



PROGRAMA FORMATIVO

Ciencia de datos en un entorno 4.0

Diciembre 2019

DATOS GENERALES DE LA ESPECIALIDAD

1. Familia Profesional: INFORMÁTICA Y COMUNICACIONES

Área Profesional: SISTEMAS Y TELEMÁTICA

2. Denominación: CIENCIA DE DATOS EN UN ENTORNO 4.0

3. Código. IFCT190PO

4. Objetivo general:

Convertir los datos de una empresa en nuevos productos y/o servicio, como Científico de Datos, Business Intelligent o Business Analytics, escribiendo código propio para analizar ingentes cantidades de datos y desarrollando dashboards interactivos para presentar la información, para lo que será imprescindible que el alumnado posea conocimientos de informática y programación a nivel de Técnico Superior de FP como mínimo o bien dos años de experiencia en lenguajes de programación.

5. Número de participantes:

Máximo 25 participantes para cursos presenciales.

6. Duración:

Horas totales: 132

Distribución horas:

- Presencial: 132

7. Requisitos mínimos de espacios, instalaciones y equipamiento

7.1. Espacio formativo:

- Aula de gestión: 45 m² para 15 alumnos.

Cada espacio estará equipado con mobiliario docente adecuado al número de alumnos, así mismo constará de las instalaciones y equipos de trabajo suficientes para el desarrollo del curso

7.2. Equipamiento:

- Aula de gestión:
 - Mesa y silla para el formador
 - Mesas y sillas para el alumno
 - Material de aula
 - Pizarra
 - PC instalados en red con posibilidad de impresión de documentos, cañón con proyector e internet para el formador
 - PCs instalados en red e internet con posibilidad de impresión para los alumnos
 - Software específico para el aprendizaje de la acción formativa

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

8. CONTENIDOS FORMATIVOS

1. Análisis de Datos con Python (16h)

- Estructuras de control y sintaxis
- Estructuras de datos
- Almacenamiento y manipulación de datos con Python
- Librería Numpy
- Pandas
- Matplotlib
- Técnicas de programación
- Casos prácticos

2. Recogida de datos (8h)

- Datos internos
Gestión de datos en el entorno empresarial
Ejercicios prácticos SQL
- Datos externos
Análisis en plataformas sociales
Análisis web
Análisis en Youtube e Instagram

3. Estructuras de Bases de datos (12h)

- Bases de datos tradicionales
Modelo Entidad-Relación
Modelo Relacional
Integridad Referencial
Programación SQL
- Nuevos tipos de bases de datos
Bases de datos nosql
MongoDB
Grafos

4. Análisis Estadístico (24h)

- Introducción, Probabilidad
- Variable aleatoria
- Inferencia
- Lenguaje R
Modelos estadísticos, Intervalos de confianza y Test de hipótesis, Regresión lineal y no lineal, Modelos lineales generalizados
- Técnicas de muestreo

5. Machine Learning (24h)

- Aprendizaje supervisado vs no supervisado
- Reglas de clasificación, selección de variables y de modelos
- Algoritmos de clasificación. Redes neuronales
- Vecino más cercano
- NaiveBayes. K-means
- Métodos aplicados a la bioinformática
- Deep Learning

6. Técnicas de Big Data (24h)

- Fundamentos de Big Data
- Adquisición de Datos y Almacenamiento (Hadoop, MapReduce y YARN, Apache Spark y HIVE)
- Interfaces de Usuario (HUE, AMBARY, SQL-Developer,..)
- Opciones de despliegue de BigData (Introducción a los servicios Cloud, Amazon AWS, Cloudera, Oracle, Hortonworks,..)

7. Power BI (8h)

- Introducción a Power Bi desktop
- Conceptos y opciones al conectarse a los datos
- Análisis. Creación de vistas
- Campos calculados
- Formato de datos
- Creación de cuadros de mando interactivos

8. Gestión de proyectos y Visualización (16h)

- Análisis descriptivo
- Técnicas de visualización. Diseño basado en datos
- Python Matplotlib
- Gestión e implementación de proyectos de analítica avanzada con metodologías Ágiles