



UNIVERSIDAD  
COMPLUTENSE  
MADRID

nticmaster

Máster de Formación Permanente en

# BIG DATA, DATA SCIENCE & INTELIGENCIA ARTIFICIAL

8<sup>a</sup>  
EDICIÓN

Facultad de Estudios Estadísticos  
Universidad Complutense de Madrid



# PROGRAMA

- Módulo . GNU/Linux y Git
- Módulo . Bases de Datos SQL
- Módulo . Business Intelligence con Tableau
- Módulo . Programación Python
- Módulo . Bases de datos NoSQL
- Módulo . Fundamentos de estadística
- Módulo . Minería de datos y modelización predictiva
- Módulo . Machine Learning
- Módulo . Visualización avanzada
- Módulo . Inteligencia Artificial con Deep Learning
- Módulo . Spark
- Módulo . Tecnologías del Big Data
- Módulo . Productivizar un modelo
- Módulo . Data Science aplicada a la empresa
- TFM . Trabajo Final de Máster



# La importancia del Big Data, Data Science e Inteligencia Artificial

## NECESIDAD DEL MERCADO

Las empresas y organismos están adoptando rápidamente la **transformación digital** para aprovechar el valor de los **datos masivos** en la toma de decisiones. El Big Data se ha convertido en un recurso valioso para la gestión empresarial, y lo que comenzó como una ventaja competitiva se ha vuelto esencial para mantenerse relevante. El dato es considerado como el petróleo del siglo XXI, y aquellos que puedan aprovechar su potencial estarán en una posición privilegiada para **innovar y liderar** en sus respectivos mercados.

## APLICACIONES PRÁCTICAS

El impacto de la información a gran escala va más allá del ámbito matemático o estadístico, ya que tiene aplicaciones prácticas en diversos campos empresariales, gubernamentales, científicos y sociales. El Big Data se ha convertido en un recurso fundamental para **afrentar situaciones complejas en tiempo real**, permitiendo tomar decisiones informadas y generar conocimiento valioso en áreas como la medicina, la seguridad, el marketing, entre otros. El Big Data es una herramienta imprescindible en el mundo actual que ayuda a resolver problemas y aprovechar oportunidades en diversos entornos y sectores.

## INNOVACIÓN IMPULSADA POR DATOS

El Big Data no solo ayuda a las empresas a ser más eficientes, sino que también impulsa la innovación. Al analizar grandes volúmenes de datos, las organizaciones pueden **descubrir patrones ocultos, nuevas oportunidades de negocio y mejorar procesos**. Esto les permite crear productos y servicios más competitivos y alineados con las demandas del mercado actual.

## MACRODATOS

El Big Data es una herramienta valiosa en diversas áreas, desde la creación de informes estadísticos hasta la identificación de nuevas oportunidades de negocio, pasando por la predicción de resultados publicitarios y la investigación médica. El análisis de grandes volúmenes de datos permite obtener información relevante y tomar **decisiones informadas en tiempo real**, lo que resulta en una ventaja competitiva para las empresas y avances significativos en la investigación científica.

## MEJORA EN LA EXPERIENCIA DEL CLIENTE

Las herramientas de Big Data permiten a las organizaciones personalizar la experiencia del cliente de manera más precisa. Al analizar patrones de comportamiento y preferencias de los usuarios, las empresas pueden ofrecer **productos y servicios más alineados con las expectativas de sus consumidores**, mejorando la satisfacción y la lealtad.



# ¿Por qué estudiar un Máster de Big Data?

Dentro del mundo empresarial, el análisis de grandes volúmenes de datos se ha vuelto indispensable para la toma de decisiones estratégicas. El Big Data permite a las organizaciones transformar datos en información valiosa para predecir tendencias, optimizar procesos y mejorar el rendimiento. Estudiar un máster en Big Data te preparará para dominar estas técnicas y herramientas, y te permitirá convertirte en un profesional clave en un campo de alta demanda.

## Alta Demanda Laboral

El campo del Big Data está en constante expansión. Las empresas de todos los sectores buscan profesionales capaces de gestionar, analizar y transformar grandes volúmenes de datos en información valiosa para la toma de decisiones. Especializarte en Big Data te posicionará en un mercado con alta demanda y numerosas oportunidades.

## Oportunidades Profesionales

Las salidas profesionales en Big Data son muy diversas, abarcando roles como analista de datos, ingeniero de datos, arquitecto de datos y científico de datos, entre otros. Un máster te preparará para enfrentar estos desafíos y desarrollar habilidades técnicas avanzadas.

## Networking

Estudiar un máster en Big Data te brindará la oportunidad de conectar con profesionales del sector, compañeros de clase, profesores expertos y empresas. Este entorno te permitirá generar relaciones clave para tu desarrollo profesional y acceder a oportunidades laborales.

## Mayor Potencial de Ingresos

Los profesionales en Big Data suelen tener un alto potencial de ingresos debido a la demanda de personas con habilidades técnicas avanzadas. Las empresas están dispuestas a pagar más por aquellos que puedan transformar datos en información estratégica que impulse la toma de decisiones.

## Actualización Constante

El campo del Big Data está en continua evolución con nuevas herramientas, tecnologías y enfoques. Un máster te permitirá mantenerte actualizado en las últimas novedades y aprender a adaptarte rápidamente a las tendencias emergentes en la industria de los datos.

## Desarrollo de Proyectos Reales

Un máster en Big Data te ofrecerá la oportunidad de trabajar en proyectos reales, aplicando tus conocimientos a situaciones del mundo empresarial. Estas experiencias prácticas serán clave para fortalecer tus habilidades y mejorar tu empleabilidad.

## Ventaja Competitiva

Tener un máster en Big Data te dará una ventaja competitiva en el mercado laboral, diferenciándote como un profesional preparado para abordar los complejos desafíos de la gestión y análisis de grandes volúmenes de datos.

## Flexibilidad Laboral

El conocimiento en Big Data ofrece flexibilidad en el tipo de empresas para las que puedes trabajar, desde startups hasta grandes corporaciones. Además, muchos roles relacionados con Big Data permiten trabajar de forma remota, lo que te da mayor control sobre tu vida profesional.

## Emprendimiento

Con los conocimientos adquiridos en un máster en Big Data, tendrás las herramientas necesarias para lanzar proyectos propios en el ámbito de la tecnología y los datos, lo que te permitirá explorar nuevas oportunidades de negocio.



## Duración

1 año académico

## Modalidades

Presencial, Semipresencial y Online

## Creditos ECTS

60

## Modalidad Presencial

Viernes tarde y sábados  
mañana en la universidad

## Modalidad Semipresencial

3 Semanas en presencial  
en la universidad y en la  
plataforma online

## Modalidad Online

100% desde nuestra  
plataforma online



# ¿Por qué estudiar en la Universidad Complutense de Madrid?

La Universidad Complutense de Madrid (UCM) es una de las instituciones educativas más destacadas de Europa, reconocida por el prestigioso QS World Ranking como la mejor de España. Ofrece una amplia gama de oportunidades y beneficios para los estudiantes, así como una excelencia académica reconocida, una calidad docente de primer nivel. Ofrece alrededor de 90 títulos de grado y más de 30 dobles grados, más de 200 programas máster, además de estudios de formación permanente. La UCM tiene más de 500 años de historia y reconocimiento social. La Universidad Complutense de Madrid es la universidad española de referencia en 5 continentes.

El prestigio de la universidad está avalado por 7 Premios Nobel, 20 Príncipes de Asturias, 7 Premios Cervantes, Premios Nacionales de Investigación y a la Excelencia. La Universidad Complutense de Madrid tiene estudiantes de más de 90 países y convenios con universidades de los 5 continentes.





# ¿Por qué estudiar un Máster en Formación Permanente?



*Si hay algo que afianza los conceptos teóricos de un programa educativo es la práctica.*

Nuestros módulos formativos combinan una base teórica con ejercicios prácticos basados en situaciones reales de las empresas. Además, todos los módulos se evalúan con tareas prácticas, no con exámenes, tratándose de un programa de configuración eminentemente práctica.

La preparación del Trabajo Final de Máster (TFM) garantiza la puesta en práctica de todos los conceptos adquiridos a lo largo del curso, capacitando definitivamente al alumno para asumir responsabilidades dentro de un entorno laboral real.



# PROGRAMA

*Los módulos de aprendizaje de Big Data más completos enseñados de forma eficaz para el alumno.*



UNIVERSIDAD  
COMPLUTENSE  
MADRID

 nticmaster





# MÓDULO

## GNU/Linux y Git

En este módulo se trabajará el entorno de trabajo con **Linux y Git**, conocimiento esencial para preparar a los estudiantes en el uso de herramientas fundamentales para la **gestión de proyectos y el análisis de datos a gran escala**.

Linux es un sistema operativo de código abierto que se utiliza ampliamente en la industria de Big Data debido a su flexibilidad, escalabilidad y capacidad para manejar grandes cantidades de datos. A través de este módulo, los estudiantes aprenderán a trabajar con la línea de comandos de Linux, lo que les permitirá realizar **tareas avanzadas de programación y análisis de datos**.

Git es un sistema de control de versiones que se utiliza para administrar el código fuente de los proyectos de software. Los estudiantes aprenderán cómo usar Git para **colaborar en equipo, compartir código y hacer un seguimiento de los cambios en el código** a medida que se realiza la edición y mejora del software.

