

➔ CONSTRUYENDO EL FUTURO

Guía en análisis de datos

Caterina Abanoni



CONTENIDO

Al final de esta guía descubrirás si el análisis de datos es para ti y estarás listo para dar el siguiente paso, **nosotros te acompañamos**.

👋 **PARTE 0: Hola**

- A. ¿Por qué leer éste ebook?

💻 **PARTE I: Fundamentos**

- A. Introducción
- B. ¿Qué son los datos?
- C. ¿Cómo los datos son útiles?
- D. Tipos de datos
- E. ¿De dónde vienen los datos?
- F. Big data vs traditional data

🎓 **PARTE II: Tu carrera en datos**

- A. ¿Qué es realmente el análisis de datos?
- B. ¿Cómo sé si esto es para mí?
- C. ¿Qué oportunidades hay?
- D. Diferencias entre ésta carrera y otras del mundo tech
- E. Beneficios de ser analista de datos

🌐 **PARTE III: Términos del mundo de los datos**

- A. Business Intelligence
- B. Data Science o ciencia de datos
- C. Análisis de datos
- D. Ingeniero de datos
- E. Inteligencia artificial
- F. Machine Learning
- G. Deep Learning

🧙 **PARTE IV: Roles en el mundo de los datos**

- A. El analista de datos
- B. El analista de business intelligence
- C. El científico de datos
- D. El ingeniero de datos

💻 **PARTE V: Conocimientos técnicos**

- A. Almacenamiento de datos
- B. ETL | Extracción, Transformación y Carga
- C. Visualización de datos
- D. Lenguajes de Programación
- E. Estadística

📍 **PARTE VI: ¿Como sigo?**

- A. ¿Qué trabajos podría conseguir si estudio datos?
- B. Aplicación según tu perfil
- G. Palabras finales



PARTE 0

> HOLA 😊

A. ¿POR QUÉ LEER ESTE EBOOK?

Hola, seguramente te estas preguntando porque deberías leer este ebook.

¿Te identificas con algunas de estas situación?

- 1 ¿Te encuentras en un **punto muerto en tu carrera profesional** y buscas una forma de avanzar hacia nuevas oportunidades?
- 2 ¿Te sientes **atrapado en una rutina profesional** que no te satisface y buscas una vía para desafiar y crecer a nivel profesional?
- 3 ¿Estás experimentando la presión de la **automatización y la digitalización** en tu industria y te preguntas cómo mantenerte relevante?
- 4 ¿Te encuentras en un **mercado laboral competitivo** y buscas una ventaja que te destaque entre otros profesionales?

5 ¿Te inquieta la **incertidumbre laboral** y te gustaría adquirir habilidades que te brinden estabilidad en tu carrera?

7 ¿Has oído hablar del análisis de datos, pero te sientes **abrumado** por no comprender exactamente en qué consiste y cómo podría beneficiarte?

8 ¿Te encuentras en un punto en tu vida en el que deseas un **cambio significativo** y consideras el análisis de datos como una vía para ello?

Entonces este ebook es para ti.

El objetivo de éste ebook es llevarte del punto A al punto B.



En el punto A, te encuentras en un lugar donde no solo careces de conocimientos sobre análisis de datos, sino que también puede que no comprendas la esencia de la profesión y las oportunidades que ofrece. Además, es posible que te sientas perdido al intentar dar tus primeros pasos en este fascinante campo.

6 ¿Sientes que estás **desaprovechando tu potencial** y te gustaría descubrir cómo liberar tu creatividad y habilidades en un nuevo campo?

EL punto B, representa un cambio significativo. Aquí, adquirirás conocimientos sólidos sobre qué es el análisis de datos y todos esos términos raros que escuchamos, tendrás una comprensión clara de las oportunidades que se presentan y sabrás diferenciar los distintos roles que existen.

¿Y porque deberías escucharme a mí?

En el 2018 yo estaba en una situación donde buscaba **crecimiento personal** y pasé por muchas de las etapas que pasan las personas que quieren aprender **análisis de datos**.

Esta **guía** está elaborada pensando en todo lo que me faltó tener cuando inicié una carrera en datos. Es una guía elaborada por una analista de datos **para otros analistas de datos**. Espero que te guste, lo hice con mucho cariño.

Te veo pronto.



PARTE I

> FUNDAMENTOS 

A. INTRODUCCIÓN

En las economías modernas, los datos están **en todas partes**. Hemos producido más datos en los últimos dos años que en la historia de la humanidad. **Usas y creas datos todos los días**.

¿Alguna vez has leído reseñas de un producto antes de decidir si comprarlo o no?

Eso es **analizar datos**. O tal vez has usado un rastreador de actividad física para contar tus pasos y mantenerte activo durante todo el día. Cada vez que usas tu teléfono, buscas algo en línea, transmites música, compras con una tarjeta de crédito, publicas en redes sociales o utilizas GPS para trazar una ruta, estás creando datos. Nuestro **mundo digital** y los millones de **dispositivos inteligentes** en su interior han hecho que la cantidad de datos disponibles sea realmente alucinante.



B. ¿QUÉ SON LOS DATOS?

Pero primero, ¿qué son exactamente los datos?

Los datos son una **colección de hechos**. Es el registro de lo que ha ocurrido.

Eso **puede responder** cosas como: quién, cuántos, qué tan grande, dónde, qué categoría, qué color, etc.

Esta **colección** puede incluir números, imágenes, videos, palabras, medidas, observaciones y más.

C. ¿CÓMO LOS DATOS SON ÚTILES?

Saber exactamente lo que ha sucedido históricamente nos permite tomar **mejores decisiones** sobre lo desconocido o en el futuro. Ver el futuro sería un poder asombroso, pero imposible. Ver claro como el cristal en el pasado para **entender patrones y tendencias** para ayudarnos a predecir el futuro es lo más cercano que tenemos, y por lo tanto tenemos datos. Los datos nos permiten ver qué **resultados** se produjeron debido a los actos y esfuerzos anteriores.

D. TIPOS DE DATOS

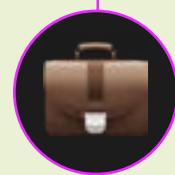


Datos estructurados



Los datos estructurados son cualquier **dato o información** que se encuentre en un campo fijo dentro de un **registro o archivo definido**. Normalmente son bases de datos y hojas de cálculo.

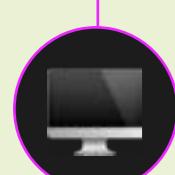
Son datos que están organizados de una manera pre-determinada, normalmente en **filas o columnas**. Los datos estructurados se manejan normalmente con **lenguaje SQL**.



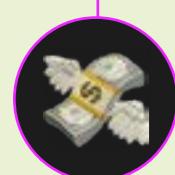
Una empresa puede sacar mucho potencial de los datos estructurados. Los ejemplos más comunes son los datos de los clientes, las ventas, financieros, informes transaccionales, número de visitas a la página web. Actualmente estos son los datos que ofrecen la mayor cantidad de información para las empresas.



