Catálogo de Especialidades Formativas

PROGRAMA FORMATIVO

ANALISTA DE DATOS BIG DATA CLOUDERA

Versión 3

Febrero 2026

GOBIERNO DE ESPAÑA
MINISTERIO DE TRABAJO Y ECONOMÍA SOCIAL
SERVICIO PÚBLICO DE EMPLEO ESTATAL
SEPE

SISTEMA NACIONAL DE EMPLEO

IDENTIFICACIÓN DE LA ESPECIALIDAD Y PARÁMETROS DEL CONTEXTO FORMATIVO

Denominación de la especialidad:	ANALISTA DE DATOS BIG DATA CLOUDERA
Familia Profesional:	INFORMÁTICA Y COMUNICACIONES
Área Profesional:	SISTEMAS Y TELEMÁTICA
Código:	IFCT35
Nivel de cualificación profesional:	3

Objetivo general

Extraer, manejar, almacenar, buscar y visualizar grandes volúmenes de datos de diferentes tipos relacionándolos entre sí para obtener información relevante mediante herramientas y soluciones Cloudera.

Relación de módulos de formación

Módulo 1	Introducción a SQL	50 horas
Módulo 2	Fundamentos de Cloudera Apache Hadoop	25 horas
Módulo 3	DANA-262: Analyzing with Cloudera Data Warehouse	75 horas

Modalidades de impartición

Presencial

Duración de la formación

Duración total	150 horas

Requisitos de acceso del alumnado

Acreditaciones / titulaciones	Cumplir como mínimo alguno de los siguientes requisitos: -Título de Bachiller o equivalente -Título de Técnico Superior (FP de Grado superior) o equivalente -Haber superado la prueba de acceso a ciclos formativos de Grado Superior -Haber superado cualquier prueba oficial de acceso a la universidad -Certificado profesional de nivel 3 -Título de Grado Universitario o equivalente -Título de Master Universitario o equivalente
Experiencia profesional	No se requiere
Otros	Se recomiendan los siguientes requisitos mínimos - Dominio de inglés a nivel de lectura. - Conocimientos básicos de sistemas operativos. - Conocimientos básicos en bases de datos relacionales.

Otros	- Conocimientos de básicos SQL. - Conocimientos sobre comandos de Linux. - Aunque no es obligatorio, se recomienda el manejo de algún lenguaje de scripting (Shell scripting, Python, Perl). Cuando el aspirante al curso no posea el nivel académico indicado, demostrará conocimientos suficientes a través de una prueba de acceso.

Justificación de los requisitos del alumnado

Para acreditar los requisitos de titulación, se documentará con la presentación de los títulos o certificados correspondientes.
-En caso de no disponer de nivel académico mínimo indicado, el aspirante demostrará conocimientos y competencias suficientes para participar en el curso con aprovechamiento mediante la superación de una prueba de acceso de competencias básicas correspondiente a la especialidad.
Además, el Centro comprobará mediante una prueba de acceso que el candidato posee conocimientos generales relacionados con la especialidad indicados anteriormente.