### Índice de contenidos:

- GNU/Linux + Git (1 semana).
- Instalación VirtualBox (Windows/ Mac).
- Instalación y configuración máquina virtual GNU/Linux.
- Interfaz gráfica vs Terminal.
- Principales comandos del sistema GitHub, GitLab y Bitbucket:
- Creación de repositorios.
- Documentación del proyecto (Markdown).
- Instalación y configuración de Git:
  - Comandos básicos en Git.
  - GitFlow - Trabajando con ramas.





# MÓDULO

## Bases de Datos SQL

**Diseño y modelización de base de datos y lenguaje de consulta estructurada**, más conocido como SQL (Structured Query Language). Estudio del modelo entidad-relación, modelo relacional, e implementación relacional con SQL.

En éste módulo de base de datos SQL, se estudiarán los fundamentos y conceptos necesarios de los **sistemas de gestión de bases de datos relacionales** y cómo se pueden implementar y administrar utilizando el lenguaje de consulta estructurado (SQL).

Los estudiantes tendrán como objetivo general en este módulo **adquirir los conceptos fundamentales de las bases de datos y sus técnicas básicas** de diseño, gestión y explotación, haciendo hincapié en el modelo de bases de datos relacionales.

Se realizarán **prácticas** para asentar correctamente estos conocimientos, pues se trata del modelo de mayor implantación en la actualidad, y sus conceptos y técnicas están presentes en todos los demás.

Podemos afirmar que las bases de datos SQL son el punto de partida para el manejo de volúmenes de datos, ya sean pequeños o grandes, y por tanto una parte crucial en la iniciación del proceso que lleva a sacar el máximo rendimiento a la inteligencia de datos para ponerla al servicio de un objetivo concreto.

### Índice de contenidos:

- **Introducción a las bases de datos**, sus sistemas de gestión y ventajas. Se profundizará en qué es una base de datos, su evolución y tipos.
- El **modelo entidad/relación** en el diseño de las bases de datos, incluyendo el modelado conceptual de datos, las entidades, especialización y generalización. Transformación del modelo entidad/relación a tablas.
- **Conceptos fundamentales del modelo relacional** como relaciones, claves y restricciones de integridad, así como normalización.
- **Implementación relacional con SQL**: lenguaje de definición de datos (DDL) y lenguaje de manipulación de datos (DML). Consultas simples, subconsultas, JOIN, EXIST y NOT EXIST, HAVING Y GROUP BY, UNION, INTERSECT Y EXCEPT, INSERT, UPDATE Y DELETE. Vistas, trigger, procedimientos almacenados y funciones.

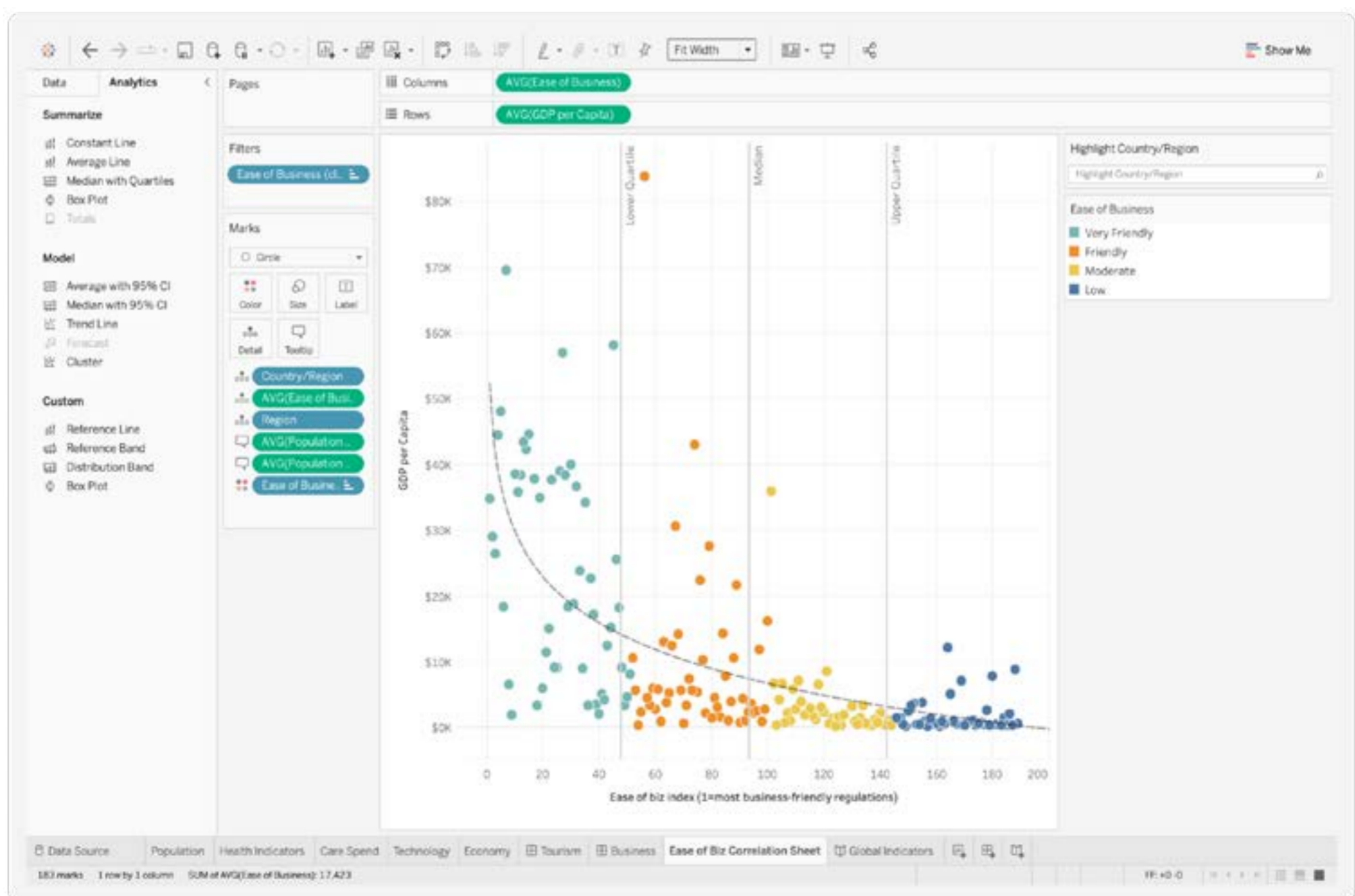
# MÓDULO

## Business Intelligence con Tableau

Este módulo plantea diferentes objetivos al alumno, con los que espera **mejorar sus capacidades analíticas**, así como sus habilidades para desenvolverse en un **entorno empresarial** dentro del ámbito del “Business Intelligence”.

El módulo introduce al alumno al concepto de Business Intelligence, diferenciando éste, del concepto de Machine Learning o de Data Science, su relación con los nuevos paradigmas de Big Data. Para pasar a un **enfoque completamente práctico** en el que el alumno aprenderá a utilizar Tableau. Se acompañará al alumno en el proceso de **descubrimiento de claves** (insights) aplicado sobre un conjunto de datos abiertos. Tableau es líder en el cuadrante de visionarios de Gartner en las plataformas de BI en los últimos años. Con esta introducción el alumno podrá realizar **análisis básicos** usando esta solución que complementarán al aprendizaje de otros lenguajes y técnicas analíticas de este Máster.

Como complemento de esta vertiente técnica, el alumno aprenderá otros **conceptos/habilidades más orientados a cómo desenvolverse en un entorno empresarial orientado a la analítica avanzada**. Uno de estos conceptos se centra en cómo han de usarse de forma efectiva y eficiente diferentes tipos de gráficos, y cómo preparar y comunicar de forma eficiente los resultados de un análisis de datos a una audiencia no-técnica, de negocio. Para los analistas de negocio estos dos elementos se están considerando como esenciales en las organizaciones.





# MÓDULO

## Programación Python

En este módulo los alumnos profundizarán en el manejo del lenguaje de programación Python.

Python es un lenguaje de programación ampliamente utilizado en el entorno de Big Data debido a su **facilidad de uso, flexibilidad y capacidad para trabajar con grandes cantidades de datos**.

Tiene una gran cantidad de **librerías especializadas** para el análisis de datos, que proporcionan herramientas para la manipulación, procesamiento, análisis y visualización de datos a gran escala.

Además, Python puede integrarse fácilmente con estas tecnologías para procesar grandes cantidades de datos de manera eficiente.

Veremos sus características, tipos de datos, estructuras de control de flujo, funciones, parámetros, manipulación de cadenas, estructuras de datos, etc.

Otro atractivo con el que cuenta este módulo es su **enfoque práctico**. Nuestro programa no solo incluye la teoría necesaria, también la visión práctica para experimentar sus técnicas básicas en una amplia gama de aplicaciones.

### Índice de contenidos:

Veremos la introducción a la programación con Python y conceptos básicos como: variables, instrucciones generales, tipos de datos y operaciones.

Estos conocimientos impartidos nos permitirá a continuación abordar otros aspectos como:

- Entrada y salida.
- Estructuras de control: selección, iterativas. Funciones. Recursividad. Orden superior en Python. Expresiones lambda.
- Estructuras de datos fundamentales: colecciones, listas, tuplas, conjuntos, diccionarios. Programación orientada a objetos.
- Librerías para el procesamiento numérico y estadístico y para el análisis de datos.

```
0 # load from the website
1 # checking response.status_code (if you get 502, try rerunning the code)
2 if response.status_code != 200:
3     print(f"Status: {response.status_code} - Try rerunning the code!")
4 else:
5     print(f"Status: {response.status_code}\n")
6
7 # using BeautifulSoup to parse the response object
8 soup = BeautifulSoup(response.content, "html.parser")
9
10 # finding Post images in the soup
11 images = soup.find_all("img", attrs={"alt": "Post image"})
12
13 # downloading images
14 for i in range(len(images)):
15     img = images[i]
16     url = img.get("src")
17     r = requests.get(url)
18     with open(f"image_{i}.jpg", "wb") as f:
19         f.write(r.content)
```





# MÓDULO

## Bases de Datos NoSQL

En éste módulo se aprenderá a **utilizar y modelar los sistemas de gestión de bases de datos NoSQL** y sus principales operaciones. Introducción a MongoDB, operaciones CRUD, dominar el Find o proyectar los campos en resultados de búsqueda.