Los datos estructurados se dice que **son solo un 20%** de los datos disponibles en el mundo. El restante 80% son datos en un formato NO estructurado.



Por ejemplo sí tiene datos estructurados sobre las visitas a su página web y le dicen que las mismas han caído un 25% en el último mes, necesitará otro tipo de datos para saber porqué.

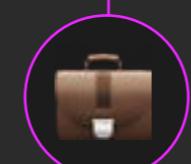


Los **aspectos positivos** es que son baratos, fáciles de almacenar y fáciles de analizar.

Datos no estructurados o semi estructurados



Sin embargo los avances, como internet, la tecnología de sensores, la computación en la nube o cloud computing y nuestra capacidad para almacenar y analizar datos ha cambiado el **tipo y cantidad** de datos que podemos recopilar.



Ahora las actividades diarias como ir por la calle, meterse en el auto o comprar algo en una tienda generan un rastro de datos cada vez mayor, tanto estructurado como no estructurados. Las compañías pueden utilizar todos estos datos diferentes para **hacer negocios**.



Los datos no estructurados se refieren a cualquier dato que no encaje bien en los formatos estructurados tradicionales.



Ejemplos: conversaciones de correo electrónico, texto de páginas web, publicaciones en redes sociales, contenido de video, de fotos y de grabaciones de audio , etc.



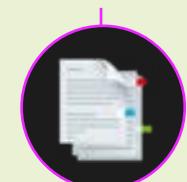
Contienen mucha **densidad de texto**, pero también pueden contener datos como fechas y número u otro tipo de datos como imágenes.



Hasta hace poco , todo lo que no encajaba en bases de datos u hojas de cálculo se rechazaba con frecuencia o almacena un papel, microfichas o archivos escaneados que no se podían analizar fácilmente.



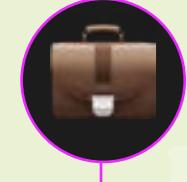
Datos internos



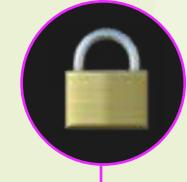
Los datos internos se refiere a **toda información** que tiene actualmente su empresa o que tiene el potencial de recopilar.



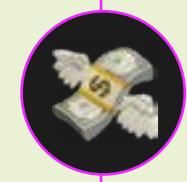
Pueden ser **estructurados** (listado de clientes o registro de transacciones) o **no estructurados** (como los datos de conversaciones de llamadas de atención al cliente o feedback a partir de entrevista con empleados.)



Se trata de los **datos privados** o en propiedad que posee una empresa. (Esto significa que solo la empresa tiene acceso a ellos)



La desventaja es que los datos internos usted es el encargado de mantenerlos y protegerlos, especialmente los personales.



Sin embargo cuando compra datos externos, el proveedor de los datos ya ha asumido esa responsabilidad y esa obligación por usted.



Datos Externos

Los datos externos se refieren al **infinito despliegue de información** que existe afuera de su organización, estos pueden estar disponibles públicamente (como datos del gobierno) o ser propiedad de terceras partes (como amazon) y también en formato **estructurado y no estructurado**.

E. ¿DE DÓNDE VIENEN LOS DATOS?

Los datos tradicionales pueden provenir de **registros básicos** de clientes o **información histórica** de precios de acciones. Los registros básicos de clientes pueden contener información como la identificación del cliente, cuántas veces han presentado una queja, la cantidad de dinero gastada en una sola compra, si son parte de un esquema de miembros, su dirección, información de contacto, etc. Los datos históricos de precios de acciones, pueden contener fechas de las observaciones, los nombres

Los datos se generan a partir de tres fuentes: Personas, Máquinas y Corporaciones.

PERSONAS

Generan datos de Redes Sociales (Facebook, Twitter, Instagram, LinkedIn) correos electrónicos, Internet, documentos, entre otros.

MÁQUINAS

Generan datos de sensores, satélites, archivos de registro de computadora, cámaras, máquinas de secuenciación genética, telescopios espaciales, sondas, entre otros.

CORPORACIONES

Datos generados de transacciones, sistema administrativo, tarjetas de crédito, sistema financiero, contabilidad, comercio electrónico, ventas, registros médicos, investigación, entre otros.

Las máquinas generan más datos que las corporaciones y las personas.

Produce **100 (cien)** veces más datos que las corporaciones y **10 (diez)** veces más datos que las personas.

En resumen:

Máquinas = 100X (cien)

Corporaciones = 1X (uno)

Personas = 10X (diez)

Algunos ejemplos de datos generados por máquinas son:

- 1 Sensores de todo tipo integrados en equipos.
- 2 Software que realizan cálculos
- 3 Sistemas de aprendizaje automático.
- 4 Sistemas de automatización.
- 5 Clústeres de computadoras.
- 6 Telemetría satelital.
- 7 Sistemas GPS.
- 8 Máquinas en general.
- 9 Chips RFID
- 10 Teléfonos celulares

