%.

Recuperación: Se estableceà una fecha para la recuperación del proyecti, tanto para la primera como la segunda entrega.

INDICADORES DE CALIFICACIÓN

La nota final de la asignatura se calcula de la sigüiente manera (entre parèntesie quadrado las notes mínimas para hacer media):

- $Pract = PG + PI$
- $Par1 = ExPar1 + Pract$ [ExPar1 ≥ 4] --- $Par2 = ExPar2$
- $Nota\ Teoria = 0,4 * (Par1) + 0,6 * (Par2)$ [Par1, Par2 ≥ 5]
- $Nota\ Proyecto = 0,4 * Prj_1 + 0,6 * Pr_2$ [Pr_1, Pr_2 ≥ 5]

$$NOTA\ FINAL\ ASIGNATURA = 0'5 * Nota\ Teoria + 0'5 * Nota\ Pràcticas$$

CRITERIOS D'AVUACIÓN

- Para poder calcular la NOTA FINAL ASIGNATURA, se deberán superar todas les actividades evaluables (exàmenes de Teoria, exámenes y entregas de pràcticas) amb un 5.
- PERMANENCIA EN LOS EXAMENS: En cualquier examen el alumnado no podrá marchar del aula durante los primeros 45 minutos del tiempo de examen, excepto por causas justificadas o por indicación del profesorado.
- En todas les recuperaciones de actividades evaluables (exámenesparciales o de prácticas, proyecto VD, etc.),la nota máxima será 8.

- **MEJORAR LA NOTA DE EXAMEN:** El alumnado se puede presentar a una recuperación para mejorar la nota de un examen. Si entrega el examen significa que renuncia a la nota anterior obtenida. En caso de que no entregue significa que acepta la nota anterior obtenida.
- **NO EVALUABLE:** Un alumno o alumna se considera No Avaluable (NA) únicamente si no ha hecho ninguna actividad de evaluación. Recordamos que la nota No Avaluable también corre convocatoria.
- **SUSPÈS:** En el caso de no conseguir al mínimo exigido en alguna de las actividades de evaluación, si el cálculo de la nota final es igual o superior a 5, se pondrá un 4 de nota en el expediente.
- **MATRÍCULA DE HONOR:** Podrá obtener una MH la alumna o alumno que tenga una nota igual o superior a los 9 puntos, a criterio del profesor. Como que el número de MH no puede superar el 5% del alumnado matriculado, se concederán a aquellos o aquellas que tengan las notas finales más altas, a criterio del profesor. Se concederá al alumnado que haya realizado todas las pruebas evaluativas en primera convocatoria.
- **REPETIDORES:** No se convalida ninguna parte aprobada por separado (Teoría, exámenes y entregas de prácticas) de un curso académico a otro.

CALENDARIO DE EVALUACIÓN:

- Exámenes de Recuperación: Según calendario académico de l'Escola de Ingeniería.
- Exámenes y entregas de prácticas: Fechas en la normativa de prácticas, que se publicará en la Plataforma Virtual.

Las fechas de evaluación continua y de entrega de trabajos se publicarán en la Plataforma Virtual y pueden estar sujetos a cambios de programación por motivos de adaptación a posibles incidencias. Siempre se informará en la Plataforma Virtual sobre estos cambios ya que se entiende que éste es el medio habitual de intercambio de información entre profesores y alumnado fuera del aula.

Sin perjuicio de otras medidas disciplinarias que se estimen oportunas, y de acuerdo con la normativa académica vigente, las irregularidades cometidas por un estudiante que puedan conducir a una variación de la calificación se calificarán con un cero (0). Las actividades de evaluación calificadas de esta forma y por este procedimiento no serán recuperables. Si es necesario superar cualquiera de estas actividades de evaluación para aprobar la asignatura, esta asignatura quedará suspendida directamente, sin oportunidad de recuperarla en el mismo año académico. Estas irregularidades incluyen, entre otras:

- La copia total o parcial de una práctica, informe, o cualquier otra actividad de evaluación
- Dejar copiar.
- Presentar un trabajo de grupo no realizado íntegramente por los miembros del grupo.
- Presentar como propios materiales elaborados por un tercero, aunque sean traducciones o adaptaciones y, en general, trabajos con elementos no originales y exclusivos de la alumna o alumno.
- Tener dispositivos de comunicación (como teléfonos móviles, *smart watches*, etc.) accesibles durante las pruebas de evaluación teórico-prácticas individuales (exámenes).

En caso de no superar la asignatura debido a que alguna de las actividades de evaluación no llega a la nota mínima requerida, la nota numérica del expediente será el valor menor entre 4 y la media ponderada de las notas. Con las excepciones de que se otorgará la calificación de "no evaluable" a los estudiantes que no participen en ninguna de las actividades de evaluación, y de que la nota numérica del expediente será el valor menor entre 3.0 y la media ponderada de las notas en caso que el estudiante haya cometido irregularidades en un acto de evaluación (y por tanto no será posible el aprobado por compensación).

EN RESUMEN: copiar, dejar copiar o plagiar en cualquiera de las actividades de evaluación equivale a un SUSPENSO con nota inferior a 3,5.

Actividades de evaluación

Título	Peso	Horas	ECTS	Resultados de aprendizaje
Nota de Teoría (exámenes parciales)	50%	4	0,16	2, 3, 4

Bibliografía

MATERIAL DE LA ASIGNATURA (apuntes, transparencias, artículos): En la Plataforma Virtual.

BIBLIOGRAFIA BÁSICA:

- A. Cairo, The functional art, <http://www.thefunctionalart.com/2012/09/download-three-chapters-of-functional.html>, 2012 (darer accés: juliol 2020).
- T. Munzner, Visualization Analysis & Design, CRC Press, 2014.

BIBLIOGRAFIA COMPLEMENTARIA:

- S. Few, Information Dashboard Design: Displaying Data for At-a-glance Monitoring, 2nd edition, Analytics Press, 2013.
- A. Cairo, The truthful art. Data, charts and maps for communication, New Riders, 2016.
- A. Kirk, Data Visualization. A handbook for Data Driven Design, 2nd edition, SAGE Publications, 2019.
- C.O. Wilke, Fundamentals of Data Visualization. A Primer on Making Informative and Compelling Figures, O'Reilly Media, 2019.

Software

En esta asignatura utilizaremos el siguiente software:

- Software R: *R is 'GNU S', a freely available language and environment for statistical computing and graphics which provides a wide variety of statistical and graphical techniques*. Nos podemos descargar el entorno en el link: <https://cran.rstudio.com/>.
- RStudio: Se trata d'un framework que trabaja sobre R y Python. Se puede descargar de forma gratuita en diferentes plataformas (Windows, MacOS, Linux) en el link: <https://www.rstudio.com/products/rstudio/>.
- Tableau es una herramienta de análisis y visualización de datos potente utilizada en el área de *Business Intelligence*. Disponemos de una licencia educativa para que la pueda utilizar nuestro alumnado. También nos podem bajar una versión de demo gratuita en el link: <https://www.tableau.com/es-es/why-tableau/what-is-tableau>.

Todo el software que se utilizará és de libre acceso o se ofrecerá gratuïtament al alumnat con licencias educacionales.