

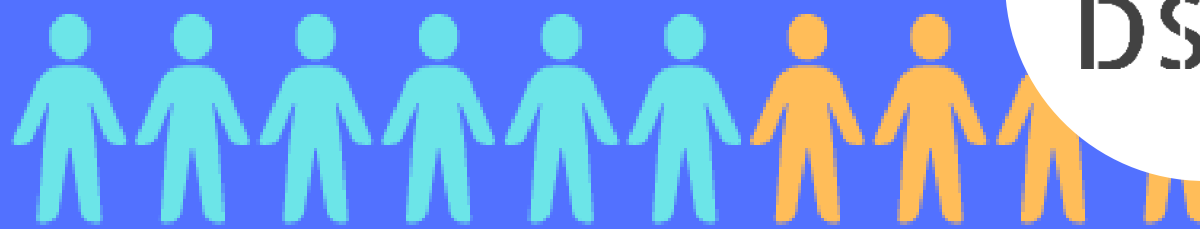
Guía Práctica - Edición 2022-2023

CÓMO EMPEZAR TU CARRERA EN DATA SCIENCE

*Para profesionales junior que quieren
comenzar su carrera en el mundo del
análisis de datos*



DS4B



Mensaje de Isaac González

Año tras año data science sigue entre los primeros puestos de las profesiones más demandadas.

Realmente no es nada nuevo.

Hace ya 20 años yo tomé una decisión (de forma totalmente aleatoria en aquel momento) de aprender lo que entonces se llamaba data mining.

Acababa de terminar la licenciatura en Psicología, y la verdad, en cuanto a perspectivas laborales no era de las titulaciones más buscadas por las empresas.

Pero durante la carrera había conocido la estadística, y me había gustado mucho, así que de nuevo, por más instinto que pericia, decidí seguir formándome por ahí y sobre todo en ese nuevo área llamado data mining.

El resultado fue que a diferencia de la mayoría de mis compañeros de promoción (que incluso años después seguían sin un empleo cualificado) yo no estuve ni un mes buscando trabajo.

El trabajo vino literalmente a buscarme a mí.

Y poco tiempo después las empresas ya no se fijaban en cual era mi licenciatura, si no en lo que sabía hacer.

Te cuento esto porque estoy viendo una segunda gran explosión en este área ahora mismo.

Más adelante en esta guía te mostraré rankings y noticias sobre esto.

Pero más allá de los datos, es algo que “se palpa” en el mercado.

Por tanto creo que si estás ahora comenzando tu carrera profesional Data Science es una de las mejores áreas en las que puedes focalizarte.

Y que concretamente 2023 será un excelente año para este ámbito.

Y espero que esta guía pueda ayudarte a decidir si esta oportunidad puede ser para ti y orientarte en tus primeros pasos.

Si te sirve me encantará saberlo: isaac@datascience4business.com

A tu disposición,

Isaac

Fundador DS4B



TABLE OF CONTENTS

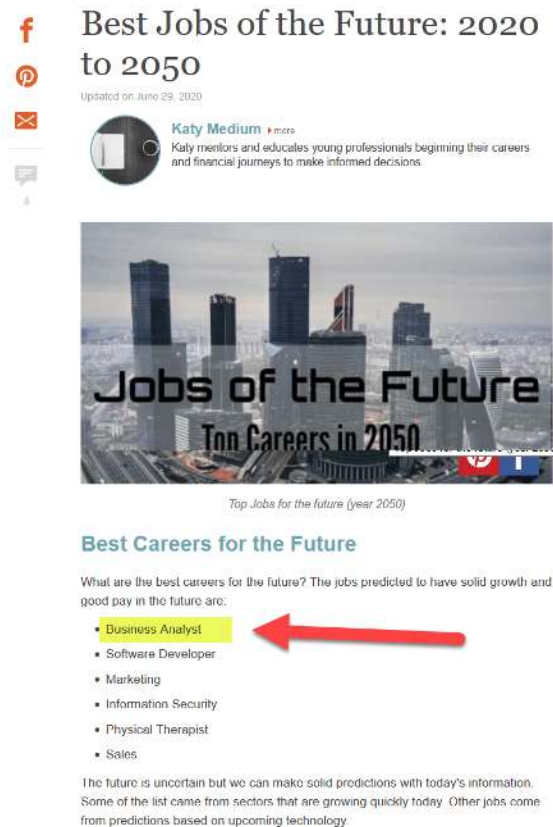
Mensaje de Isaac González	1
¿Es Data Science una buena opción laboral en 2022?	4
¿Es Data Science una buena opción para COMENZAR tu carrera?	6
Los dos tipos de Data Scientist	7
¿Qué herramientas debe dominar un Data Scientist?	10
¿Cómo conseguir un trabajo de Data Scientist si estás empezando tu carrera?	13
Próximos pasos	21

¿Es Data Science una buena opción laboral en 2023?

