

Data Scientist (Cientifico de datos)

Que es/hace un Data Scientist

A grandes rasgos un **data scientist** toma los datos e informacion de diversas fuentes como fuentes abiertas o internas de la organizacion y los usa para la creacion de **machine learning** y **encontrar informacion valiosa** Esto con el fin de toma de desiciones basadas en datos

Incorporar datos a los productos **Como por ejemplo, en los software de ventas incluir los productos mas comprados o con mayor demanda**

Trabajo de un Cientifico de Datos

- Extraer/Obtener y la limpieza de los datos
- Diseñar e implementar **machine learning** para el analisis de los datos
- Monitorear los datos (Monitorizacion del machine learning)
- Automatizacion del proceso de recoleccion y extraccion/limpieza de datos
- Creacion de reportes de informacion (Tableros, Tablas, Graficas)
- Incorporar los datos extraidos a los productos

Diferencias entre Analista de Datos y Cientifico de Datos

Data Scientist

- Incorporacion de datos a los productos de software, trabajo de la mano con el equipo de ingenieria
- Programacion mas avanzada (Programacion orientada a objetos POO)
- Machine Learning y estadistica avanzada (Al ver y trazar trayectoria al futuro)

Data Analyst

- Se enfoca en responder preguntas del negocio
- Programacion fundamental (En casos no se usa ni programacion)
- Analizar el presente de la organizacion

Herramientas y Tecnologias Para Data Scientist

- **Lenguajes de Programacion**
- Python y R (Incluir la programacion Orienta a Objetos)
- Jupyter NoteBooks (Con este se crea el codigo para posteriormente verlo representado en graficas)
- Librerias para la manipulacion y analisis de datos (Pandas, Numpy)
- Modelos de Machine Learning y llevarlos a la practica (Scikit-Learn y TensorFlow)
- Bases de datos NoSQL y SQL (Consultas de datos)

Matematicas en Data Scientist implementadas a la practica

- Algebra (Representacion de Planos y funciones)
- Probabilidad y Estadistica Descriptiva y Diferencial
- Algebra Lineal (Funcionamiento de Machine Learning)
- Calculo (Formular Funciones para los algoritmos de Machine Learning)

Comenzar y dar Inicio

- Saber como las empresas y organizaciones usan los datos
- Programacion en python y sus librerias para data science
- Creacion de Proyectos con **Jupyter NoteBooks**
- Estadisticas y Probabilidad (aplicado a Data Scientist)

