

Análisis e Interpretación de Datos

MÁSTER UNIVERSITARIO EN ANÁLISIS Y VISUALIZACIÓN DE DATOS
MASIVOS /VISUAL ANALYTICS AND BIG DATA

Yeray Mezquita

Presentación de la asignatura

Presentación de la asignatura

Presentación del profesor: Yeray Mezquita

Formación académica:

- Doctor en Ingeniería Informática.
- Master en Sistemas Inteligentes.
- Ingeniero Informático.

Presentación de la asignatura

Presentación del profesor: Yeray Mezquita

Trayectoria académica:

- Profesor de: Informática aplicada a la Estadística, bases de datos, fundamentos de sistemas blockchain, y Análisis y tratamiento de datos, entre otros.
- Director de varios TFGs y TFM.
- Investigador y coautor de publicaciones científicas.

Presentación de la asignatura

Presentación del profesor: Yeray Mezquita

Trayectoria profesional:

- Sector energético, aviación...
- Ingeniero de datos.
- Consultor de tecnologías distribuidas disruptivas (aprendizaje federado, blockchain).

Presentación de la asignatura

Asignatura Obligatoria Máster (Oficial) universitario en Análisis y Visualización de Datos Masivos

6 créditos ECTS



Presentación de la asignatura

Contenido resumido

- ▶ EDA (Análisis exploratorio de datos).
- ▶ Conceptos de probabilidad (Eventos y Variables).
- ▶ Inferencia Estadística (Estimación y Contrastes).
- ▶ Análisis de regresión y extracción de componentes principales.



Word cloud from <https://elitedatascience.com/learn-statistics-for-data-science>

Presentación de la asignatura

Objetivos

- ▶ Entender los conceptos básicos de la estadística: muestreo, población, variable, etc.
- ▶ Identificar y aplicar los elementos mínimos necesarios para realizar un análisis exploratorio de datos.
- ▶ Aprender a aplicar todos los pasos necesarios que conlleva la realización de un razonamiento estadístico.
- ▶ Conocer los conceptos básicos de la probabilidad y la combinatoria como fundamento del análisis de datos.
- ▶ Estudiar y aplicar los conceptos y herramientas de la estadística inferencial.
- ▶ Aprender técnicas de análisis de datos que nos permitan conocer sus propiedades y establecer hipótesis sobre ellos.

Presentación de la asignatura

Sesiones de tutoría y consulta:

- ✓ Sesiones presenciales virtuales (15)
- ✓ Foro (48h máx para tener respuesta). Importante: colaboración entre alumnos.
- ✓ 1 clase de presentación de la asignatura
- ✓ 1 clases de laboratorio
- ✓ 1 clase de repaso para explicar cómo será el examen
- ✓ 3 clases de refuerzo- explicar la solución de actividades y laboratorios

Presentación de la asignatura

Contenidos

Bloque 1: Fundamentos de Estadística y Estadística Computacional.

- ▶ Tema 1. Introducción a la estadística. Análisis exploratorio y descriptivo de datos.
- ▶ Tema 2. Estadística Computacional. Herramientas.

Presentación de la asignatura

Contenidos

Bloque 2: Análisis de datos, medidas estadísticas y de regresión.

- ▶ Tema 4. Regresión y correlación.
- ▶ Tema 3. Medidas que resumen la información.

Bloque 3: Probabilidades y distribuciones.

- ▶ Tema 5. Probabilidad condicional y variables aleatorias.
- ▶ Tema 6. Distribución en el muestreo.
- ▶ Tema 7. Intervalos de confianza.

Presentación de la asignatura

Contenidos

Bloque 4: Modelos estadísticos y aplicaciones avanzadas.

- ▶ Tema 8. Contrastes de hipótesis.
- ▶ Tema 9. Regresión.
- ▶ Tema 10. Análisis de componentes principales.

Desarrollo de la asignatura

Software de apoyo



R y R Studio

Desarrollo de la asignatura

Modelo 4P

MU Análisis y Visualización de Datos Masivos Estructura del Proyecto - Entregables

1Q

Bases de Datos NoSQL

Equipo de Trabajo: **Gestión de Datos**

Entregables:

- Diseño de un esquema de base de datos NoSQL adaptado a las necesidades del proyecto.
- Implementación de la base de datos NoSQL (ejemplo: MongoDB, Cassandra, Neo4J).
- Documentación sobre las estructuras de datos y esquemas utilizados.



Ingeniería para el Procesado Masivo de Datos

Equipo de Trabajo: **Procesamiento de Datos**

Entregables:

- Desarrollo de un pipeline de procesamiento de datos utilizando herramientas como Apache Hadoop o Apache Spark.
- Integración del pipeline con la base de datos NoSQL.
- Documentación técnica del pipeline y guía de uso para los analistas.



Análisis e Interpretación de Datos

Equipo de Trabajo: **Análisis de Datos**

Entregables:

- Implementación de algoritmos para el análisis estadístico y exploratorio de los datos.
- Desarrollo de modelos predictivos utilizando técnicas de regresión.
- Reportes de análisis y conclusiones derivadas de los datos.



Técnicas de IA

Equipo de Trabajo: **Inteligencia Artificial**

Entregables:

- Creación de modelos de IA avanzados para tareas específicas (p.ej., clasificación, regresión, clustering).
- Integración de los modelos de IA en el pipeline de procesamiento de datos.
- Documentación sobre los modelos desarrollados y las métricas de rendimiento.