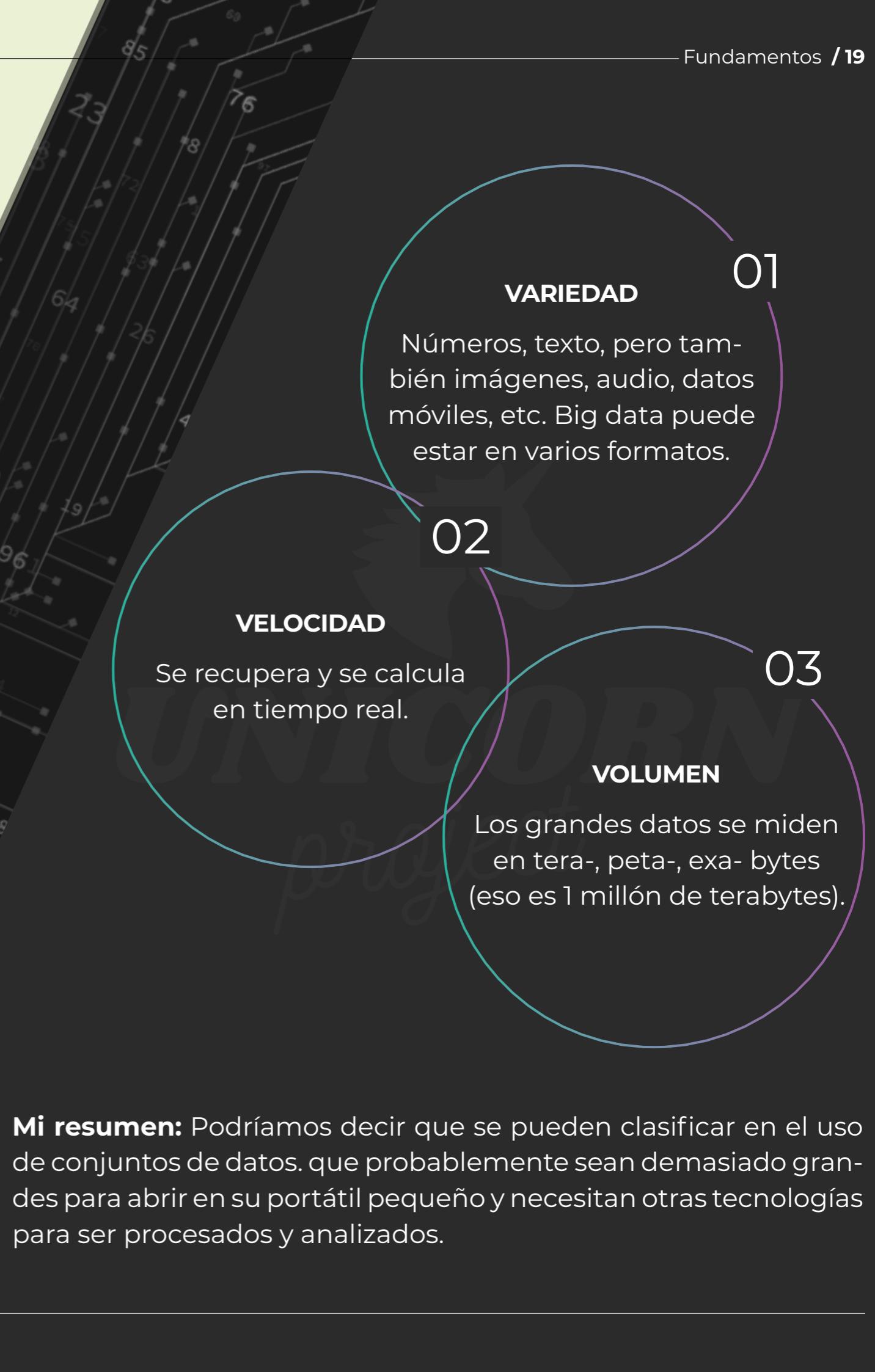


F. BIG DATA VS TRADITIONAL DATA

Los datos son la base de la ciencia de datos; es el material en el que se basan todos los análisis. En el contexto de la ciencia de datos, hay dos tipos de datos: **tradicional y big data**.

Los datos tradicionales son datos **estructurados** y almacenados en **bases de datos** que se pueden administrar desde una computadora; está en **formato de tabla**, que contiene valores numéricos o valores de texto. El término “datos tradicionales” no es parte del oficial vernáculo. Es algo que estamos introduciendo con claridad. Creemos que ayuda a enfatizar la distinción entre big data y... non-big data.;

Big data es más grande que los datos tradicionales, pero no en el sentido banal. Se trata de **datos extremadamente grandes**, distribuidos a través de una red de computadoras, pero no es solo caracterizado por su volumen. Estos datos pueden estar en varios formatos; puede ser **estructurado, semiestructurado o no estructurado**. A menudo verá grandes datos caracterizados por la letra “V”. Esto se deriva de “**las 3V de los grandes datos**”:



Mi resumen: Podríamos decir que se pueden clasificar en el uso de conjuntos de datos que probablemente sean demasiado grandes para abrir en su portátil pequeño y necesitan otras tecnologías para ser procesados y analizados.

► TU CARRERA EN DATOS

PARTE III

A. ¿QUÉ ES REALMENTE EL ANÁLISIS DE DATOS?

El análisis de datos es una disciplina o rama dentro del mundo de los datos.

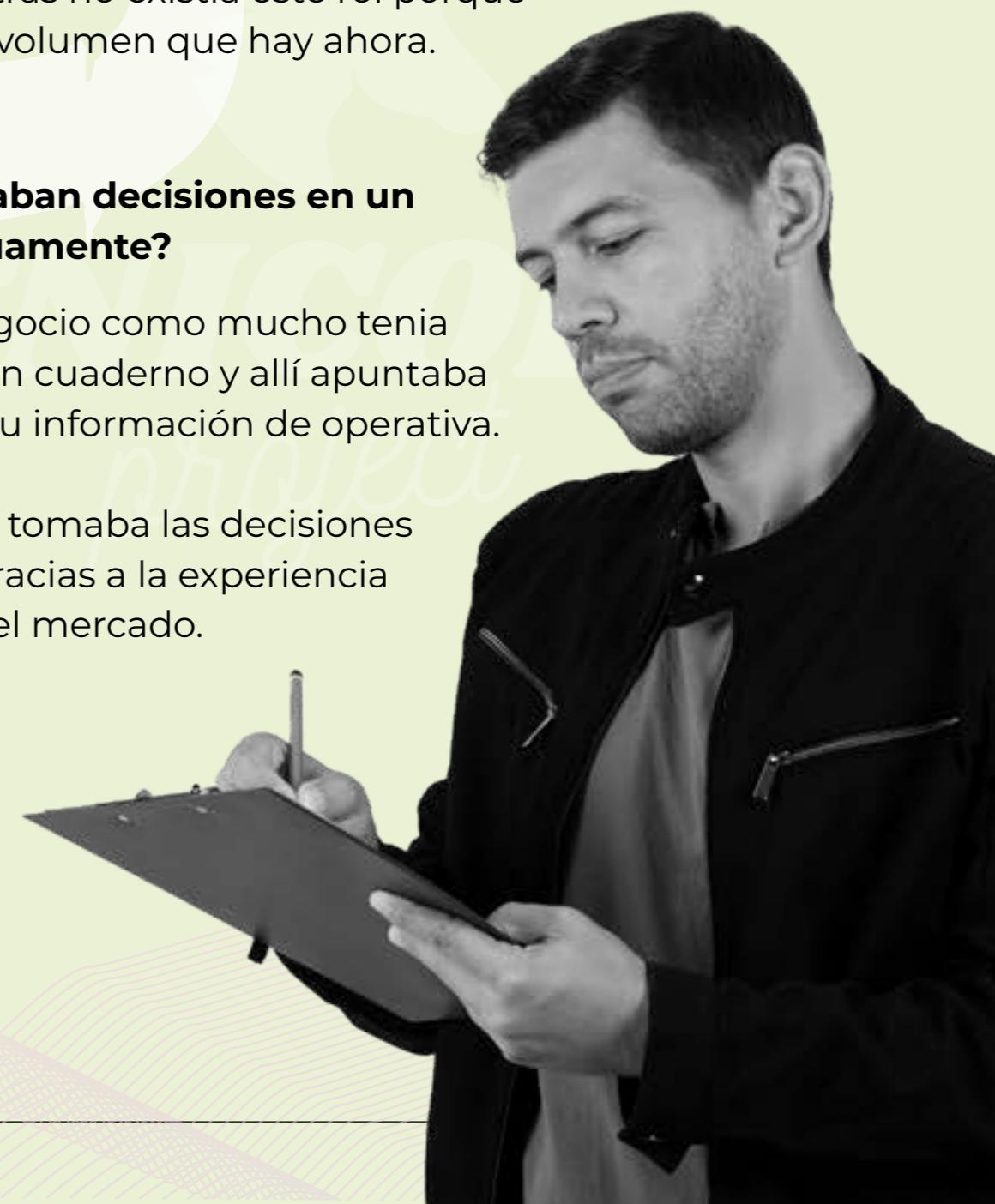
Este tipo de roles enfocados a extraer valor de los datos surgen por la **cantidad** de datos que existen hoy en día.