Cuando hablamos de bases de datos NoSQL, nos estamos refiriendo a aquellas cuya característica más destacable es que no requieren de estructuras fijas como tablas, a diferencia de las bases de datos tradicionales que requieren de SQL como lenguaje principal de consultas. Sin embargo, esto no significa que no soporten dicho lenguaje SQL.

### Índice de Contenidos

Introducción a las **bases de datos NoSQL**:

- Un poco de historia sobre las BD relacionales.
- Modelo de datos (entidad relación vs agregación).
- BD Relacionales VS NoSQL (ACID vs BASE).
- Diferentes tipos de BD NoSQL (key-value, documentales, XML, Grafos, Column Family).
- Cuando utilizar NoSQL.
- Teorema del CAP.
- Mongo VS Cassandra.

Instalar **MongoDB**:

- Cómo instalar servidor MongoDB en WIN/Mac/Linux.
- Cómo conectarse a mongo con diferentes clientes.
- Conceptos de JSON y Arrays.
- Manejo básico en la base de datos/colecciones y tipos de datos.
- Importación y Exportación de ficheros.
- Operaciones CRUD.

Create – insert, insertOne, insertMany:

- Read – find, findOne.
- Update – update, updateOne, updateMany, findAndModify y operadores.
- Delete – remove, deleteOne, deleteMany.

Dominar el **Find**:

- Aplicar filtros avanzados.
- Arrays y Documentos Embebidos y operadores.
- Ordenar los resultados de búsqueda.
- Paginar y limitar los resultados de búsqueda.
- Otros cursores importantes.

**Índices y Agregaciones:**

- Índices.
- Agregaciones, pipeline, etapas, operadores.

**Modelo de datos:**

- Cómo modelar.
- Denormalización.



# MÓDULO

## Fundamentos de Estadística

La estadística permite **sistematizar la información proporcionada por las bases de datos** cuya procedencia, como sabemos, puede ser muy variada: datos económicos de empresas, datos sobre la realidad social, datos sobre la economía, sobre deportes, finanzas, opinión pública y política, etc. El tratamiento estadístico de los datos, con frecuencia, no acaba en la obtención de un resumen de la información (Estadística Descriptiva) sino que sirve para **tomar decisiones** en múltiples ámbitos, **resolver diversos tipos de problemas** y **generar modelos predictivos**.

Gracias a este módulo los alumnos serán capaces de manejar las herramientas que les permiten la **estimación de los parámetros de los modelos**, los **contrastos de hipótesis** sobre dichos parámetros, así como el acceso a la metodología que permite el rigor en la toma de decisiones en Estadística.

Se persigue la consolidación de los conocimientos de estadística necesarios para abordar el resto de las áreas del máster.

El módulo está dividido en dos temas:

### Estadística descriptiva:

Se centra en la **descripción de las variables estadísticas univariantes**, que estudian una característica en particular del objeto de estudio, pueden ser numéricas o de otra índole (como por ejemplo sexo, nivel de estudios o sector profesional).

Se profundizará en diversos tipos de medidas estadísticas de centralización, dispersión, asimetría y curtosis. Además, se tratarán temas como:

- Descripción de variables bidimensionales. A diferencia de las univariantes, tienen en cuenta dos caracteres del mismo sujeto de estudio y las posibles relaciones entre dos objetos distintos.
- Análisis de la vinculación.
- Medidas de asociación.
- Regresión.

### Inferencia estadística:

Cuando hablamos de inferencia en el contexto estadístico, nos estamos refiriendo a las diversas técnicas y metodologías a través de las cuales, en base a una información con la que previamente contamos (los **datos**), se realizan **modelos de predicción**. Además es muy relevante **controlar el margen de error en las decisiones**, margen que necesariamente existe dada la aleatoriedad intrínseca en los fenómenos que analiza la estadística.

Un ejemplo muy cotidiano al respecto son las encuestas electorales de intención de voto, donde con una base de encuestados que busca ser lo más heterogénea posible, se intenta determinar cuál será el reparto de votos entre los distintos partidos políticos que se presentan a unas elecciones. Pero también el análisis del riesgo a la hora de conceder un crédito personal o hipotecario, el uso de datos para el diagnóstico diferencial en medicina preventiva, o la información que recibe un portero de fútbol sobre las habilidades y tendencias del delantero que le tirará el penalti.

Algunos de los puntos que se tratarán en este módulo en relación a la inferencia estadística serán:

- Variables aleatorias.
- Modelos de distribución de probabilidad.
- Estimulación puntual de parámetros.
- Estimulación por intervalos de confianza.
- Contrastos paramétricos.
- Contrastos no paramétricos.



# MÓDULO

## Minería de Datos y Modelización Predictiva

En este módulo los alumnos aprenderán técnicas de minería para **descubrir patrones, tendencias y relaciones útiles y significativas en grandes conjuntos de datos utilizando técnicas avanzadas de análisis** para explorar grandes cantidades de datos y extraer información valiosa de ellos. Esta información puede ser utilizada para tomar decisiones informadas, identificar oportunidades de negocio, mejorar la eficiencia de los procesos empresariales y optimización en diferentes industrias, entre otros.

Mediante la modelización predictiva, utilizarán los datos históricos y los patrones descubiertos a través de la minería de datos para predecir eventos futuros. La modelización predictiva utiliza **algoritmos estadísticos y de aprendizaje automático** para construir modelos que puedan predecir con precisión los resultados futuros que pueden ser utilizados para tomar decisiones informadas, identificar oportunidades de negocio, prevenir riesgos, entre otros.

Se estudiarán conceptos desde una variedad de temas esenciales en el campo, desde codificación, depuración, descripción e integridad de datos hasta técnicas avanzadas como regresión lineal y logística, reducción de dimensionalidad y análisis y predicción con series temporales.

Los estudiantes también aprenderán acerca de clasificación no supervisada y análisis de clusters, y cómo construir scorecards y modelos de gestión efectivos utilizando datos. Además, se explorará la estadística geoespacial, lo que permitirá a los estudiantes aplicar técnicas avanzadas de análisis espacial a datos geográficos.

A lo largo del curso, los estudiantes trabajarán en **ejemplos prácticos** para mejorar su comprensión de los conceptos y técnicas clave. Al finalizar el curso, los estudiantes estarán entrenados con habilidades esenciales en data mining para analizar, procesar y utilizar datos de manera efectiva para tomar decisiones y resolver problemas complejos.





# MÓDULO

## Machine Learning

Machine Learning se refiere al conjunto de **técnicas y algoritmos que permiten a los sistemas informáticos aprender y mejorar automáticamente a partir de los datos**, sin ser explícitamente programados.

El módulo se enfoca en introducir a los estudiantes en los conceptos de Machine Learning, incluyendo las diferentes técnicas y algoritmos utilizados, como la regresión lineal, la clasificación, el clustering y el aprendizaje profundo (Deep Learning). También se cubren temas como la evaluación y selección de modelos, el preprocesamiento de datos, la validación cruzada y el overfitting.

Los alumnos aprenderán diversas herramientas y aplicaciones de **Machine Learning y modelización predictiva avanzada**.

Estas técnicas de **aprendizaje automático** permiten a las máquinas ser capaces de asimilar una serie de comportamientos generalizados para realizar diversas acciones que toman, como ejemplo de referencia, las pautas indicadas previamente.

Es importante ser conscientes, y así se les hace saber a nuestros estudiantes, que el Machine Learning está presente en buena parte de elementos con los que interactuamos diariamente en nuestra vida cotidiana.

Teléfonos móviles o sistemas de navegación, por poner únicamente dos ejemplos, cuentan con este sistema de aprendizaje automático. Igualmente, aplicaciones de detección de fraudes, diagnósticos médicos o el propio buscador de Internet, que te devuelve resultados a una consulta, tienen presente el mismo.

### Índice de Contenidos

Debido a la importancia de este módulo dentro de los que componen el temario del máster, se repasarán a lo largo del mismo los siguientes puntos:

- Introducción.
- Redes neuronales y Deep Learning.
- Árboles de decisión.
- Random forest.
- Gradient Boosting.
- Support Vector Machines.
- Algoritmo KNN.

Dentro del aprendizaje automático, las tecnologías usadas para llevar a cabo proyectos de Machine Learning tienen igualmente y como es lógico un gran protagonismo.

Es por ello que el programa en este punto no se detendrá únicamente a dar a conocer a los alumnos las características más comunes del aprendizaje automático, sino que pondrá también en su mano las herramientas necesarias para poder **desarrollar su propio proyecto al respecto**.







# MÓDULO

## Visualización avanzada

El módulo de visualización tiene como objetivo **definir las claves para la eficacia en la presentación de información visual en el entorno del macrodato**. Cualquiera que sea la herramienta que se utilice, la visualización de datos es una herramienta crítica para la comunicación de la realidad latente.