Prescripciones de formadores y tutores

Acreditación requerida	Cumplir como mínimo alguno de los siguientes requisitos: Cumplir como mínimo alguno de los siguientes requisitos: - Licenciado/a, Ingeniero/a, Arquitecto/a o el Título de Grado correspondiente u otras titulaciones equivalentes en los ámbitos de conocimiento de ingeniería, informática y de sistemas, ingeniería eléctrica, ingeniería electrónica, ingeniería de la telecomunicación, ingeniería industrial, matemáticas, estadística. - Diplomado/a, Ingeniero/a Técnico/a, Arquitecto/a Técnico/a o el Título de Grado correspondiente u otras titulaciones equivalentes en los ámbitos de conocimiento de ingeniería, informática y de sistemas, ingeniería eléctrica, ingeniería electrónica, ingeniería de la telecomunicación, ingeniería industrial, matemáticas, estadística. - Técnico superior de las familias profesionales: Informática y Comunicaciones, Electricidad y Electrónica.
Experiencia profesional mínima requerida	Al menos 1 año en ocupaciones relacionadas con la especialidad
Competencia docente	Se requiere acreditar un mínimo de 500 horas de experiencia docente o 250 horas si la experiencia está relacionada con especialidades o certificados de profesionalidad de la misma familia profesional en la que está encuadrada la presente especialidad. O estar en posesión de alguna de las siguientes titulaciones: -Certificado de Profesionalidad de Docencia de la Formación Profesional para el empleo (SSCE0110) o equivalente. -Certificado de Aptitud Pedagógica (CAP). -Título de licenciado en Pedagogía, Psicopedagogía o de Maestro en cualquiera de sus especialidades, graduado universitario en el ámbito de la Psicología o de la Pedagogía, o

Competencia docente	un título universitario oficial de posgrado en los citados ámbitos. -Título profesional de Especialización Didáctica, el Certificado de Cualificación Pedagógica, el Máster Universitario habilitante para el ejercicio de las Profesiones Reguladas de Profesor de Educación Secundaria Obligatoria y Bachillerato, Formación Profesional y Escuelas Oficiales de idiomas.
Otros	El formador deberá estar homologado como instructor en la correspondiente tecnología específica del fabricante. Excepcionalmente, cuando el profesor no acredite el requisito de titulación académica requerido, el centro formativo podrá valorar la posibilidad de acreditarlo mediante la emisión por parte del fabricante de una carta de validación como docente

Justificación de las prescripciones de formadores y tutores

Para acreditar tanto los requisitos de formación como de experiencia impartiendo formación en las horas mínimas estipuladas, se documentará con la presentación de títulos y certificados oficiales, vida laboral, certificados de empresa o contratos más liquidación (que demuestre que se han ejecutado).
Además, deberá disponer de la alguna de las siguientes certificaciones vigentes a la fecha de comienzo de la formación:
- Instructor de Cloudera.

Requisitos mínimos de espacios, instalaciones y equipamientos

Espacios formativos	Superficie m² para 15 participantes	Incremento Superficie/ participante (Máximo 30 participantes)
Aula de informática	45.0 m²	2.4 m² / participante

Espacio formativo	Equipamiento
Aula de informática	- Mesa y silla para el formador. - Mesas y sillas para el alumnado. - Material de aula. - Pizarra. - Impresora láser con conexión a red. - Pantalla y cañón de proyección. - PC instalado en red con posibilidad de impresión de documentos, cañón con proyección e Internet para el formador. La configuración mínima de los equipos que se recomienda es: CPU de 64 bits Intel Core i5 de novena generación o superior, o AMD similar; 1 disco duro SSD de 512GB o superior, máquina virtual en remoto con memoria RAM 104GB, 2 discos SSD y procesador Intel AMD (24 cores); Pantalla 21´o superior, resolución de pantalla:1600*1080 para interfaces de gráficas de usuario, Gigabit Ethernet, soporte USB3 , teclado multimedia USB, ratón sensor óptico USB de 2 botones y rueda de desplazamiento. - PCs instalados en red e Internet con

	posibilidad de impresión para los alumnos. La configuración mínima de los equipos que se recomienda para cada alumno es: CPU de 64 bits Intel Core i5 de novena generación o superior, o AMD similar; 1 disco duro SSD de 512GB o superior, máquina virtual en remoto con memoria RAM 104GB, 2 discos SSD y procesador Intel AMD (24 cores); Pantalla 21´o superior, resolución de pantalla: 1600*1080 para interfaces de gráficas de usuario, Gigabit Ethernet, soporte USB3 , teclado multimedia USB, ratón sensor óptico USB de 2 botones y rueda de desplazamiento. - Software específico para el aprendizaje de cada acción formativa: o Licencia sistema operativo o Licencia del software antivirus o Licencias del software y herramientas necesarias para la impartición del curso (versión actualizada) o Acceso a los sistemas oficiales de Cloudera configurados específicamente con los ejercicios prácticos del curso aportados por el fabricante.