Hace 50 años atrás no existía este rol porque no teníamos el volumen que hay ahora.

¿Como se tomaban decisiones en un negocio antiguamente?

El dueño de negocio como mucho tenía un bolígrafo y un cuaderno y allí apuntaba sus finanzas o su información de operativa.

Pero la persona tomaba las decisiones por intuición, gracias a la experiencia que llevaba en el mercado.



Vamos a suponer que tenemos un dueño de un supermercado que se llama Ricardo.

Ricardo lleva operando este supermercado durante 20 años.

Cuando le preguntas a Ricardo como sabe que los viernes debe tener más stock en su supermercado te responde: "Yo llevo 20 años aquí y por eso sé que los viernes es el día que mas vendemos".

Pero hoy en día no necesitamos 20 años de experiencia para saber eso. Y así con todo.

Nuestra capacidad como humanos para recolectar la información se ha disparado y cada vez irá a más.

Aquellas personas que logren sacar provecho de toda esa información serán los que tendrán **el poder**.

Hoy en día hay diferentes formas de generar datos , pero en el mundo empresarial lo más común es con las **herramientas o software tecnológicas** que usa la empresa en su día a día.

Todas estas herramientas generan información mucho mas rápido de lo que puede un humano si lo hace manualmente.

Se genera con código y de forma automática.

Por eso **cada vez generamos mas información**.

El análisis de datos consta de muchas tareas diferentes pero todas están enfocadas a hacer crecer un negocio.

Estas son algunas de las formas en las que un analista de datos puede ayudar en una empresa:



Eficiencia operativa:

Analizar los datos nos permite identificar áreas de ineficiencia en las operaciones de una empresa, lo que permite optimizar procesos, reducir costos.



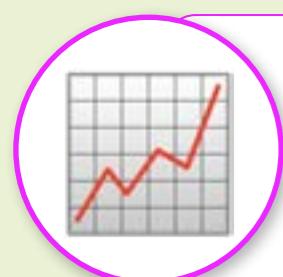
Predicción y pronóstico:

Un objetivo importante es utilizar datos históricos para predecir el futuro. Esto puede incluir pronósticos de ventas, demanda de productos, tendencias del mercado y más.



Identificar oportunidades:

Se puede descubrir nuevas oportunidades de mercado, nichos de clientes, tendencias emergentes y demandas no satisfechas, lo que puede llevar a la expansión del negocio y al aumento de los ingresos.



Tomar mejores decisiones:

Tomar decisiones basadas en evidencia y números. No en intuición como nuestro amigo Ricardo



Mejorar la calidad de productos y servicios:

El análisis de datos recopila retroalimentación y opiniones de los clientes, lo que ayuda a mejorar la calidad y a responder a las necesidades del mercado de manera más efectiva.



Detección de fraudes y riesgos:

En sectores como las finanzas, el análisis de datos es esencial para detectar actividades fraudulentas y gestionar riesgos. Permite identificar patrones anómalos que podrían indicar fraude o riesgo crediticio, lo que protege los activos de la empresa.



Medición de Resultados y Rendimiento:

Los analistas de datos desarrollan métricas y KPIs para evaluar el éxito de los proyectos y estrategias. Esto permite a la empresa medir su rendimiento y ajustar sus acciones en consecuencia.

B. ¿CÓMO SÉ SI ESTO ES PARA MI?

La mayoría de las personas se cree que para saber análisis de datos tenés que ser ingeniero de la nasa o de carreras STEM (Ciencia, matemáticas, ingeniería o tecnología) **La realidad es que no.**

El mundo de los datos ofrece una **versatilidad importante de roles y oportunidades** que hacen que tengamos para elegir entre roles poco técnicos hasta roles mucho más técnicos.

Un analista de datos es un rol que tiene:



¿Qué quiere decir esto?

Que es un **mix de varios conocimientos**.

Entonces pongamos un ejemplo, Marta trabaja en **marketing** y no sabe si puede ser analista de datos.

Ella trae conocimientos previos en marketing y ha trabajado para una empresa de ecommerce.

¿Marta puede ser analista de datos?

Claro que si , lo único que tiene que trabajar Marta son las **herramientas nuevas** porque el conocimiento de industria de ecommerce ya lo trae y podría trabajar como **analista de datos en una ecommerce**.

Porque conoce el negocio, como funciona la empresa. Y traé conocimientos en un área funcional. Que es marketing.

Lo mismo sucede para personas que trabajan en contabilidad, administración, logística, áreas comerciales, atención al cliente.

Ya traen unos conocimientos de **industria** y de **área**.

Solo necesitamos aprender a combinarlos bien.

Lo único que le falta a la mayoría de las personas son las herramientas técnicas, saber como explotar esas habilidades que ya traen con conocimientos en **tecnología adaptada a procesos de datos**.

Aún así les dejo algunas características de personas que les suele gustar el análisis de datos por si te identificas:

- 01 Intolerancia a la ineficiencia, incompetencia o vagos.
- 02 Obsesionados con los patrones.
- 03 Seres pensantes y resolutivos.
- 04 Amantes de la automatización.
- 05 Obsesionados con productividad.
- 06 Increíblemente curiosos.
- 07 Amantes de Excel.
- 08 Independientes.
- 09 Amantes de lo racional.
- 10 Perfeccionistas.

Así que en resumen:

Cualquier persona que venga de carreras empresariales o derivados puede ser analista de datos.



C. ¿QUÉ OPORTUNIDADES HAY?

1. La generación de datos

Statista en su reporte informa que para 2025, los ingresos anuales por la generación de datos alcanzarán los **68 mil millones de dólares** con la creación de **181 zettabytes de datos**.

Los datos de los clientes se han convertido en la moneda más rica del mundo. A partir de él, **las empresas pueden determinar todo tipo de información como:**

02

¿A qué mensajes y campañas responden bien?



01

Patrones de comportamiento de sus consumidores.



03

¿Qué plataformas resultan más beneficiosas?



04

Qué datos demográficos muestran respuestas positivas y mucho más.



Toda esta información se recopila sin problemas y por escala. Por lo tanto, es crucial que las empresas puedan darle sentido a estos datos y, para hacerlo, necesitan analistas de datos talentosos.

La información es el petróleo del siglo XXI y la analítica es el motor de combustión. - Peter Sondergaard

2. Los analistas de datos están en demanda

La Oficina de Estadísticas Laborales de EE. UU. predice que la demanda de trabajos de análisis de datos **crecerá un 23 % entre 2021 y 2031**, mucho más rápido que el promedio del **5 % para todas las demás industrias**.

Además de esto, un informe realizado por Analytics Insight predijo ofertas de trabajo para más de 3 millones de profesionales de datos el año pasado.

Los últimos años, especialmente desde la pandemia, han dado paso a un aumento astronómico en la forma en que se recopilan e interpretan los datos para la inteligencia empresarial y la toma de decisiones.

