

La Ciencia de Datos (Data Science) es una de las profesiones más prometedoras de hoy en día.

¿Qué es un científico de datos?



Yo lo definiría como la intersección de tres campos fundamentales: Matemáticas, Estadística y Programación. Un científico de datos tiene que ser capaz de **analizar e interpretar grandes bases de datos, y extraer la información fundamental de esos datos para resolver los posibles problemas que le pueda interesar a la empresa.** Por lo cual, el científico de datos es uno de los profesionales más importantes en cualquier empresa hoy en día. De ahí la cantidad de ofertas de trabajo y las remuneraciones tan altas.

¿Hay trabajo?



La respuesta es sí, ahora y en el futuro. La tecnología actual no solo necesita del mejor talento sino de **datos y de resultados basados en ellos**. Hoy en día una empresa que no haga estudios con datos es casi imposible de visualizar. Todas necesitan data scientists. Por no hablar de la importancia que tienen los datos no solo en los negocios sino también en Medicina, detección de cáncer, neurociencias, bioquímica, en el área de investigación tanto académica como no, y en muchas otras ciencias.

El problema es que, de repente, ha surgido una gran demanda de un perfil que hasta hace poco prácticamente no existía. Recordemos que se precisan conocimientos estadísticos que un programador no suele tener y conocimientos informáticos que un estadístico no suele ni siquiera imaginar. Hoy en día cada vez van surgiendo más grados o máster de formación más enfocados a esa intersección de áreas, pero aún la demanda supera a la oferta

formativa, sobre todo por el precio. Lo bueno es que todo se puede solucionar con **formación autodidacta**.

¿Cómo formarme para ser un Data Scientist?



Hoy por hoy, podemos encontrar una gran diversidad de perfiles profesionales en el mundo de la ciencia de datos, que vienen básicamente de las matemáticas y la estadística, o de la ingeniería informática, pero cada vez es más usual encontrar perfiles de otras ingenierías.

Grados

Hoy en día, existen **algunos grados dobles** en ingeniería informática y matemáticas que parecen la mejor opción si nos planteamos esta especialización. También están los 'grados en ciencia de datos' que irán surgiendo poco a poco en el futuro.

Posgrados

El de los posgrados es **un mundo muy diverso**. Podemos encontrar posgrados, másteres o cursos de especialización en

casi todas las universidades y una oferta privada realmente desmesurada. La clave está en buscar complementar nuestra formación previa y, en ese sentido, la diversidad es una buena noticia.

Lo que sí podemos encontrar en la formación de posgrado que no podemos encontrar en la formación previa es el componente de '**orientación de negocio**'. No debemos olvidar que la mayor parte del trabajo de los científicos de datos está en empresas que buscan rentabilizar sus bases de datos, porque lo que la orientación al mercado es algo muy recomendable. La pega principal es el precio de estos estudios, ya que como están muy demandados, el precio suele ser bastante alto.

CURSOS ONLINE

Uno de los recursos más interesantes que puedes encontrar son los cursos online. Hoy en día son muchos los que se aprovechan de esta **ventajosa forma de aprendizaje**. Una de las plataformas más usadas, con cursos de precios accesibles, pero de gran calidad es **Udemy**. Te recomiendo que si optas por este tipo de formación, primero identifiques lo que tienes aún por mejorar y aprender, y luego empieces a buscar la mejor opción para formarte. Los cursos online de Udemy son una buena opción porque tienen un precio muy reducido, por un contenido de mucha calidad, profesores muy experimentados que estarán encantados de contestar cualquier duda que tengas. Lo mejor es que puedes hacerlos a tu ritmo, descargar los materiales, verlos cuando quieras porque son de acceso ilimitado de por vida y si no estás muy seguro te dan hasta 30 días para el reembolso del precio del curso por si acaso no te convence algo. Es una plataforma en crecimiento, cada vez hay más oferta y más estudiantes apuntados. Por último, por si no lo sabías, al finalizar el curso te dan un certificado de finalización oficial.

¿Qué base teórica tengo que tener?



La base teórica que debe tener un buen científico de datos es una base muy potente en matemáticas y estadística. De la primera rama te aconsejaría: álgebra y cálculo. De la segunda, que creo que es más importante aún dominarla de verdad, te aconsejo una base fuerte en: probabilidades, análisis descriptivos, inferencia, métodos de estimación y modelización (regresión lineal, clasificación, clúster, detección de atípicos, etc).

Es muy importante que la base teórica sea fuerte. Que domines todos los conceptos y no sea de memoria, sino que sepas interpretarlos, que sepas su importancia y cómo y por qué se utilizan. Es como un edificio en construcción, seguramente se termine destruyendo si los cimientos no están afianzados.