Data Science está continuamente en los rankings de las profesiones más demandadas y con mayor crecimiento.

Pero eso, a nivel de concepto seguramente ya lo sabes, o por lo menos lo intuyes.

Así que más que decírtelo de mi boca, que podrás creer más o menos, prefiero ponerte directamente unas muestras de algunos de esos rankings.



The Jobs Landscape in 2022

emerging roles, global change by 2022

133 Million

Top 10 Emerging

1. Data Analysts and Scientists
2. AI and Machine Learning Specialists
3. General and Operations Managers
4. Software and Applications Developers and Analysts
5. Sales and Marketing Professionals
6. Big Data Specialists
7. Digital Transformation Specialists
8. New Technology Specialists
9. Organisational Development Specialists
10. Information Technology Services

Número #1

Top 10 Declining

1. Data Entry Clerks
2. Accounting, Bookkeeping and Payroll Clerks
3. Administrative and Executive Secretaries
4. Assembly and Factory Workers
5. Client Information and Customer Service Workers
6. Business Services and Administration Managers
7. Accountants and Auditors
8. Material-Recording and Stock-Keeping Clerks
9. General and Operations Managers
10. Postal Service Clerks

declining roles, global change by 2022

75 Million

BBC Menú

NEWS | MUNDO

Noticias América Latina ¿Hablas español? Internacional Economía Tecnología Ciencia

Las 10 habilidades más demandadas por las empresas, según LinkedIn

Redacción
BBC News Mundo

18 enero 2019

f b t e Compartir

Habilidades duras

- Manejo de datos en la nube
- Inteligencia artificial
- Razonamiento analítico
- Manejo de personas
- Diseño UX (User Experience Design, en inglés)

El diseño UX, conocido como el diseño de la experiencia del usuario, es la destreza de adaptar los desarrollos tecnológicos y hacerlos accesibles o de fácil

Evolución de perfiles más demandados



¿Es Data Science una buena opción para COMENZAR ahora tu carrera?

¿Y qué pasa si estás en los primeros años de tu carrera profesional?

Si quizá acabas de terminar tu formación superior y estás decidiendo qué quieres hacer el resto de tu vida.

O si las has terminado hace 3 o 4 años pero todavía no has podido conseguir un trabajo cualificado y estable.

¿Sería adecuado en este caso?

De nuevo la respuesta es que sí.

Si hay algo seguro es que la digitalización y el avance de la tecnología siempre va a generar más datos.

Y por tanto ya siempre serán necesarios perfiles que sepan explotarlos y usarlos para mejorar los resultados de negocio de las empresas.



Los dos tipos de Data Scientist

Si en algún momento has mirado ofertas de puestos para data science posiblemente te hayas fijado en que parece que hay mucha variedad.

Mientras que unas tienen un marcado sesgo técnico y se centran en tecnologías como Hadoop, Spark, bbdd NoSql, Cloud, e incluso cosas como Java, C++, html o css.

Hay otras que se centran más en metodologías de análisis y aplicaciones funcionales como segmentación de clientes, prevención de abandono, identificación de fraude, etc.

Los motivos son varios.

Por un lado es cierto que muchas empresas no tienen claro lo que buscan, y no saben diferenciar entre los distintos roles (hablaremos más adelante sobre esto).

Pero por otro también es cierto que al final están persiguiendo 2 grandes objetivos distintos cuando buscan data scientist, pero a todo le llaman data scientist.

Hace aprox un año escuché un podcast americano (no recuerdo el nombre la verdad) en la que le pusieron un nombre a este efecto.

Y creo que realmente es muy descriptivo y ayuda, así que vamos a usarlo aquí también.

Se podría decir que hay 2 tipos de data scientist:

El tipo A, que viene de “Analyst”, que está más escorado hacia el negocio.

Se centra en crear valor y resultados de negocio analizando datos. Suele trabajar en áreas de negocio o en Centros de Excelencia.

Y las herramientas que domina están más orientadas al análisis que a la puesta en producción.

Y el tipo B, que viene de “Builder”, que está más escorado hacia la tecnología y se centra en crear sistemas o aplicaciones que utilicen data science o machine learning.