En este módulo, los estudiantes aprenderán acerca de las visualizaciones para comunicar. Se explorarán herramientas para visualización de datos con librerías como **matplotlib** y **seaborn**, entre otras, y se discutirá la **gramática de los gráficos**, la cual es importante para la comprensión de cómo se construyen y comunican los gráficos.

Los estudiantes también aprenderán sobre los componentes básicos para la visualización, incluyendo el **uso del color** y **tipos de gráficos** para tipos de datos y propósitos concretos.

Además, se presentarán los **entornos python, notebooks** y **entornos virtuales**, los cuales se utilizarán en los ejemplos propuestos y en el trabajo con datos propios.

En la segunda semana, los estudiantes aprenderán sobre **gráficos interactivos** utilizando **Plotly**, así como las animaciones y la interactividad en los gráficos. También se presentarán otras herramientas como Bokeh y Pygal, y se explorarán dashboards en python, Dash y Glean.

Por último, los estudiantes aprenderán sobre mapas en python utilizando librerías como **Geoplotlib** y **Leaflet/Folium**. A través de ejemplos prácticos, los estudiantes podrán aplicar lo aprendido en el módulo a sus propios datos y visualizaciones de manera efectiva y eficiente.

### Índice de Contenidos

#### 1. Introducción: claves para la eficacia en la visualización de datos

- Cualquiera que sea la herramienta que se utilice.
- Visualización para la comunicación a lo largo de la historia.
- Visualizaciones para comunicar y visualizaciones para confundir.
- Herramientas que veremos en el módulo.
- Datos que usaremos.
- Entornos python: notebooks, entornos virtuales.

#### 2. Componentes básicos para la visualización

- El color.
- Tipos de gráficos para tipos de datos y propósitos concretos.

#### 3. Gráficos estáticos en Python

- Librería matplotlib.
- Librería seaborn.

#### 4. Gramática de los gráficos: ggplot

- Por qué una gramática de los gráficos.
- Capas.
- Librerías plotnine, ggplot2.
- Trabajo del módulo: ejemplos propuestos y trabajo con datos propios.

#### 5. Introducción a los gráficos interactivos: plotly

- Por qué y para qué interactividad.
- Animaciones.

#### 6. Bokeh y pygal

#### 7. Dashboards en Python

- Dash.
- Glean.

#### 8. Mapas en Python

- Geoplotlib.
- Leaflet / Folium.



# MÓDULO

## Inteligencia Artificial con Deep Learning

El Deep Learning es una rama de la inteligencia artificial que utiliza **redes neuronales artificiales para aprender y realizar tareas complejas**.

En este módulo los estudiantes aprenderán técnicas de Deep Learning, incluyendo la **arquitectura y funcionamiento de las redes neuronales**.

También se explorarán **aplicaciones prácticas** de la tecnología en diferentes áreas, como el procesamiento de imágenes con redes convolucionales, el procesamiento de secuencias con redes recurrentes, la predicción de series temporales, el procesamiento de lenguaje natural (NLP) y los modelos generativos, así como los últimos avances en deep learning (ChatGPT, DALL-E ...) poniendo el foco en sus aplicaciones comerciales.

### Índice de contenidos:

#### 1. Deep Learning. Introducción y fundamentos de las redes neuronales:

El deep learning difiere por su capacidad para trabajar con datos no estructurados, como imágenes y texto, y para aprender y mejorar con más datos utilizando una estructura jerárquica de capas de procesamiento que imita la forma en que el cerebro humano procesa la información. En este bloque los alumnos aprenderán técnicas para analizar y aprender patrones complejos en grandes conjuntos de datos. Se estudiará el modelo bio-inspirado, perceptrones, funciones de activación, Keras y Tensorflow 2 entre otros.

#### 2. Inteligencia Artificial. Procesamiento de imágenes con redes convolucionales (CNN):

Este bloque está diseñado para proporcionar una comprensión sólida de los principios fundamentales del procesamiento de imágenes con redes convolucionales, así como para proporcionar habilidades prácticas para aplicar estas técnicas en el mundo real. Los estudiantes aprenderán a diseñar y entrenar redes convolucionales para tareas de clasificación y segmentación de imágenes, y adquirirán una comprensión profunda de los enfoques más avanzados en el campo de las CNNs.

Se estudiará computer vision, convolución, capas (pooling, rectificación, flatten), transfer learning, entre otros.

#### 3. Inteligencia Artificial. Predicción de series temporales con redes recurrentes (RNN):

Este bloque está diseñado para proporcionar una comprensión sólida de los principios fundamentales del uso de RNNs para la predicción de series temporales, así como para proporcionar habilidades prácticas para aplicar estas técnicas en el mundo real. Los estudiantes aprenderán a diseñar y entrenar RNNs para realizar predicciones precisas de series de tiempo en diferentes aplicaciones. También adquirirán una comprensión profunda de los enfoques más avanzados y recientes en el campo de las RNNs

Se estudiará procesamiento de secuencias, introducción a las redes recurrentes y tipos de RNNs: LSTM, GRU, Bidirectionals.

#### 4. Inteligencia Artificial. Introducción al procesamiento de lenguaje natural (NLP):

Este bloque está diseñado para proporcionar una comprensión sólida de los principios fundamentales del procesamiento de lenguaje natural, como rama de la IA que se centra en la interacción entre los humanos y las máquinas mediante el lenguaje natural.

Se estudiará Words embeddings, aplicación del deep learning al NLP, modelos de secuencia a secuencia, modelos de atención y mecanismos de atención, transformers, últimos modelos y estado del arte (GPT3, chatGPT).

#### 5. Inteligencia Artificial. Modelos generativos:

Los estudiantes aprenderán a entrenar y evaluar modelos generativos en diferentes dominios, como imágenes, texto y sonido. También se explorarán las aplicaciones prácticas de los modelos generativos, como la generación de imágenes y la síntesis de voz. Entre otros: autoencoders, GANs., etc.



# MÓDULO

## Spark

En este módulo repasaremos las **tecnologías Big Data** y su motivación en el contexto actual de la era digital y las necesidades de las empresas. Proporciona a los estudiantes una comprensión profunda de cómo funcionan estos sistemas de procesamiento de datos distribuidos y cómo aprovecharlos para **procesar grandes cantidades de datos de manera eficiente** y efectiva.

Tras describir brevemente el **manejo de HDFS**, el curso se centrará en **Apache Spark**, sin duda la tecnología más demandada para procesamiento de grandes volúmenes de datos. Describiremos su filosofía peculiar basada en un grafo de ejecución (DAG) y profundizaremos en cada uno de los módulos, en especial Spark SQL, MLlib y Structured Streaming. Se usará el servicio **Dataproc** de Google Cloud para que cada alumno pueda desplegar un cluster de forma sencilla y gratuita.



### Índice de contenidos:

- Introducción a las tecnologías Big Data.
- HDFS.
- Apache Spark.
- Arquitectura de Spark.
- Spark SQL.
- Spark MLlib.
- Spark Streaming.
- GraphFrames.



# MÓDULO

## Tecnologías del Big Data

El alumno aprenderá a descubrir y valorar oportunidades presentes en el día a día del ciudadano y de la empresa.

Se verán temas como el **Internet de las cosas** como servicio, **conectividad** con fuentes de datos heterogéneas a través de brókers de mensajes y hubs con dispositivos, **Hadoop**, **Spark** y diversas técnicas de **visualización** y análisis de información por parte de los usuarios finales. También se tratará la aproximación a la aportación del IOT al mundo del macrodato y la conectividad con fuentes de datos heterogéneas a través de **bróker de mensajes** y **hubs** con dispositivos

Otros temas a tratar serán:

- ETL as a Service.
- Gobierno de los datos as a Service.
- Stream Analytics, CEP análisis en memoria en tiempo real en los eventos complejos as a Service.

### Índice de contenidos:

#### Hadoop:

- Introducción a Hadoop, entorno de trabajo que soporta diversas aplicaciones distribuidas bajo una licencia libre.
- Almacenamiento HDFS para grandes volúmenes de datos en Apache Hadoop.
- Acceder a los datos almacenados en HDFS como si fuera una base de datos: introducción a HIVE.

#### Spark:

- Introducción a Spark, sistema de computación basado en el paradigma de Hadoop que destaca por su velocidad gracias al almacenamiento de datos en memoria.
- Clúster de Spark.
- Desarrollo de aplicaciones con Spark.
- RDD, Datasets y transformaciones.
- Spark Streaming.

#### Tecnologías Big Data en Google Cloud:

- Ingesta y procesamiento en batch: Cloud Dataproc.
- Ingesta en streaming: Cloud PubSub.
- Almacenamiento de ficheros: Google Cloud Storage.
- Almacenamiento de datos: BigQuery.
- Desarrollo y ejecución de pipelines: Cloud Data Fusion.
- Visualización de datos: Cloud Data Studio.
- Analítica avanzada de datos: Vertex AI.
- Orquestación de datos: Cloud Composer.
- Catálogo y gobierno del dato: Cloud Dataplex.



# MÓDULO

## Productivizar un modelo

Productivizar un modelo es esencial para aplicar de manera adecuada los conocimientos adquiridos durante los anteriores módulos del máster, así como los modelos de machine learning de manera eficiente y efectiva en el mundo real.