A medida que surjan nuevas tecnologías digitales para ayudar y mejorar las empresas, los analistas de datos serán cada vez más valiosos, lo que conducirá a una fantástica progresión profesional y seguridad laboral.



Los salarios son buenos y tienes la posibilidad de trabajar en tecnología



La remuneración promedio de efectivo adicional para un analista de datos en Argentina es de \$ 337.054, con un rango de entre \$ 181.418 y \$ 541.934. Las estimaciones de sueldos se basan en 887 sueldos enviados anónimamente a Glassdoor por empleados con el cargo de analista de datos en Argentina.



El sueldo promedio de analista de datos es COP 7,415,351 por mes en Bogota, Colombia. La remuneración promedio de efectivo adicional para un analista de datos en Bogota, Colombia es de COP 3,671,226, con un rango de entre COP 1,047,115 y COP 9,789,935. hace 6 días



glassdoor.com.mo

<https://www.glassdoor.com.mx> › [Sueños](#) › [analista-de-d...](#)



El salario analista de base de datos promedio en México es de \$ 120,000 al año o \$ 61.54 por hora. Los cargos de nivel inicial comienzan con un ingreso de \$ 87,005 al año, mientras que profesionales más experimentados perciben hasta \$ 180,000 al año.

 Talent.com
<https://mix.talent.com/salary/jobanalista-de-base-de-...>



La remuneración promedio de efectivo adicional para un analista de datos en Montevideo, Uruguay es de **UYU 64.000**, con un rango de entre **UYU 45.000** y **UYU 65.000**. Las estimaciones de sueldos se basan en 34 sueldos enviados anónimamente a Glassdoor por empleados con el cargo de analista de datos en Montevideo, Uruguay. 1 nov 2023



classadoor.com.ar

<https://www.glassdoor.com.br> | Sobre | Sair



¿Cuánto gana un Analista de datos en España? El salario analista de datos promedio en España es de € 23.300 al año o € 11,95 por hora. Los cargos de nivel inicial comienzan con un ingreso de € 19.500 al año, mientras que profesionales más experimentados perciben hasta € 38.849 al año.

Talent.com
<https://es.talent.com/salary/jobanalista+de+datos>



Sueldos para Analista en Caracas, Venezuela

El sueldo promedio de un Analista es de Bs. F. 3.050.000 en Caracas.

14 of 22



glassdoor.com.au

<https://www.classroom.com.ar> > Sustentos > carreco-avistel

Sueldo: Analista en Caracas, Venezuela 2023 - Glassdoor

3. Cada vez más empresas están adoptando tecnologías de análisis de datos.

La nueva tecnología hace que la vida sea mucho más fácil para realizar tareas complejas y que requieren mucho tiempo.

Las empresas más grandes pueden tener cantidades diversas y francamente asombrosas de conjuntos de datos para clasificar.

La razón por la que estas mismas empresas buscan adoptar tecnologías que respalden sus estrategias es la siguiente: la implementación de herramientas de análisis les brinda una ventaja competitiva.



El mercado mundial de inteligencia empresarial valió **24 mil millones de dólares** en 2021.



Prácticamente todos los sectores utilizan análisis a cierta escala, desde la automoción hasta la moda, pasando por el comercio minorista y el comercio electrónico; no faltan opciones para elegir en qué industria le gustaría trabajar.

Estas estadísticas de DataProt reflejan esto:



La tasa global de adopción de inteligencia de negocios **es del 26%**.



El **52%** de las empresas de software y el **50%** de las financieras utilizan herramientas de BI.



Las organizaciones dejan sin utilizar el **97%** de los datos recopilados.

Muchas oportunidades de crecimiento.

Debido a la gran cantidad de industrias que adoptan la gestión de datos, siempre tendrá opciones para el **movimiento lateral y ascendente**.



D. DIFERENCIAS ENTRE ESTA CARRERA Y OTRAS DEL MUNDO TECH

Analista de Datos

Fusión de conocimientos en datos y tecnología para informar decisiones empresariales.

Desarrollador Web

Creación de aplicaciones y sitios web funcionales utilizando tecnologías web.

Desarrollador Full Stack

UX Designer: Experiencia del usuario y flujo.
UI Designer: Diseño visual e interacción de la interfaz de usuario.

Diseñador UX/UI

Desarrollo completo de aplicaciones y coordinación. Habilidades tanto en desarrollo frontend como backend.

Gerente de Proyectos

Supervisión, gestión y coordinación de proyectos tecnológicos.

Arquitecto de Soluciones

Diseño de arquitecturas tecnológicas a gran escala. Diseño de sistemas, integración tecnológica y visión estratégica.

Especialista en Cloud Computing

Implementación y gestión de soluciones en la nube. Conocimientos en servicios en la nube y seguridad.

Desarrollador de Aplicaciones Móviles

Creación y diseño de aplicaciones para plataformas móviles.





El análisis de datos es una carrera muy ligada a los **conocimientos de dominio** (Área funcional y negocio o industria) y los **conocimientos en tecnología**.

Si sos una persona que le gusta la parte de negocio de **análisis** y además la **tecnología** seguro el análisis de datos es para vos. Piensen en el analista como un **consultor de negocios** pero con tecnología.

El resto de las carreras digitales como: Desarrollador, UX o UI designer, Full stack no ofrecen el lado de negocio que ofrece el analista. No son una especie de consultor de negocio, sino que las tareas son más de **programar o ejecutar proyectos** mas cerrados. El analista tiene que ser un rol mas estratégico, un aliado que ayuda a todos los departamentos a que crezcan mediante la información.

A continuación, te proporciono una breve comparación de principales roles:

Arquitecto de Soluciones (Solution Architect):

- Enfoque:** Diseña arquitecturas de software y sistemas tecnológicos a gran escala para satisfacer las necesidades empresariales.
- Habilidades clave:** Diseño de sistemas, integración de tecnología, gestión de proyectos y visión estratégica.
- Resultados:** Define la arquitectura tecnológica para proyectos y soluciones empresariales.

Analista de Datos:

- Enfoque:** Se centra en la recopilación, limpieza, análisis y visualización de datos para tomar decisiones empresariales basadas en datos.
- Habilidades clave:** Conocimientos básicos en estadísticas, herramientas de análisis de datos, programación (por ejemplo, Python o R), habilidades de comunicación y habilidades empresariales.
- Toma de decisiones:** Ayuda a la toma de decisiones empresariales basadas en insights derivados de datos.

Desarrollador Web:

- Enfoque:** Diseña y crea aplicaciones web y sitios web utilizando tecnologías web como HTML, CSS, JavaScript y marcos de trabajo (frameworks).
- Habilidades clave:** Conocimientos en lenguajes y tecnologías web, resolución de problemas, programación, diseño web y habilidades de colaboración.
- Toma de decisiones:** Desarrolla sitios y aplicaciones web interactivos y funcionales.