Su objetivo no es tanto llegar a los resultados de negocio, si no construir esos sistemas con los que otros puedan llegar a los resultados de negocio. Suele estar situado en IT.

Y las herramientas que domina son más tecnológicas.

Data Scientist Tipo A

- Analyst
- Escorado hacia negocio
- Crear valor y trasladar a resultados
- Situado en áreas de negocio o COEs
- Python, SQL, R, Tableau, ...

Data Scientist Tipo B

- Builder
- Escorado hacia la tecnología
- Crear sistemas y data products
- Situado en IT
- Cloud, Big Data, Python, NoSQL, ...

Es fundamental que domines este concepto porque te va a permitir:

1. Saber a cual de los 2 se acerca más tu perfil
2. Identificar en una oferta laboral cual de los 2 están realmente buscando
3. No frustrarte si ves un montón de requisitos para los que no cualificas porque realmente un data scientist “del otro tipo” es como un role diferente, aunque como te decía antes las empresas lo mezclen y no sean capaces de expresar claramente lo que buscan

El data scientist de tipo B se corresponde de forma más ajustada con el rol que se llama “data engineer”. Incluso últimamente está surgiendo un término para ponerle nombre a la posición que estaría en medio de data engineer y data scientist de tipo B, que se llama “machine learning engineer”

Pero todo esto es muy dinámico y los nombres cambian. Por eso lo importante es que conozcas estos conceptos base para poder “leer entre líneas” lo que se está buscando en cada caso.



¿Qué herramientas debe dominar un Data Scientist?

Si haces caso de lo que oyes en el mercado un data scientist debería dominar:

- Herramientas analíticas: Python, R, SQL...
- Metodologías: estadística, data science, machine learning, computer vision, NLP, análisis de series temporales, ...
- Herramientas Big Data: Hadoop, Spark, Hive, Pig, Scoop, Flume, Storm, ...
- Plataformas Cloud: AWS, Azure, GCP, ...
- Herramientas de Visualización: PowerBI, Tableau, Clik, ...
- Bases de datos relacionales: Oracle, SQL Server, MySQL, ...
- Bases de datos NoSql: Cassandra, MongoDB, Neo4j
- Desarrollo de software: Java, C++
- Desarrollo web: Html, CSS
- Metodologías de negocio: marketing, CRM, Supply Chain Maganement, Riesgos, Fraude, mantenimiento preventivo, ...
- Capacidad de convertir resultados en negocio
- Capacidad de comunicación a stakeholders
- Etc

¿En serio crees que alguna persona domina todo eso?

¿En serio crees que Gustavo o Marta, que fueron contigo a la facultad y ahora ves en LinkedIn que son data scientist dominan todo eso?

Este es quizá el principal mito de la profesión, alimentado por todo esa confusión y ruido del mercado.

Es necesario pasar del mito a la realidad.

Y lo vamos a hacer con lo que mejor se nos da: analizar datos.

Por eso en DS4B hemos decidido invertir unos cuantos días en analizar realmente qué es lo que el mercado está pidiendo a un data scientist en términos de herramientas, competencias y habilidades.

Y hemos analizado las más de 2.000 ofertas existentes en LinkedIn para la búsqueda “data scientist en España”

Los resultados son extremadamente reveladores.

Entre otras cosas los datos identifican claramente y sin lugar a dudas que de todas las tecnologías y competencias de las que hablábamos al principio de esta sección se puede optar a más del 70% de las ofertas de data science con sólo 6 de ellas.

- Python (o R)
- Data Science
- Machine Learning
- Estadística
- Sql
- Tableau

MUY IMPORTANTE: Por favor pon mucha atención en lo que acabas de leer, ya que es una clave que te va a ahorrar años de trabajo, y es posiblemente el consejo más valioso de toda esta guía.

Lo repito: puedes optar a más del 70% de las ofertas de data science con solo tener las 6 capacidades indicadas.

No es algo fácil, pero es algo factible en un tiempo razonable de 6-12 meses.

Céntrate en ello, consigue tu primer puesto de Data Scientist, y a partir de ahí ya seguirás aprendiendo.