La superficie de los espacios e instalaciones estarán en función de su tipología y del número de participantes. Tendrán como mínimo los metros cuadrados que se indican para 15 participantes y el equipamiento suficiente para los mismos.

En el caso de que aumente el número de participantes, hasta un máximo de 30, la superficie de las aulas se incrementará proporcionalmente (según se indica en la tabla en lo relativo a m²/participante) y el equipamiento estará en consonancia con dicho aumento.

No debe interpretarse que los diversos espacios formativos identificados deban diferenciarse necesariamente mediante cerramientos.

Las instalaciones y equipamientos deberán cumplir con la normativa industrial e higiénico-sanitaria correspondiente y responderán a medidas de accesibilidad y seguridad de los participantes.

En el caso de que la formación se dirija a personas con discapacidad se realizarán las adaptaciones y los ajustes razonables para asegurar su participación en condiciones de igualdad.

Aula virtual

Características
• La impartición de la formación mediante aula virtual se ha de estructurar y organizar de forma que se garantice en todo momento que exista conectividad sincronizada entre las personas formadoras y el alumnado participante así como bidireccionalidad en las comunicaciones. • Se deberá contar con un registro de conexiones generado por la aplicación del aula virtual en que se identifique, para cada acción formativa desarrollada a través de este medio, las personas participantes en el aula, así como sus fechas y tiempos de conexión.

Vinculaciones con capacitaciones profesionales

CDP Data Analyst CDP-4001

Ocupaciones y puestos de trabajo relacionados

- 27111019 ANALISTAS DE SISTEMAS, NIVEL MEDIO (JUNIOR)
- 27111028 ANALISTAS DE SISTEMAS, NIVEL SUPERIOR (SENIOR)
- 27111037 INGENIEROS INFORMÁTICOS
- 27121021 ANALISTAS DE APLICACIONES, NIVEL SUPERIOR (SENIOR)
- 27121049 INGENIEROS TÉCNICOS EN INFORMÁTICA DE GESTIÓN
- 27191013 AUDITORES-ASESORES INFORMÁTICOS
- 27191022 INGENIEROS TÉCNICOS EN INFORMÁTICA, EN GENERAL

Requisitos oficiales de las entidades o centros de formación

Estar inscrito en el Registro de entidades de formación (Servicios Públicos de Empleo).

Los centros impartidores de formación Oficial de Cloudera deben cumplir el requisito: Acreditación de encontrarse autorizado por Cloudera Authorized Training Center de Cloudera.
Los requisitos mínimos tanto Hardware como Software serán los que marque el fabricante como recomendados en cada momento para las versiones actualizadas.
A los alumnos se les proporcionará la documentación oficial necesaria para el seguimiento del curso.

MÓDULO DE FORMACIÓN 1: Introducción a SQL

OBJETIVO

Conocer los fundamentos y dominar las acciones operaciones necesarias que permitan interactuar con bases de datos relacionales utilizando el lenguaje SQL.

DURACIÓN:

50 horas

RESULTADOS DE APRENDIZAJE

Conocimientos / Capacidades cognitivas y prácticas

- Introducción a las BBDD:
  Conceptos básicos de bases de datos.
  Utilidad y relevancia del lenguaje SQL.
- Creación y estructura de una base de datos:
  Estructura básica de una base de datos.
  Creación de una base de datos usando SQL.
- Realización de Consultas básicas:
  Búsqueda y recuperación de datos básicos.
  Manejo de consultas simples.
- Mantenimiento de la Base de datos
  Copia de tablas y modificaciones de columnas.
  Índices y restricciones.
  Eliminación o modificación de filas de la tabla de datos.
  Creación de objetos de BBDD (tablas, índices, vistas…)
- Utilización de Funciones:
  Uso y tipos de funciones.
  Funciones predeterminadas.
- Exportación e importación de datos:
  Exportación de datos, consultas y utilidades.
  Importación de datos y archivos de datos
  Importación con sentencias y utilidades
- Utilización de Joins para la extracción de datos
  Definición y tipo de Joins
  Uso de joins para combinar datos de múltiples tablas
- Ejecución y diseño de Subconsultas:
  - Tipos de subconsultas
    Subconsultas generales y básicas.
    Subconsulta no correlacionada y correlacionada.
    Modificación de la tabla con subconsultas.
- Realización de operaciones avanzadas:
  Creación de Vistas.
  Gestión de transacciones.
- Utilización del diccionario de Metadatos:
  Uso de metadatos para obtener información sobre los objetos de la base de datos..

Habilidades de gestión, personales y sociales

- Actitud activa y concienciación de la importancia del lenguaje SQL.
- Interés por la utilización de bases de datos relacionales como almacenamiento y explotación de datos estructurados.

MÓDULO DE FORMACIÓN 2: Fundamentos de Cloudera Apache Hadoop

OBJETIVO

Aprender los conceptos básicos de Big Data, conocer de la arquitectura de Hadoop e Identificar los componentes de Big Data y su integración en el ecosistema de Hadoop.

DURACIÓN:

25 horas

RESULTADOS DE APRENDIZAJE

Conocimientos / Capacidades cognitivas y prácticas

- Identificación y asimilación de conceptos
  Introducción a Big Data
  La motivación por Apache Hadoop
  Conceptos básicos de Hadoop
- Visión del Ecosistema de Hadoop
  Soluciones del ecosistema de Hadoop
  Aplicaciones comunes y usos especiales de Hadoop.
- Gestión de Hadoop
  - Uso de herramientas para la gestión del almacenamiento
  - Ejecución de aplicaciones con procesamiento distribuido

  Ingesta de datos en Hadoop
  Explotación y análisis de datos en Hadoop.

Habilidades de gestión, personales y sociales

- Capacidad para obtener, procesar y asimilar nuevas habilidades y conocimientos durante todo el proceso de desarrollo.

MÓDULO DE FORMACIÓN 3: **DANA-262: Analyzing with Cloudera Data Warehouse**

OBJETIVO

Aplicar las habilidades tradicionales de análisis de datos e inteligencia empresarial a grandes volúmenes de datos conociendo las herramientas que se necesitan para acceder, manipular, transformar y analizar datos de diferentes tipos utilizando SQL y lenguajes de scripting.

DURACIÓN:

75 horas

RESULTADOS DE APRENDIZAJE

Conocimientos / Capacidades cognitivas y prácticas

- Fundamentos para el Análisis de Big Data
  Visión General del Análisis de Big Data
  Almacenamiento de Datos: HDFS
  Procesamiento Distribuido de Datos: YARN, MapReduce y Spark
  Procesamiento y Análisis de Datos: Hive e Impala
  Integración de Bases de Datos: Sqoop
  Otras Herramientas de Datos
- Introducción a Hive e Impala
  ¿Qué es Hive?
  ¿Qué es Impala?
  ¿Por qué utilizar Hive e Impala?
  Schema y almacenamiento de datos
  Comparación entre Hive y bases de datos tradicionales
  Casos de uso
- Consultas con Hive e Impala
  Tablas y bases de datos
  Sintaxis básica en consultas Hive e Impala
  Tipos de datos
  Empleo de Hue para ejecutar consultas
  Empleo de Beeline (la Shell de Hive)
  Empleo de la Shell de Impala
- Operadores comunes y funciones integradas
  Operadores
  Funciones escalares
  Funciones de agregado
- Administración de datos
  Almacenamiento de datos
  Creación de bases de datos y tablas
  Carga de datos
  Modificación de bases de datos y tablas
  Simplificación de consultas con vistas
  Almacenamiento de resultados de consultas
- Almacenamiento de datos y rendimiento
  Particionamiento de tablas
  Carga de datos en tablas particionadas
  Cuándo utilizar particionamiento
  Selección del formato de archivo
  Uso de los formatos de archivo Avro y Parquet
- Trabajando con múltiples Datasets