Desarrollador Full Stack:

- Enfoque:** Posee habilidades tanto de desarrollo frontend (interfaz de usuario) como de desarrollo backend (servidor y base de datos).
- Habilidades clave:** Conocimientos en lenguajes y tecnologías frontend y backend, resolución de problemas, desarrollo de

Cloud Computing:

- Enfoque:** Diseña, implementa y administra soluciones en la nube utilizando servicios de proveedores como AWS, Azure o Google Cloud.
- Habilidades clave:** Conocimientos en servicios en la nube, escalabilidad, seguridad en la nube y gestión de recursos en la nube.
- Resultados:** Optimiza la infraestructura tecnológica utilizando

Desarrollador de Aplicaciones Móviles:

- Enfoque:** Crea aplicaciones móviles para plataformas como iOS y Android, utilizando lenguajes y marcos específicos.
- Habilidades clave:** Desarrollo móvil, diseño de aplicaciones, pruebas en dispositivos móviles y publicación en tiendas de aplicaciones.
- Resultados:** Desarrolla aplicaciones móviles que ofrecen funcionalidades específicas para dispositivos móviles.

Gerente de Proyectos (Project Manager):

- Enfoque:** Supervisa proyectos de tecnología, coordina equipos, gestiona recursos y garantiza que los proyectos se completen a tiempo y dentro del presupuesto.
- Habilidades clave:** Habilidades de liderazgo, gestión de proyectos, comunicación, toma de decisiones y habilidades empresariales.
- Resultados:** Asegura la entrega exitosa de proyectos tecnológicos.

UI/UX Designer:

- Enfoque:** - UX Designer: Se concentra en la experiencia del usuario, la investigación, la usabilidad y el flujo del usuario.
- UI Designer: Se enfoca en el diseño visual, la apariencia y la interacción de la interfaz de usuario.
- Habilidades clave:** Conocimientos en diseño de experiencia de usuario y diseño de interfaz de usuario, herramientas de diseño, empatía con el usuario y habilidades de creatividad.
- Resultados:** Crea interfaces y experiencias de usuario atractivas y funcionales.

Cada uno de estos roles tiene su **propio enfoque y conjunto de habilidades**, y contribuye de manera única a la industria tecnológica. La elección de carrera dependerá de tus intereses, habilidades y objetivos profesionales. Además, es importante destacar que en el campo tecnológico, la **colaboración** y la **interacción** entre estos roles a menudo son necesarias para el desarrollo y la implementación exitosa de proyectos.



E. BENEFICIOS DE SER ANALISTA DE DATOS

Ser analista de datos no solo te da la satisfacción a nivel profesional porque las tareas que realizas son retadoras, divertidas y aportan valor. Sino que tienen una serie de beneficios psicológicos.

02

Libertad geográfica

Vive donde quieras con acceso a internet.



01

Trabajo Remoto

Aprovecha la tendencia actual del trabajo desde casa.

03

Mejor salario

Mayor demanda significa mejores salarios.

04

Horarios flexibles

Adapta tu jornada y equilibra vida y trabajo.



05

Red Global

Conéctate profesionalmente en todo el mundo.



07

Independencia

Trabaja como freelance para mayor control.



06

Crecimiento

Aprende y crece constantemente.



08

Sector Versátil

Encuentra oportunidades en diversas industrias.

> TÉRMINOS DEL MUNDO DE LOS DATOS 

PARTE III

A. BUSINESS INTELLIGENCE

La **inteligencia comercial (BI)** es un proceso impulsado por la tecnología para analizar datos y brindar información procesable que ayuda a los ejecutivos, gerentes y trabajadores a tomar decisiones de negocio informadas.

Como parte del proceso de BI, las organizaciones recopilan datos de los sistemas de TI internos y fuentes externas, los preparan para el análisis, ejecutan consultas contra los datos y crean visualizaciones de datos, tableros e informes de BI para que los resultados analíticos estén disponibles para los usuarios comerciales para la toma de decisiones operativas. Elaboración y planificación estratégica.

🎯 El **objetivo final** de las iniciativas de BI es:

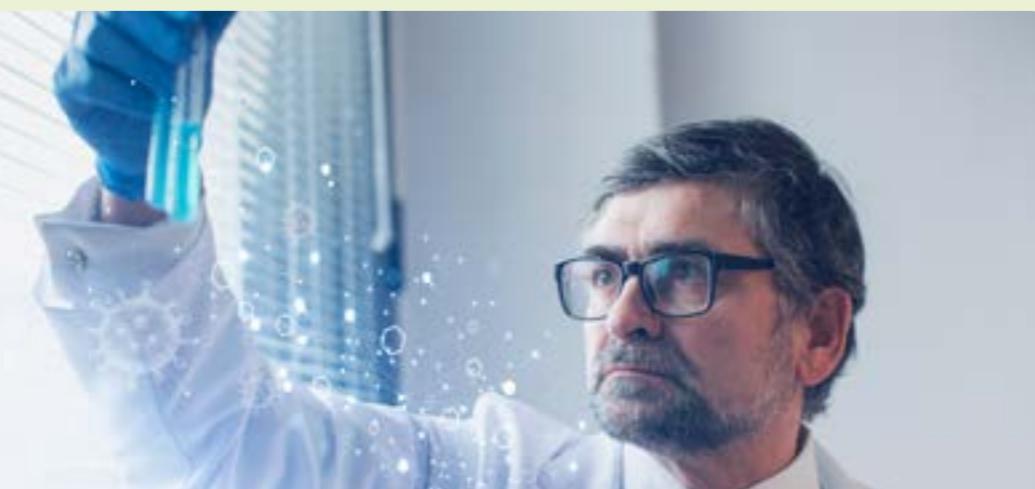
1. Impulsar mejores decisiones comerciales que permitan a las organizaciones aumentar los ingresos.
2. Mejorar la eficiencia operativa y obtener ventajas competitivas sobre los rivales comerciales.

Para lograr ese objetivo, BI incorpora una combinación de herramientas de análisis, gestión de datos e informes, además de varias metodologías para gestionar y analizar datos.

DATA SCIENCE O CIENCIA DE DATOS

Data Science es un término general para información a partir de datos. Data Science es un conjunto de métodos y prácticas para **recopilar conocimientos** (información, aprendizajes, etc.) a partir de datos. Los datos pueden ser cualquier cosa (precios de acciones, grabaciones de voz, datos de sensores de pluviómetros, imágenes de satélite, etc.). La ciencia de datos puede incluir el **procesamiento de los datos**, la realización de **análisis estadísticos de los datos**, la **presentación de los datos** de manera que otros puedan entenderlos (lo que se denomina narración de datos), etc. A veces, estos análisis son simples (como la precipitación promedio). A veces son mucho más complejos. Pero todo es ciencia de datos.

- En términos simples, la ciencia de datos es el **proceso de extraer información útil de datos no estructurados y estructurados**.
- Es un **enfoque interdisciplinario** que combina varios campos de la informática, la estadística y los procesos y métodos científicos para sacar conclusiones a partir de puntos de datos sin procesar.



ANÁLISIS DE DATOS

Aunque muchos grupos, organizaciones y expertos tienen diferentes formas de abordar el análisis de datos, la mayoría de ellas se pueden resumir en una única definición.