Llegados a este punto, y lo que ya llevas leído de esta guía, espero que ya tengas más claridad sobre:

- Si data science puede ser una opción deseable para reconvertir tu carrera o simplemente iniciarla
- Por qué estamos en el mejor momento para hacer esa transición
- Por qué ves tanto ruido en el mercado y tanta información diferente e incluso contradictoria
- Los 2 tipos fundamentales de data scientist que existen
- A cual te acercas más o a cual te quieres dirigir
- Cómo leer entre líneas para interpretar lo que están buscando las empresas en cada caso
- Por qué no debes creer todo ese bullshit del “data scientist moderno” que tiene que saber todo de todo
- Las 6 competencias concretas que te permiten optar al 70% de las ofertas de data science

Pero posiblemente estés pensando “vale, ya tengo las cosas mucho más claras, pero cómo puedo usar toda esta información para transformarla en pasos operativos para los próximos meses”

Así que voy a darte el framework que suelo utilizar cuando alguien me dice que quiere seguir este camino.



¿Cómo conseguir un trabajo de Data Scientist si estás empezando tu carrera?

Tu gran problema es básicamente que no tienes nada de experiencia. No tienes nada que ofrecer más allá de tu conocimiento técnico.

E incluso con él, la pregunta es “¿y cómo me diferencio frente a otros candidatos que también tienen un gran conocimiento técnico como yo?”

Por tanto en este caso los pasos que tienes que dar son 3:

1. Elegir un rol y capacitarte técnicamente para él
2. Darte una capa de negocio que no tengan otros candidatos
3. Demostrar con pruebas lo que sabes hacer

Vamos a ver este framework paso a paso.



Paso 1: Elegir un rol y capacitarte técnicamente para él

Dado que el componente técnico sí va a tener mucho peso en tu caso tu gran reto va a ser cómo conseguirlo en el menor tiempo posible.

Y para eso vas a tener que hacer dos cosas:

1. Seleccionar el rol que más se ajuste a tu perfil y apostar por él
2. Conocer lo que está pidiendo el mercado para separar el trigo de la paja y centrarte justo en aquello que te dará más oportunidades laborales

Sobre el punto 1 ya hemos hablado, es básicamente decidir entre data scientist de tipo A o de tipo B.

Y sobre el punto 2 también. Esperablemente me habrás hecho caso y te habrás centrado en empezar por dominar las 6 competencias comentadas.



Paso 2: Darte una capa de negocio que no tengan otros candidatos

Asumimos que ya tienes un conocimiento técnico potente.

Pero también que será difícil que seas quien tiene el mayor conocimiento técnico de todos los candidatos.

Por tanto por ahí es difícil destacar.

Pero hay otra ruta más inteligente.

Una que la mayoría de los data scientist dejan desatendida: la dimensión de negocio.

Y aquí no se trata tanto del conocimiento de sectores, mercado, etc, que lógicamente todavía no tienes.

En un junior lo que se busca en este punto es el enfoque y la forma de pensar.

Y créeme que es más difícil de encontrar de lo que parece.

Y las empresas cuando lo encuentran no lo quieren soltar.



La mayoría de los junior tienen un pensamiento técnico.

Piensan que su trabajo es únicamente ejecutar lo que les dicen aplicando su conocimiento técnico.

Y no se dan cuenta de que realmente esas tareas en la empresa se hacen siempre con un objetivo.

Que suele ser ganar dinero en alguna de sus formas.

Puede ser generando nuevos ingresos.

Puede ser reduciendo los datos.

O lo que sea, pero al final las cosas se hacen con un objetivo.

Si tu eres capaz de demostrar que tienes los conocimientos técnicos pero además que entiendes que tienen que estar al servicio de un resultado serás increíblemente magnético para una empresa.

Cuando yo empecé a trabajar en esto lo que me gustaba era hacer modelos.

No quería mayores salarios, no quería promociones.

Básicamente lo único que me importaba era que alguien me pagara un sueldo que me permitiera vivir mientras dedicaba todo el día a hacer lo que me gustaba (los modelos).

Una vez mi gerente me dijo “en la empresa no hay que tener el mejor modelo, si no el mejor modelo posible teniendo en cuenta los recursos que tenemos para hacerlo”.

Cuando por fin entendí aquello (y lo extrapolé a todo lo que hacía) fue cuando mi carrera se empezó a disparar.

Y al final seguía haciendo lo que me gustaba, pero acompañado de mejores salarios y promociones, que al final a nadie le amarga un dulce.

Por tanto, como paso 2 del framework tienes que incorporar la dimensión de negocio desde el principio.

Aprende a identificar objetivos, aprende a aterrizarlos en métricas medibles, aprende a conseguirlos mediante técnicas de data science y aprende a comunicarlos en términos de negocio.