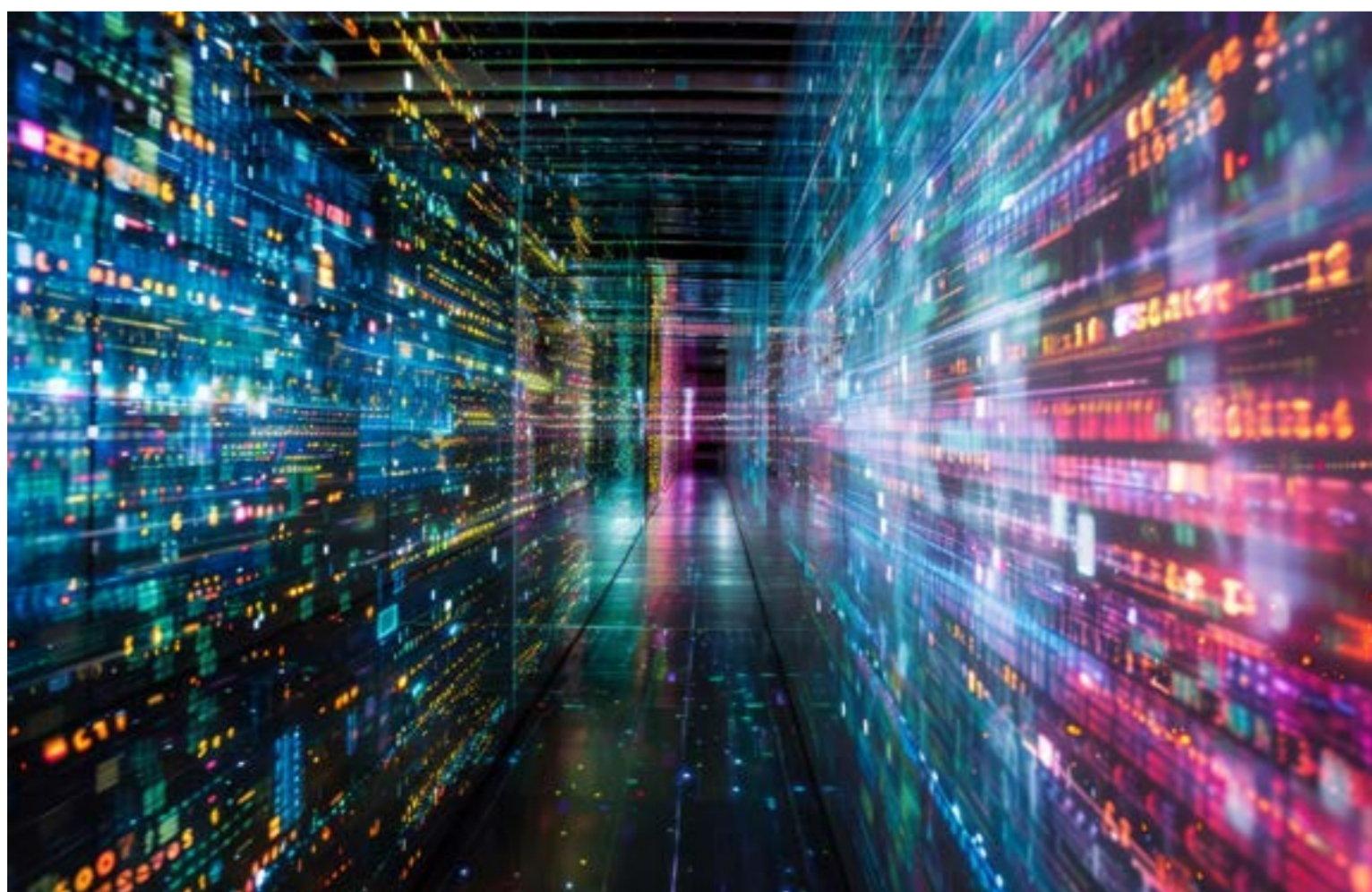
Una vez que se ha desarrollado un modelo, el siguiente paso es **implementarlo en un sistema de producción para que pueda ser utilizado para hacer predicciones o clasificaciones en tiempo real**. El módulo de productivizar un modelo es la herramienta que ayuda a automatizar este proceso de implementación.

Este módulo es responsable de **integrar el modelo en el flujo de trabajo del sistema de producción y asegurar que el modelo pueda funcionar** correctamente en un entorno en vivo. Algunas de las tareas que el módulo de productivizar un modelo realiza son:

- **Preprocesamiento de datos:** el módulo prepara los datos de entrada para el modelo, lo que puede incluir el procesamiento de datos faltantes, la normalización de datos, la selección de características y el muestreo de datos.
- **Integración del modelo:** el módulo integra el modelo de machine learning en el sistema de producción y establece las interfaces necesarias para que el modelo pueda ser utilizado por otros componentes del sistema.

- **Monitoreo del modelo:** el módulo de productivizar un modelo monitorea constantemente el rendimiento del modelo y alerta al equipo de cualquier problema o desviación en los resultados de las predicciones.
- **Escalado y actualización:** el módulo de productivizar un modelo también es responsable de garantizar que el modelo pueda manejar grandes volúmenes de datos y de actualizar el modelo a medida que se reciben nuevos datos para mejorar su precisión.

El módulo de productivizar un modelo es una herramienta crítica para llevar los modelos de machine learning a la producción y hacerlos útiles en el mundo real. Este módulo ayuda a **automatizar el proceso de integración del modelo** y garantiza que funcione de manera efectiva en un entorno en vivo.





# MÓDULO

## Data Science aplicada a la Empresa

Visión integral de las **empresas orientadas al dato**, creación de equipos de científicos de datos y estructuración de un proyecto Data Science. Este módulo aborda igualmente aspectos de la **comunicación** personal con individuos, la pública ante grupos, y la mediática para audiencias.

El objetivo principal es proporcionar a los alumnos una visión integral de las empresas orientadas al dato, formarles en la creación de equipos de Data Science y enseñarles a **estructurar un proyecto de ciencia de datos**.

Es importante que los estudiantes conozcan el amplio campo de acción de esta nueva ciencia y sus ventajas y aplicaciones prácticas en una estrategia digital a nivel global dentro de entornos empresariales. Los procesos, métodos científicos y sistemas estarán muy presentes a lo largo de este bloque del programa del máster.

El perfil del **Data Scientist** es crucial en todo el proceso que envuelve al macrodato. Es habitual que sea el científico de datos el profesional encargado de sacar el máximo provecho a los datos de los que dispone la empresa.

Contar con una excelente **capacidad de análisis** es muy útil en el perfil que nos ocupa, pero también es interesante que disponga de los conocimientos necesarios para, a través de las herramientas existentes a tal efecto, llegar a esas conclusiones analíticas.

El Data Science permite ir un paso más allá en el campo del análisis de datos. Tras lo aprendido con los fundamentos estadísticos, el data mining, el aprendizaje automático o los métodos de análisis predictivo, es lo que ha otorgado mayor protagonismo a este campo de conocimiento en los últimos años: ha sido la evolución natural derivada de la relación aplicada del dato al objetivo corporativo de la empresa.



# TFM

## Trabajo Fin de Máster

Asimilados todos los conceptos previos, llega el momento de poner a prueba los conocimientos adquiridos en el máster.

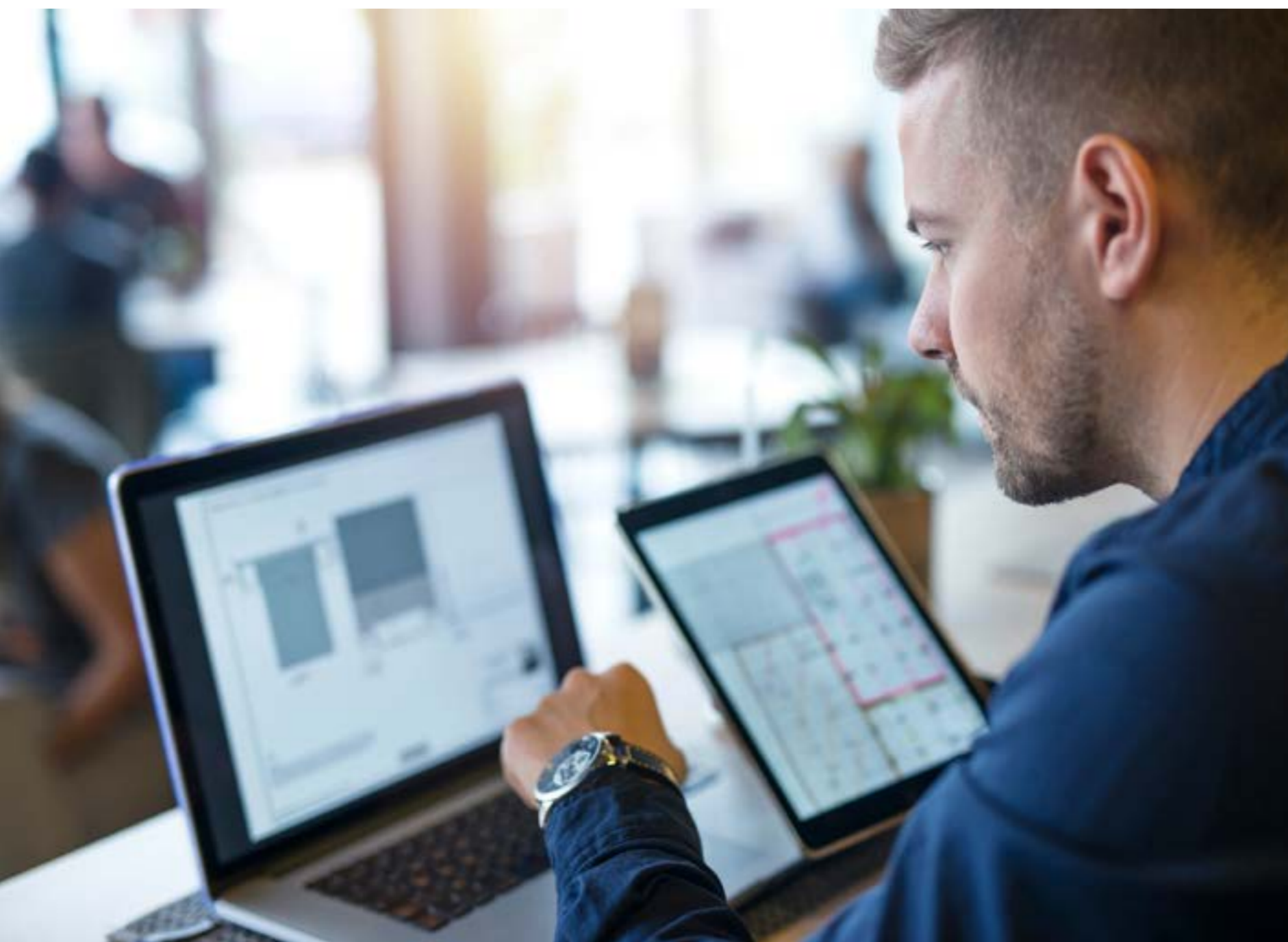
El alumno planteará una estrategia global de inteligencia de datos para una empresa, basándose en diferentes técnicas y softwares de apoyo de entre los existentes en el mercado.