 El análisis de datos es el proceso de **limpieza, cambio y procesamiento de datos sin procesar**, y la extracción de información relevante y procesable que ayuda a las empresas a tomar **decisiones informadas**.

 El procedimiento ayuda a reducir los riesgos inherentes a la toma de decisiones al proporcionar información y estadísticas útiles, que a menudo se presentan en **cuadros, imágenes, tablas y gráficos**.

 **Un ejemplo** sencillo de análisis de datos lo podemos ver siempre que tomamos una decisión en nuestro día a día evaluando lo que ha pasado en el pasado o lo que pasará si tomamos esa decisión. Básicamente, este es el proceso de analizar el pasado o el futuro y tomar una decisión basada en ese análisis.



INGENIERO DE DATOS

La **ingeniería de datos** es un campo de la informática dedicado al **diseño y desarrollo de sistemas** para **adquirir, transformar, almacenar y gestionar** grandes conjuntos de datos.

Los ingenieros de datos crean procesos **ETL** para limpiar y estructurar datos provenientes de diversas fuentes, implementan sistemas de almacenamiento eficientes, diseñan procesos de procesamiento de datos y aseguran la seguridad y privacidad de la información.

Su objetivo es:

01

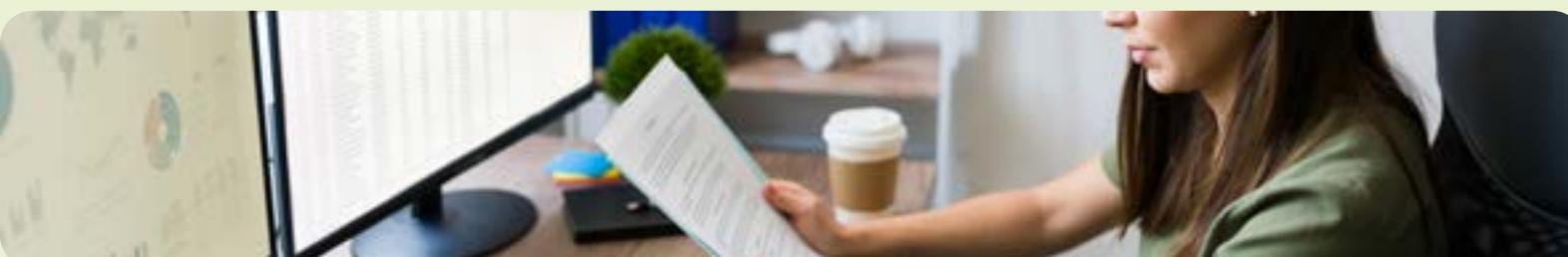
Proporcionar una **infraestructura robusta**

02

Permitirle a las organizaciones obtener **conocimientos valiosos** a partir de sus datos

03

Facilitar la toma de decisiones informadas y la optimización de operaciones



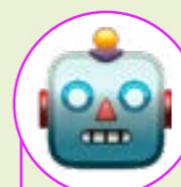
INTELIGENCIA ARTIFICIAL

La **inteligencia artificial (IA)** es un término general para cualquier tecnología en la que un **programa de computadora** intenta tareas que son naturales para el cerebro humano.

Habilidades como comprender el lenguaje escrito, detectar el habla, reconocer objetos a partir de imágenes y hacer planes para optimizar el tiempo, son ejemplos de inteligencia que los humanos muestran todos los días.



- **Nuestros cerebros** aprenden la mayoría de forma natural a medida que crecemos e interactuamos con el mundo, y luego se refinan y avanzan mediante el aprendizaje formal.



- La inteligencia artificial es la habilidad de **una máquina** de presentar las mismas capacidades que los seres humanos:
 - El razonamiento
 - El aprendizaje
 - La creatividad
 - La capacidad de planear



MACHINE LEARNING

El aprendizaje automático implica que una computadora reconozca patrones mediante ejemplos, en lugar de programarla con reglas específicas. Estos patrones se encuentran dentro de Data. Machine Learning es una gran rama de la inteligencia artificial.

 **Máquina** = Su máquina o computadora

 **Aprendizaje** = Encontrar patrones en los datos

El aprendizaje automático se trata de: crear algoritmos (un conjunto de reglas) que aprenden de funciones complejas (patrones) de datos para hacer predicciones sobre ellos.

En resumen, las máquinas pueden predecir el futuro, siempre que el futuro no se vea muy diferente del pasado.

Esencialmente, se puede resumir en 3 pasos:



Aplicaciones del aprendizaje automático

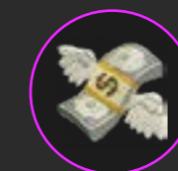
Antes de comenzar, aquí hay una descripción general rápida de lo que es capaz de hacer el aprendizaje automático:



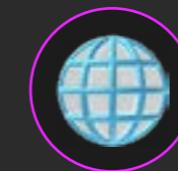
Atención médica: Predicción de diagnósticos de pacientes para que los médicos los revisen.



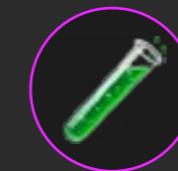
Red social: Predecir ciertas preferencias de coincidencia en un sitio web de citas para una mejor compatibilidad.



Finanzas: Predicción de actividad fraudulenta en una tarjeta de crédito.



Comercio electrónico: Predicción de la rotación de clientes.



Biología: Encontrar patrones en mutaciones genéticas que podrían representar cáncer.

¿Cómo aprenden las máquinas?

Para simplificar las cosas, sepa que las máquinas “aprenden” encontrando patrones en datos similares. Piense en los datos como información que adquiere del mundo. Cuantos más datos se le dan a una máquina, más “inteligente” se vuelve.

Pero no todos los datos son iguales. Imagina que eres un pirata y tu misión en la vida es encontrar el tesoro enterrado en algún lugar de la isla. Para encontrar el tesoro, necesitará suficiente **información**. Al igual que los datos, esta información puede llevarlo en la dirección correcta o en la dirección incorrecta.

Cuanto mejor sea la información/datos que se obtengan, más incertidumbre se reducirá, y viceversa. Por lo tanto, es importante tener en cuenta el **tipo de datos** que le está dando a **su máquina** para que los aprenda.



No obstante, después de proporcionar una cantidad suficiente de datos, la máquina puede hacer **predicciones**. **Las máquinas** pueden predecir el futuro, siempre que el futuro no se vea muy diferente del pasado.

La máquina realmente “aprende” usando datos antiguos para obtener información sobre cuál es la mayor probabilidad de que suceda.

Si los datos antiguos se parecen mucho a los datos nuevos, entonces las cosas que puede decir sobre los datos antiguos probablemente serán relevantes para los datos nuevos. Es como mirar hacia atrás para mirar hacia adelante.

Tipos de aprendizaje automático

Hay tres categorías principales de aprendizaje automático:



- **Aprendizaje supervisado**

La máquina aprende de los datos etiquetados. Normalmente, los datos son etiquetados por humanos.