De nuevo, si haces eso, te diferenciarás ampliamente del resto de juniors y tendrás un gran atractivo para las empresas.

Ahora bien, no sólo hacer falta saber hacer todo lo anterior, si no también poder demostrarlo, y eso lo hacemos en el paso 3 del framework.



Paso 3: Demostrar con pruebas lo que sabes hacer

Cada vez más los procesos de selección se hacen por competencias.

Es decir, no importa tanto lo que digas, ni siquiera importa tanto los títulos que tengas, lo que importa es lo que sabes hacer.

Pero tienes que demostrar que sabes hacerlo.

Para ello es posible que te planteen tres tipos de pruebas:

1. Pruebas técnicas in situ
2. Prueba para hacer en casa
3. Evaluación de competencias durante la entrevista

La primera consiste en tener que resolver pequeños problemas, normalmente muy centrados en la parte de programación, o también en resolver en tiempo real sencillos casos analíticos, como hacer un modelo de XGBoost con Sklearn por ejemplo.

Éstas son las más sencillas si tienes el conocimiento técnico, porque al final si sabes hacerlo simplemente te pones en el ordenador y lo resuelves.



En la segunda, prueba para hacer en casa, no importa tanto la parte técnica, ya que tienes varios días para hacerla, puedes mirar en internet, preguntar a tu amigo, etc.

Por tanto lo que importa es más la integración del conocimiento técnico con la dimensión de negocio que te explicaba antes.

Lo que quieren ver es “cómo piensas” y cómo planteas la resolución de un problema.

Usando una metodología, poniendo foco en los objetivos y aplicando los conocimientos técnicos para generar un entregable o un resultado.



Y la tercera consiste en hacer esa resolución de competencias en la propia entrevista.

Es posible que también te planteen un caso y sirva con que expliques con lápiz y papel cómo lo resolverías.

Pero aquí también tienes otro gran recurso que te permitirá diferenciarte de forma mucho más profesional y atractiva: el portfolio.

Tener un portfolio es algo habitual en otras profesiones como arquitecto o diseñador web.

Y cada vez empieza a ser más usado por data scientist, especialmente los junior.

La lógica es muy clara: si no tengo experiencia previa real que enseñarte al menos te enseñé proyectos que he hecho por mi cuenta.

Ello demuestra que sabes hacerlos, porque es algo que puedes enseñar.

Pero también demuestra que estás comprometido, y que, aunque ahora mismo no tengas trabajo todavía, no te pasas el día jugando a la play o de compras, si no que inviertes el tiempo en seguir aprendiendo y mejorando.

Son mensajes muy positivos para una empresa.

Vale, pero ¿cómo creo un portfolio?

Pues sí, lo has adivinado, también tengo un framework para ello 😊

1. Decide si eres tipo A o tipo B
2. Elige un par de sectores en los que te gustaría trabajar
3. Vete a [kaggle.com](https://www.kaggle.com), crea una cuenta gratuita, vete a la sección de datasets y busca unos datos del sector elegido en el punto 2 y que te permitan demostrar las habilidades del tipo elegido en el punto 1
4. Resuelve el caso aplicando una óptima integrada entre negocio y técnica
5. Publica tu caso, puede ser en wordpress, en github, o en lo que te resulte más sencillo
6. Imprímelo y encuadérnalo para llevarlo también físicamente a las entrevistas por si no hay internet o no funciona

Próximos pasos

Bueno, pues hemos llegado al final de esta guía.

Sé que han sido muchos conceptos en poco tiempo.

Pero realmente son conceptos muy valiosos. Léelo una segunda vez si es necesario e intenta comprenderlos de forma profunda.

Y sobre todo ponlos en práctica.

Por último recuerda que en DS4B estamos para acompañarte en este camino.



Al haber descargado esta guía ya estás en nuestra NEWSLETTER. Recibirás aprox un email cada semana. Por supuesto puedes darte de baja cuando quieras pero te recomiendo estar ahí porque es el hub desde el que comparto todo el contenido y recursos de alto valor que te ayudarán enormemente en este viaje



Si quieres ver más contenido relacionado con data science, tanto técnico como de desarrollo de carrera suscríbete a nuestro canal de [YOUTUBE](#)



Si estás empezando en Data Science / Big Data y quieres una visión general que te ayude a situarte te recomiendo mi libro [Big Data para CEOs y Directores de Marketing](#).

Y para cualquier otra cosa puedes localizarme en isaac@datascience4business.com

Un placer y hasta la próxima!

Isaac