UNION y Joins
Manejo de valores NULL en Joins
Joins avanzados

- Funciones analíticas y Windowing
  Utilización de funciones analíticas comunes
  Otras funciones analíticas
  Sliding Windows
- Datos complejos
  Datos complejos con Hive
  Datos complejos con Impala
- Análisis de texto
  Empleo de expresiones regulares con Hive e Impala
  Procesamiento de texto con SerDes en Hive
  Análisis de Sentimiento y n-grams en Hive
- Optimización de Apache Hive
  Comprender cómo se ejecutan las consultas
  Optimización basada en Costes y Estadísticas
  Bucketing
  Optimizaciones de ficheros ORC
  Indexación de datos
  Hive en Spark
- Optimización de Apache Impala
  Cómo Impala ejecuta las consultas
  Mejorar el rendimiento de Impala
- Extendiendo Hive e Impala
  Customizar SerDes y formatos de archivo en Hive
  Transformación de datos con Scripts personalizados en Hive
  Funciones definidas por el usuario
  Consultas parametrizadas
- Selección de la Mejor Herramienta para cada Tarea
  Comparación entre MapReduce, Hive, Impala, y bases de datos relacionales
  ¿Cuál elegir?

Habilidades de gestión, personales y sociales

- Habilidades de gestión en la resolución de incidencias, conflictos y problemas en la creación de cuadros de mando y creación de informes.
- Demostración de una actitud responsable y ética a la hora de utilizar de implementar aplicaciones con Cloudera.

ORIENTACIONES METODOLÓGICAS

- La formación ha de ser eminentemente práctica.
- Los conceptos y contenidos a adquirir han de ir acompañados de ejemplos prácticos.
- El formador/a utilizará el método demostrativo que consiste en que 1º el formador/a muestra el uso de las funciones en la plataforma y 2º da tiempo a los alumnos para que ellos lo realicen después.
- Todas las unidades de aprendizaje tienen que ir acompañadas de ejercicios planteados por el profesorado, de los que después se mostrará la solución.
- La evaluación formativa o control de la comprensión durante la impartición es imprescindible para que los alumnos avancen eficazmente y el formador/a realice los ajustes necesarios, si fuera preciso.

EVALUACIÓN DEL APRENDIZAJE EN LA ACCIÓN FORMATIVA

La evaluación tendrá un carácter teórico-práctico y se realizará de forma sistemática y continua, durante el desarrollo de cada módulo y al final del curso.

Puede incluir una evaluación inicial de carácter diagnóstico para detectar el nivel de partida del alumnado.

La evaluación se llevará a cabo mediante los métodos e instrumentos más adecuados para comprobar los distintos resultados de aprendizaje, y que garanticen la fiabilidad y validez de la misma.

Cada instrumento de evaluación se acompañará de su correspondiente sistema de corrección y puntuación en el que se explicite, de forma clara e inequívoca, los criterios de medida para evaluar los resultados alcanzados por los participantes.

La puntuación final alcanzada se expresará en términos de Apto/ No Apto.

CERTIFICACIÓN DE FABRICANTE

La ejecución y financiación del programa formativo incluye la presentación de los alumnos que han realizado el curso con aprovechamiento al examen para obtener la certificación oficial del fabricante, que gestionará el centro y que en ningún caso supondrá coste alguno para el alumno.
En concreto, para esta acción formativa está incluido el siguiente examen de certificación oficial de Cloudera, o el que lo sustituya actualizado al momento de su impartición:
CDP Data Analyst CDP-4001