Visualización Interactiva de Datos

Equipo de Trabajo: **Visualización de Datos**

Entregables:

- Desarrollo de un dashboard interactivo para la visualización de datos (utilizando herramientas como Tableau, Power BI, o D3.js).
- Integración del dashboard con la base de datos y el pipeline de procesamiento.
- Manuales de usuario para la navegación y utilización del dashboard.
- Establecer mecanismos para recoger comentarios y experiencias de turistas y proveedores, utilizando esta información para adaptar y mejorar continuamente los servicios.



Entrega de Actividades

Actividades (8 puntos)

Transversal
4P (5)
(16/12/2024)

Grupal (3)
(31/01/2025)

Laboratorio (5 puntos)

(23/02/2025)



Asistencia

Test (1,5)

**Solo puntúan actividades
con notas por encima del
5 entregadas dentro de
plazo**

- ▶ Entrega en PDF
- ▶ Aula Virtual > Actividades
- ▶ Plagios (por mínimo que sea)
 - ▶ 1ª vez: 0 puntos
 - ▶ 2ª vez: asignatura suspendida

- Transversal: Estadística descriptiva: información a partir del análisis crítico de datos.
- Grupal: Definición de un problema estadístico: modelización y propuesta de soluciones.
- Laboratorio: Estadística inferencial: del análisis muestral a la predicción poblacional.

Evaluación

Evaluación continua (40% - 4 puntos)

Actividad	Tipo	Puntuación máxima
Transversal	Grupal	5.00
Grupal	Grupal	3.00
Laboratorio	Individual	5.00
Test (0.2x9)	Individual	2.00
TOTAL		15.0



Solo suman actividades con calificación mayor o igual a 5 puntos dentro de plazo

Satura en 10 puntos

Examen final (60% - 6 puntos)

Evaluación

Continua
(40%)

Actividades

Laboratorios

Asistencia

Test

Examen
(60%)

obligatorio

Asignatura aprobada si
Examen presencial ≥ 5
y Nota final ≥ 5

Convocatoria única
para evaluación
continua

Continua: 10
Examen: 4.5
SUSPENSO

Continua: 4
Examen: 5
Nota final: 4,6
SUSPENSO

Continua: 5
Examen: 5
Nota final: 5
APROBADO

Evaluación

Examen final

Desarrollos prácticos

Hacer cálculos, analizar resultados, redactar un informe, todo al nivel de lo que discutiremos en clases.

Test

Preguntas de teoría de tipo test.

Desarrollo de la asignatura

Metodología



Material audiovisual



Apuntes de la asignatura



Presentaciones de los temas



Bibliografía



Programación semanal



Foros



Tutores



Profesor

Desarrollo de la asignatura

Metodología



Foros

- ☆ [Dudas y comentarios generales](#)
- ☆ [Tema 1. Introducción a la estadística](#)
- ☆ [Tema 2. Estadística computacional](#)
- ☆ [Tema 3. Medidas que resumen la información](#)
- ☆ [Tema 4. Regresión y correlación](#)
- ☆ [Tema 5. Probabilidad condicional y variables aleatorias](#)
- ☆ [Tema 6. Distribución en el muestreo](#)
- ☆ [Tema 7. Intervalos de confianza](#)
- ☆ [Tema 8. Contrastes de hipótesis](#)
- ☆ [Tema 9. Regresión](#)
- ☆ [Tema 10. Análisis de componentes principales](#)
- ☆ [Fe de erratas](#)
- ☆ [Dudas Actividad Individual. Estadística descriptiva: Información a partir del análisis crítico de...](#)
- ☆ [Dudas Actividad Grupal. Definición de un problema estadístico: modelización y propuesta de ...](#)
- ☆ [Dudas laboratorio. Estadística Inferencial: del análisis muestral a la predicción poblacional](#)
- ☆ [Dudas ejercicios o tests](#)

Programación semanal

Semana	Tema		Actividad/Lab	Entrega
1 (04/11/2024-08/11/2024)	Introducción a la estadística	Presentación+1(1)		
2 (11/11/2024-15/11/2024)	Introducción a la estadística	1(2)		
3 (18/11/2024-22/11/2024)	Estadística computacional	2(1)		
4 (25/11/2024-29/11/2024)	Estadística computacional	2(2)		
5 (02/12/2024-06/12/2024)	Medidas que resumen la información	3+ Act Transversal		
6 (09/12/2024-13/12/2024)	Regresión y correlación	4(1)		
7 (16/12/2024-20/12/2024)	Regresión y correlación	4(2)		Transversal
8 (06/01/2025-10/01/2025)	Probabilidad condicional y variables aleatorias	5(1)+ Act. Grupal	Ref. Transversal	
9 (13/01/2025-17/01/2025)	Probabilidad condicional y variables aleatorias	5(2)		
10 (20/01/2025-24/01/2025)	Distribución en el muestreo	6		Act grupal
11 (27/01/2025-31/01/2025)	Intervalos de confianza	7	Ref. Act grupal	
12 (03/02/2025-07/02/2025)	Contrastes de hipótesis	8(1)+Laboratorio		
13 (10/02/2025-14/02/2025)	Contrastes de hipótesis	8(2)		
14 (17/02/2025-21/02/2025)	Regresión	9		Laboratorio
15 (24/02/2025-28/02/2025)	PCA	10	Ref. Laboratorio + exámen	



www.unir.net