El trabajo de fin de máster es una **parte crucial del programa**, ya que permite a los estudiantes aplicar todos los conocimientos adquiridos en el curso en un proyecto práctico y relevante en el mundo real.

Proporciona una oportunidad para que los estudiantes demuestren su capacidad para **analizar, procesar y utilizar datos de manera efectiva para resolver problemas complejos y tomar decisiones informadas**.

También les permite desarrollar habilidades de presentación y comunicación al presentar sus hallazgos y resultados a una audiencia de expertos.

Además, el trabajo de fin de máster puede ser una oportunidad para que los estudiantes trabajen en **colaboración con empresas u organizaciones**, lo que les permite obtener experiencia práctica en un entorno profesional y crear conexiones valiosas para su carrera.





# EQUIPO DOCENTE

*Clases y tutorías con grandes  
profesionales del sector del Big  
Data.*



UNIVERSIDAD  
COMPLUTENSE  
MADRID

 nticmaster



# EQUIPO DOCENTE

## Directivos



Javier  
Portela García-Miguel

Director. Profesor Titular UCM, doctor en Ciencias Matemáticas UCM, licenciado en Ciencias Matemáticas UCM. En el campo de la docencia, dirige diversos estudios en Data Mining y Business Intelligence en la UCM.



Jose Carlos  
Soto Gómez

Co-Director del Máster. Socio Fundador de NTIC Master y Aplimovil. Amplia experiencia en proyectos nacionales e internacionales en IT y analítica en empresas como Banco de España, NEC, Telefónica, Vodafone, Orange, medios de comunicación...

## Coordinadores



David  
del Ser

### Coordinador Máster

David es Lic. en Marketing por ESIC, Honours Degree in Business Administration por Humberstone University, MBA por UNED, Máster Dirección Financiera, Máster Marketing Digital, Máster en Big Data. Especialista en el desarrollo de negocio y transformación digital en Ntic Master. Gran experiencia profesional trabajando en Grupo Iberostar, Grupo Avintia, entre otras.



Cristóbal  
Martínez Martínez

### Coordinador Máster

Cristóbal es Ingeniero informático. Director de IT en Aplimovil y Ntic Master. Profesor máster marketing digital de la UCM, UNED, Cámara de Comercio y CEEIC. Experto en sistemas y procesos informáticos. Gran experiencia profesional trabajando en empresas referentes como NEC, BNP Paribas, Banco de España, Vodafone.



# EQUIPO DOCENTE

## Docentes



Cristóbal  
Pareja Flores

Catedrático EU en la UCM

Con más de 30 años como docente, Cristóbal es matemático especializado en Ciencias de la Computación, Doctor en Informática. Además, es Decano de la Facultad de Estudios Estadísticos y Vicedecano de Postgrado e Investigación.



Carlos  
Ortega Fernández

Expert Data Scientist

Carlos es Lic. en Físicas y trabaja como Lead Specialist en SDG Group España. Tiene una dilatada experiencia en el sector con aplicaciones avanzadas basadas en datos y Machine Learning en múltiples sectores.



Javier  
Monjas

Analytical Lead en Aggity

Javier es un referente en el sector Big Data, especialista en Machine Learning e Inteligencia artificial y modelos predictivos. Graduado en estadística y profesor asociado de la UCM. Ha trabajado en dirección de proyectos analíticos en Sanitas, Finect.



Conrado Miguel  
Manuel García

Catedrático EU en la UCM

Conrado es Lic. en matemáticas, Doctor y Director de Departamento de estadística y ciencia de datos en la Facultad de Estudios Estadísticos. Integrante de grupos de investigación sobre teoría de juegos y aplicaciones.



Daniel  
Gómez González

Profesor titular en la UCM

Daniel es Lic. en Matemáticas, Estadística e Investigación operativa y Doctor por la UCM. Profesor catedrático en UCM, Coordinador del programa de doctorado en Data Science, Experto en SPSS, estadística y ciencias de la computación.



Charles  
Flores Espinoza

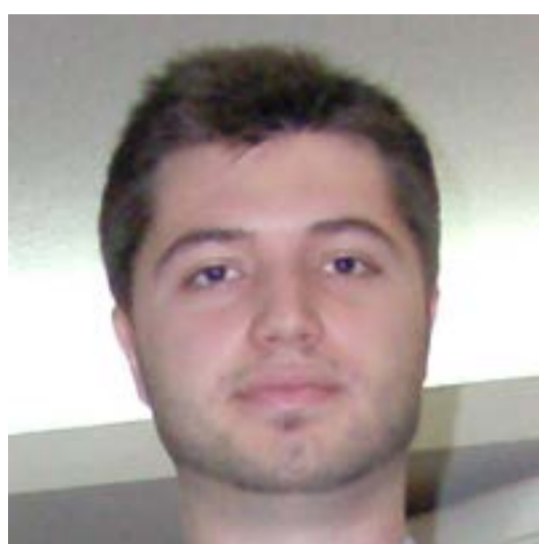
Big Data Engineer en  
Mercedes-Benz AG

Charles es Ingeniero informático. Experto Data Engineer & Data Architect, Scala, ML. Gran experiencia en el sector en empresas como StratioBD, Oliver Wyman, Vass, Bayes.



# EQUIPO DOCENTE

## Docentes



Pablo J.  
Villacorta

Científico de Datos en  
Santalucía Seguros

Pablo es Doctor en Ciencias de la Computación e IA, Ingeniero informático y licenciado en Estadística por la Univ. de Granada. Desarrollador Certificado en Spark 2.x por Databricks y autor de varios paquetes de R publicados en CRAN.



Santiago  
Mota Herce

Corporate Advisor

Santiago es senior data scientist y consultor freelance en Business Intelligence, Machine Learning y estrategia TI. Ha asesorado a empresas como Repsol, Orange, Bankia, Vodafone, Teradata o The Boston Consulting Group.



Isabel  
Riomoros Callejo

Profesora Titular en la UCM

Con más de 30 años como docente. Isabel es Lic. en Matemáticas, Doctora por la UCM. Además es Coordinadora del Campus Virtual (Facultad de Estudios Estadísticos) y Delegada del Decano para temas informáticos y de nuevas tecnologías.



Álvaro  
Bravo Acosta

Ingeniero Técnico Informático  
en Sistemas

Experto en Tecnologías Big Data, BI y Analítica. Gran experiencia en consultoras como Minsait, Sopra Steria, Everis o NTT, para externos como ISBAN y BBVA. Actualmente, DevOps en equipo de producto Frameworks en Strato.



Lorenzo  
Escot Mangas

Profesor Titular de la UCM

Además de su experiencia de más de 20 años como docente en la Complutense, Lorenzo es Codirector del Grupo de Investigación «Análisis de Datos en Estudios de Sociales, de Género y Políticas de Igualdad».



Cristóbal  
Martínez Martínez

IT Consultor

Cristóbal cuenta con amplia experiencia como profesor en diversos másteres de la UCM, la UNED y la Complutense de Cartagena. También ha sido consultor para proyectos de NEC, Vodafone y Movistar.



# EQUIPO DOCENTE

## Docentes



Alberto  
Ezpondaburu  
**NLP Specialist**

Alberto es ingeniero de telecomunicaciones, Lic. en Matemáticas. Experto en NLP y en la aplicación de técnicas de Deep Learning e Inteligencia Artificial. Ha trabajado como Data Scientist en varias empresas y creado empresas del ámbito AI.



Luis  
Leite  
**Business Intelligence Consultant**

Experto en la transformación de los datos con experiencia en diferentes áreas como las telecomunicaciones o industria farmacéutica. Posee un master en Física aplicada y otro en Big Data.



Carlos  
Rodríguez  
**Lead NLP Engineer**

Carlos es ingeniero de telecomunicación y máster en IA por la UPM. Es Lead NLP Engineer en Fujitsu. Anteriormente trabajó en empresas como Telefónica, EY, o Vodafone.