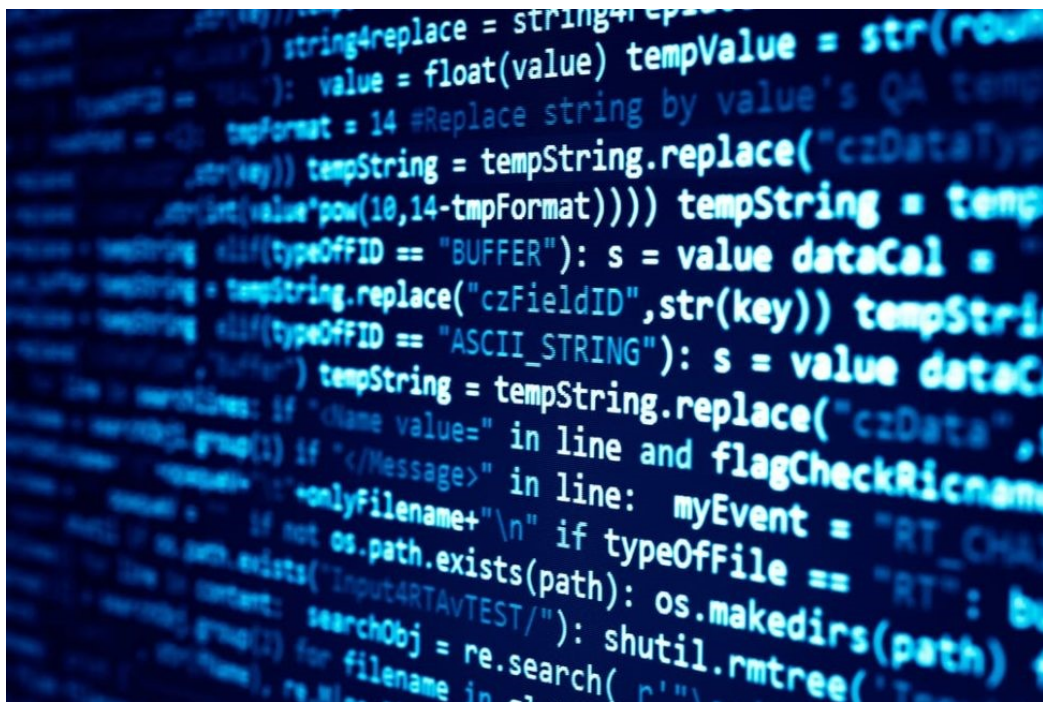
¿Quieres afianzar esa base Estadística? Tengo un descuento para ti, de mi curso online en Udemy “Curso completo de Estadística a nivel universitario” en este enlace:

<https://www.udemy.com/course/curso-completo-de-estadistica-a-nivel-universitario/?referralCode=5A1CC426596B6A1496B3>

Y si quieres aprender Data Science más avanzada y práctica, te dejo el descuento del nuevo curso que he lanzado "Análisis robusto de datos en R y Matlab":

<https://www.udemy.com/course/analisis-robusto-de-datos-en-r-y-matlab/?referralCode=D6897428B53B045636D9>

¿Qué lenguajes hay que aprender?



Depende, hay algunos más populares que otros, pero nunca se sabe cuál van a pedirte que sepas de cara a conseguir trabajo en una empresa. Los más usuales son los siguientes.

- **SQL:** Es **algo necesario** no sólo por la inmensa cantidad de datos de los que hablamos sino porque la mayor parte de los datos que usa un científico de datos profesional provienen de internet. Sin embargo, no es imprescindible,

porque como es fácil de aprender, pueden contratarte sin saber SQL y ya aprenderás en el día a día.

La gran división: los tres mosqueteros

- **R:** Tiene a su favor que ha sido el lenguaje estadístico por excelencia durante muchos años y podemos encontrar códigos y paquetes para casi cualquier cosa que se nos ocurra. Es el lenguaje de los que se acercan desde un **background científico**.
- **Python:** Es la némesis de R en este caso: tiene una sintaxis muy buena y moderna pero aún queda mucho trabajo por hacer desarrollando su ecosistema. No obstante, para ser justos, Python cada vez es más competitivo e iniciativas como SciPy están poniéndole las cosas muy difíciles a R. Es el lenguaje de los que se acercan desde un **background informático**. Además, tiene la ventaja de que con Google Colab puedes trabajar con código de Python conectado a un ordenador de Google que te ahorra tiempo de ejecución porque son mucho más potentes, y totalmente gratis.
- **Matlab:** Algunos lenguajes o entornos gozan de cierto éxito empujados por la inercia corporativa: es el caso del clásico **Matlab** pero progresivamente va perdiendo peso.

El mercado laboral: sueldos y oportunidades



Los sueldos, como en general en el mundo del desarrollo de software, cambian mucho dependiendo del **lugar, las funciones y el empleador**. No obstante, ahora mismo está muy bien pagado. A nivel general los sueldos/ingresos están en una media de 141.000 dólares para freelance, 107.000 para asalariados, 90.000 para trabajadores gubernamentales o en el sector sin ánimo de lucro; 70.000 dólares para trabajo en universidades.

No obstante, estos sueldos medios hay que tomarlas con mucha prudencia. Mientras el salario medio en Estados Unidos está entre 103.000 y 131.000 dólares, en Europa Occidental está entre 54.000 y 82.000 dólares. En España, estamos en **cifras similares** porque, pese a nuestro (cada vez menor) déficit de empresas de producto, tenemos grandes empresas (sobre todo bancarias) que se han volcado en este campo. Se puede decir que **es el momento para subirse a la ola de la ciencia de los datos**. Dentro de un par de años el mercado habrá madurado y las oportunidades estarán en otro lugar.

Espero que ahora tengas más claro qué tienes que hacer para convertirte en Data Scientist y qué futuro te espera.

¿Te animas?

iii Espero que sí !!!