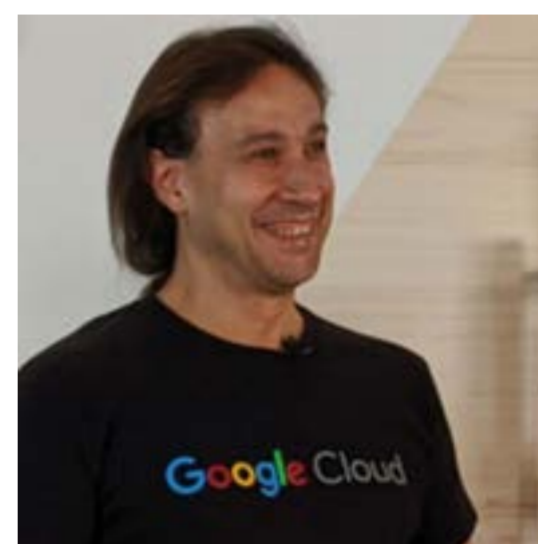
Luis  
Gascó Sánchez  
**NLP Research Engineer**

Senior Machine Learning Engineer en Avature. Doctor en Ingeniería por la Universidad Politécnica de Madrid con una larga trayectoria profesional en proyectos de I+D+i relacionados con IA y NLP. Ha trabajado en centros de investigación de prestigio como Barcelona Supercomputing Center, Nokia Bell Labs (UK) o Telecom ParisTech (Francia), entre otros.



Eduardo  
Fernández Carrión  
**Data Scientist, PhD**

Eduardo es ingeniero informático y doctor en métodos estadísticos matemáticos para el tratamiento computacional de la información. Data Scientist & ML Engineer en Santalucia seguros. Experto Data Scientist habiendo desarrollado su carrera profesional en StratioBD, Visavet, Everis, etc.



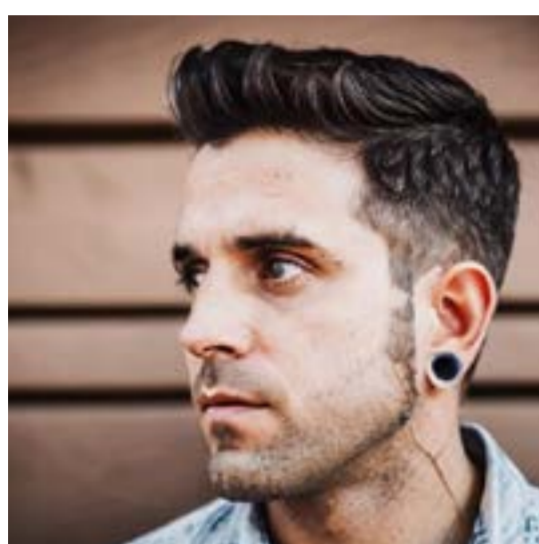
Pedro Pablo  
Malagón Amor  
**Principal Architect en Google**

Pedro Pablo es ingeniero de sistemas. Experto en Big Data, Data Science, Data Engineer. Tiene una gran experiencia profesional de más de 23 años en Microsoft como Principal Architect y actualmente en Google.



# EQUIPO DOCENTE

## Docentes



Juan Fernando  
Sánchez Martínez

**Data & Analytics Trainer en  
Telefónica**

Profesor de Data & Analytics experto en Bases de Datos SQL, Business Intelligence e Inteligencia Artificial. Ingeniero Industrial Químico (UPV), posee una sólida experiencia en investigación adquirida en la Universidad de Castilla-La Mancha y en la Universidad La Sapienza di Roma. Colabora con reconocidas empresas como Telefónica, BBVA y SDG Group.



Rosa  
Espinola

**Docente en la UCM**

Rosa es Doctora y Vicedecana de ordenación académica y estudiantes. Es miembro de proyectos de investigación del Ministerio de Educación y Ciencia para técnicas de Machine Learning y participante de grupos de investigación sobre teoría de juegos y aplicaciones.



Javier  
Domínguez Gómez

**Cybersecurity Engineer**

Javier es Ingeniero informático con certificaciones en CISSp, Cisco, entre otras. Experto en Cberseguridad y Data Science. Desarrolla su carrera como Cybersecurity Engineer | D&R para BBVA.



Javier  
Castro Cantalejo

**Docente en la UCM**

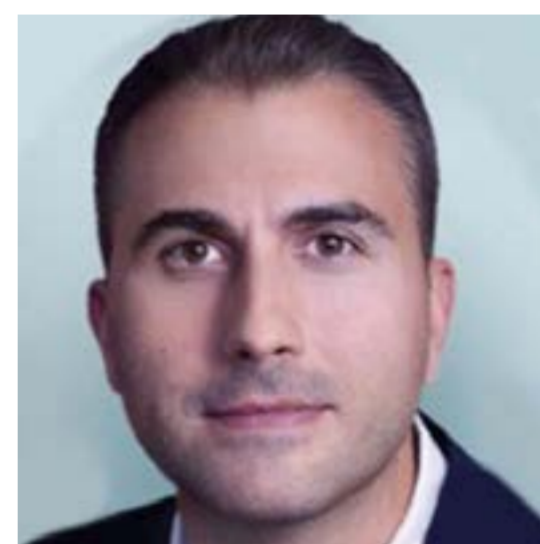
Javier es Doctor por la Universidad Complutense de Madrid con la tesis Reparto de holguras y costes en una red Pert 2007. Javier tiene experiencia, como miembro de grupos de investigación. Experto en «Text Mining y Redes Sociales».



Marlon  
Cárdenas Bonett

**Líder de Data Science en Sopra  
Steria**

Marlon es responsable de Data Science en Sopra Steria, liderando varias iniciativas de proyectos en el área de la analítica avanzada en sectores diversos. Además, es arquitecto de soluciones especializado en bases de datos.



Jorge  
Centeno

**Head of Data en Inditex  
Logistics SA**

Jorge es ingeniero de software con más de 15 años de experiencia. Ha desarrollado arquitecturas de datos en múltiples industrias como redes sociales, viajes, sector público y seguros, como arquitecto de soluciones y liderando equipos de Ingenieros de datos.



# EQUIPO DOCENTE

## Docentes



Pablo  
Hidalgo

**Databricks Champion ML  
Engineer en Capgemini**

Desarrollador de soluciones escalables para la conversión de texto a SQL mediante algoritmos de vanguardia, ingeniería rápida, LLM, interfaces de usuario y Databricks (catálogo y almacenes de Unity).



Tú

**Futuro experto en Big Data y  
Data Science**

Porque con nosotros aprenderás Big Data, Data Science e Inteligencia Artificial. Pero al final el camino tienes que recorrerlo tú y quizás muy pronto estés aquí como nuestro profesor.



# EQUIPO DOCENTE

## Ponentes Masterclass



José Ignacio  
Carrión

**Private Equity Mandate Officer  
at European Investment Fund**

José Ignacio se dedica al desarrollo, implementación y gestión de mandatos de inversión con clientes institucionales públicos. Supervisa y gestiona las inversiones en fondos de fondos privados a nivel europeo.



Ángel Araujo

**Southern Europe  
General Manager**

Responsable del negocio de Evaneos para España e Italia. Responsable de la P&L y de un equipo de 35 personas. Diseño y ejecución de las diferentes palancas de la plataforma: Oferta, Demanda y Marketing. . Managing Partner y Co-Founder de Unlock Management.



Matteo  
Consonni

**Former startup founder**

Matteo cuenta con experiencia en desarrollo de negocios, gestión de programas y financiación en 4 países. Actualmente, dirige las actividades de aceleración de Startups para Health Kic del Instituto Europeo de Innovación y Tecnología (EIT)



Guillermo  
Campoamor

**CEO de Meep**


Tras pasar gran parte de su vida laboral viajando por más de 22 países, decidió emprender en el sector de la movilidad. Desde entonces, Meep ha levantado más de 2 millones de euros y ha sido elegida como una de las 10 startups más relevantes en movilidad por #SouthSummit18



Armando  
Heras

**Chieff Data Officer**

Armando Heras, Director de la Estrategia digital de CMC XXI, especializado en finanzas, auditorías y financiación. Su función reside, entre otras cosas, en crear ecosistemas emprendedores a través de la inversión, creación de empresas y desarrollo de proyectos.



# SALIDAS PROFESIONALES

*El sector de Tecnologías de la Información continuará siendo el motor principal del crecimiento laboral, encabezando el top de empleos más demandados y mejor remunerados por las empresas.*



UNIVERSIDAD  
COMPLUTENSE  
MADRID

 nticmaster





# Salidas Profesionales

El Big Data ofrece una amplia variedad de salidas profesionales debido a su creciente relevancia en la estrategia empresarial moderna. Desde la analítica avanzada hasta la ingeniería de datos, estas son solo algunas de las muchas oportunidades profesionales que podrás explorar en el ámbito del Big Data:

*Con el avance de la tecnología y la evolución constante del entorno digital, es probable que sigan surgiendo nuevas especialidades y roles.*

## Gestor y auditor de infraestructuras para Big Data

Responsable de planificar, gestionar y auditar los sistemas que permiten almacenar y procesar grandes volúmenes de datos de manera eficiente.

## Ingeniero de aprendizaje automático

Desarrolla y optimiza algoritmos de machine learning para que los sistemas puedan aprender y mejorar automáticamente a partir de datos.

## Especialista en inteligencia de datos y de BI

Se enfoca en analizar datos y generar informes para ayudar en la toma de decisiones estratégicas basadas en la inteligencia empresarial.

## Chief Data Officer

Ejecutivo responsable de la estrategia de datos de una organización, asegurando su correcta gestión y uso para impulsar los objetivos de negocio.

## Data Analyst

Analiza conjuntos de datos grandes y complejos para extraer conclusiones útiles que apoyen la toma de decisiones en las empresas.

## Data Consultant

Proporciona asesoramiento experto en la gestión, análisis y uso estratégico de los datos para mejorar el rendimiento empresarial.

## Data Scientist

Utiliza estadísticas, machine learning y otras herramientas avanzadas para interpretar datos complejos y ofrecer soluciones a problemas empresariales.

## Data Mining

Especialista en técnicas de minería de datos que busca patrones, tendencias y relaciones ocultas en grandes bases de datos para extraer información útil.

## Analista de seguridad de datos

Responsable de proteger los datos frente a amenazas externas, asegurando la privacidad y la integridad de la información.

## Ingeniero de Procesamiento de Lenguaje Natural (NLP)

Se especializa en desarrollar sistemas que permitan a las máquinas entender y procesar el lenguaje humano, un campo clave en la IA.

## Especialista en análisis predictivo

Utiliza técnicas avanzadas para predecir comportamientos futuros y tendencias en diferentes sectores, como marketing, ventas o logística.



# ADMISIONES

*Tanto la preinscripción como la pre matrícula quedan abiertas hasta comenzar el curso académico o completar plazas.*



UNIVERSIDAD  
COMPLUTENSE  
MADRID

 nticmaster



# Proceso de admisión



## Preinscripción

Envío de solicitud para evaluar candidaturas.

1. Preinscribirse cumplimentando el formulario ubicado en la pestaña "Preinscripción" de la web [www.masterbigdataucm.com](http://www.masterbigdataucm.com)
2. Enviar la documentación requerida a fin de evaluar la candidatura.
3. Entrevista con el solicitante.
4. Confirmación de selección.
5. Realización de un pago inicial.



## Evaluación

Evaluación de candidaturas.



## Pre-admisión

Confirmación como alumno del candidato.



## Admisión

Confirmación de plaza y formalización de la matrícula.

Tanto la preinscripción como la pre matrícula quedan abiertas hasta comenzar el curso académico o completar plazas, estableciéndose lista de espera si procede. Los alumnos deberán ingresar 600 euros en concepto de pago inicial para el Máster Presencial y 500 euros en concepto de pago inicial para el Máster Semipresencial y el Máster Online, los cuales les serán descontados del importe total de la matrícula. En ningún caso se tendrá derecho a devolución de dicha cantidad, a excepción de que no se llegara a celebrar el curso.

# Documentación requerida

## Alumnos con titulación de **España**

Los documentos identificativos requeridos para la inscripción en el Máster son:

- Fotocopia del documento de identidad/pasaporte.
- Certificado de notas oficial.
- Título universitario o resguardo de solicitud de título.
- Currículum Vitae.

## Alumnos con titulación de **Unión Europea**

- Currículum Vitae.
- Pasaporte/NIE (no válidas las cédulas de identificación de fuera de España).
- Título universitario (no es valido el certificado del título).
- Certificado oficial de notas.

\*La documentación debe estar traducida al castellano por un traductor jurado homologado. (Solicitar listado oficial)

## Alumnos con titulación de **Fuera de la Unión Europea**

- Currículum Vitae.
- Pasaporte/NIE (no válidas las cédulas de identificación de fuera de España).
- Título universitario legalizado con la Apostilla de la Haya (no es valido el certificado del título).
- Certificado oficial de notas.

\*La documentación debe estar traducida al castellano por un traductor jurado homologado. (Solicitar listado oficial)



# MODALIDADES

*La evaluación de los alumnos se realizará a lo largo de todo el programa a través de ejercicios y casos prácticos.*



UNIVERSIDAD  
COMPLUTENSE  
MADRID

 nticmaster



# Máster Presencial



## Duración

**1 Año / 520 horas**  
**60 ECTS**

Inicio: septiembre de 2025  
Fin: octubre de 2026

Viernes: de 16:00 a 21:00 h  
Sábados: de 09:00 a 14:00 h



## Lugar

**Facultad de Estudios Estadísticos**

Ciudad Universitaria Avenida  
Puerta de Hierro, s/n, 28040  
Madrid



## Precio

**6.700 €**  
**+ 40 € de tasas de secretaría**

Pregunta por nuestras becas, facilidades de pago, prácticas en empresas y bolsa de trabajo.

*Una vez finalizados y superados estos estudios, la Universidad Complutense de Madrid emitirá el título, conforme a las normas de admisión y matriculación de los títulos de Formación Permanente de la UCM*

## Metodología Presencial

El curso se impartirá en aulas de la Universidad Complutense de Madrid, en la Facultad de Estudios Estadísticos los viernes y sábados con masterclass impartidas por diferentes expertos. La formación se realizará de forma tutorizada por los profesores. También se utilizará una plataforma de formación virtual para la comunicación entre los alumnos y profesores, creando una comunidad virtual de trabajo. Los distintos profesores de cada módulo, guiarán a los alumnos proponiendo actividades adicionales dependiendo del temario que se esté cubriendo en cada momento.

## Características plataforma On-line

La plataforma actuará como vía de comunicación entre el alumno y el entorno global de formación.

El estudiante tendrá información actualizada sobre los conceptos que se estén estudiando en cada momento, como enlaces a contenidos adicionales incluyendo noticias, artículos, etc.

Los alumnos deberán realizar y aprobar todas las prácticas de los distintos módulos, y realizar el trabajo fin de máster para poder aprobar el máster.

### La plataforma cuenta con:

- Mensajería individualizada para cada alumno.
- Vídeos de las clases y de casos prácticos.
- Tutorías online con el profesorado.
- Documentación, noticas y contenidos.
- Foro de los módulos del máster.
- Comunicación con los profesores vía mensajería.
- Chat entre alumnos.



# Máster Semipresencial



## Duración

**1 Año / 520 horas  
60 ECTS**

Inicio: octubre de 2025  
Fin: octubre de 2026

Jueves y viernes: de 16:00 a 21:00 h  
Sábados: de 09:00 a 14:00 h



## Lugar

**Online con  
presencialidad de 3  
semanas**

**Facultad de Estudios  
Estadísticos**

Ciudad Universitaria Avenida  
Puerta de Hierro, s/n, 28040  
Madrid



## Precio

**5.600 €  
+ 40 € de tasas de  
secretaría**

Pregunta por nuestras becas,  
facilidades de pago, prácticas en  
empresas y bolsa de trabajo.

*Una vez finalizados y superados estos estudios,  
la Universidad Complutense de Madrid emitirá  
el título, conforme a las normas de admisión  
y matriculación de los títulos de Formación  
Permanente de la UCM*

## Metodología Semipresencial

La formación se realizará de forma tutorizada por los profesores. Se utilizará una plataforma de formación virtual para la comunicación entre los alumnos y profesores, creando una comunidad virtual de trabajo. Los distintos profesores de cada módulo, guiarán a los alumnos proponiendo actividades adicionales dependiendo del temario que se esté cubriendo en cada momento. La modalidad semipresencial contempla la realización de 3 semanas presenciales con master class impartidas por diferentes expertos para preparar los TFM y hacer networking.

## Características plataforma On-line

La plataforma actuará como vía de comunicación entre el alumno y el entorno global de formación.

El estudiante tendrá información actualizada sobre los conceptos que se estén estudiando en cada momento, como enlaces a contenidos adicionales incluyendo noticias, artículos, etc.

Los alumnos deberán realizar y aprobar todas las prácticas de los distintos módulos, y realizar el trabajo fin de máster para poder aprobar el máster.

### La plataforma cuenta con:

- Mensajería individualizada para cada alumno.
- Vídeos de las clases y de casos prácticos.
- Tutorías online con el profesorado.
- Documentación, noticas y contenidos.
- Foro de los módulos del máster.
- Comunicación con los profesores vía mensajería.
- Chat entre alumnos.



# Máster Online



## Duración

**1 Año / 520 horas**  
**60 ECTS**

Inicio: octubre de 2025  
Fin: octubre de 2026



## Lugar

**Plataforma Online**



## Precio

**4.650€**  
**+ 40€ de tasas de**  
**secretaría**

Pregunta por nuestras becas,  
facilidades de pago, prácticas en  
empresas y bolsa de trabajo.

*Una vez finalizados y superados estos estudios,  
la Universidad Complutense de Madrid emitirá  
el título, conforme a las normas de admisión  
y matriculación de los títulos de Formación  
Permanente de la UCM*

## Metodología 100% Online

La formación se realizará de forma tutorizada por los profesores. Se utilizará una plataforma de formación virtual para la comunicación entre los alumnos y profesores, creando una comunidad virtual de trabajo. Los distintos profesores de cada módulo, guiarán a los alumnos proponiendo actividades adicionales dependiendo del temario que se esté cubriendo en cada momento.

## Características plataforma On-line

La plataforma actuará como vía de comunicación entre el alumno y el entorno global de formación.

El estudiante tendrá información actualizada sobre los conceptos que se estén estudiando en cada momento, como enlaces a contenidos adicionales incluyendo noticias, artículos, etc.

Los alumnos deberán realizar y aprobar todas las prácticas de los distintos módulos, y realizar el trabajo fin de máster para poder aprobar el máster.

### La plataforma cuenta con:

- Mensajería individualizada para cada alumno.
- Vídeos de las clases y de casos prácticos.
- Tutorías online con el profesorado.
- Documentación, noticas y contenidos.
- Foro de los módulos del máster.
- Comunicación con los profesores vía mensajería.
- Chat entre alumnos.



# Contacto

## Teléfono de información

+34 687 30 04 04

## Teléfono de admisiones

+34 667 89 05 83

## Correo electrónico

info@masterbigdataucm.com

## Sitio Web

www.masterbigdataucm.com

\*La dirección del máster se reserva el derecho de modificar, suprimir y actualizar los profesores, la información y el programa del máster.







UNIVERSIDAD  
COMPLUTENSE  
MADRID

 nticmaster