- **Aprendizaje no supervisado**

La máquina aprende de datos no etiquetados. Es decir, no se le da una respuesta “correcta” a la máquina para que aprenda, pero es de esperar que la máquina encuentre patrones a partir de los datos para encontrar una respuesta.

- **Aprendizaje por refuerzo**

La máquina aprende a través de un sistema basado en recompensas.

DEEP LEARNING

El deep learning es un tipo de machine learning que **entrena a una computadora** para que realice tareas como las hacemos los seres humanos, como el reconocimiento del habla, la identificación de imágenes o hacer predicciones.



Está basado en redes neuronales artificiales en las que se utilizan múltiples capas de procesamiento para extraer características de nivel progresivamente más alto de los datos.



Machine learning vs Deep Learning

- El aprendizaje profundo se distingue del aprendizaje automático clásico por el **tipo de datos** con los que trabaja y los **métodos** con los que aprende.
 - Los algoritmos de aprendizaje automático aprovechan los datos estructurados y etiquetados para hacer **predicciones**, lo que significa que las características específicas se definen a partir de los datos de entrada para el modelo y se organizan en tablas. Esto no significa necesariamente que no utilice datos no estructurados; solo significa que, si lo hace, generalmente pasa por un **procesamiento** previo para organizarlo en un **formato estructurado**.
- El aprendizaje profundo elimina parte del pre-procesamiento de datos que suele estar relacionado con el aprendizaje automático. Estos algoritmos pueden **ingerir y procesar datos** no estructurados, como texto e imágenes, y automatizan la extracción de funciones, eliminando parte de la dependencia de expertos humanos. Por ejemplo, supongamos que tenemos un conjunto de fotos de diferentes mascotas y queremos categorizarlas por “gato”, “perro”, “hámster”, etcétera. Los algoritmos de aprendizaje profundo pueden determinar qué características (por ejemplo, las orejas) son las más importantes para distinguir a un animal de otro. En el aprendizaje automático, esta **jerarquía de funciones** la establece manualmente un experto humano.

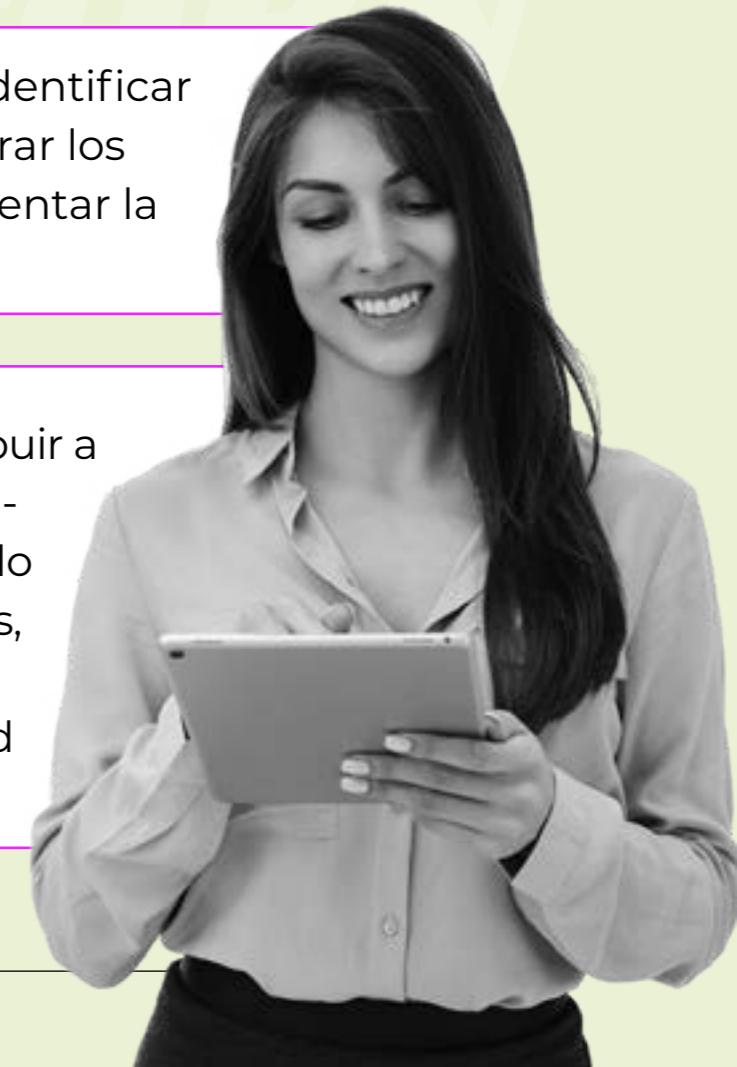
> ROLES EN EL MUNDO DE LOS DATOS 

PARTE IV

A. EL ANALISTA DE DATOS

🎯 Objetivos del Analista de Datos:

- 1 Obtener Perspicacia:** El objetivo principal del analista de datos es obtener perspicacia empresarial a partir de los datos. Esto implica identificar patrones, tendencias y relaciones en los datos que puedan beneficiar a la organización.
- 2 Apoyo en la Toma de Decisiones:** Ayudar a la empresa a tomar decisiones informadas basadas en evidencia y datos sólidos, en lugar de suposiciones o intuición.
- 3 Optimización de Procesos:** Identificar áreas donde se pueden mejorar los procesos empresariales, aumentar la eficiencia y reducir costos.
- 4 Mejora de Resultados:** Contribuir a la mejora de los resultados empresariales, ya sea aumentando los ingresos, reduciendo costos, optimizando la retención de clientes o mejorando la calidad de productos y servicios.



Tareas del Analista de Datos:



RECOPILACIÓN DE DATOS

Adquirir datos de diversas fuentes, como bases de datos, archivos, sensores, encuestas y redes sociales.



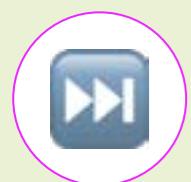
LIMPIEZA Y PREPARACIÓN DE DATOS

Depurar y organizar los datos para eliminar errores, valores atípicos y redundancias.



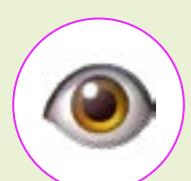
ANÁLISIS EXPLORATORIO DE DATOS

Aplicar técnicas exploratorias para identificar patrones, correlaciones y tendencias en los datos.



MODELADO Y PREDICIONES

Crear modelos y algoritmos que permitan predecir resultados futuros o tomar decisiones basadas en datos históricos.



VISUALIZACIÓN DE DATOS

Utilizar herramientas de visualización para representar gráficamente los resultados del análisis y hacerlos comprensibles para los no expertos.



COMUNICACIÓN DE RESULTADOS

Presentar los hallazgos y las recomendaciones a los equipos de liderazgo y otros interesados de manera clara y efectiva.

Entregables y Proyectos:

01

Informes de análisis:

Resumen los resultados de sus análisis y **proporcionan recomendaciones**.

02

Cuadros de mando y paneles:

Permiten a las partes interesadas **explorar datos y métricas clave** por sí mismas.

03

Modelos Predictivos:

Predicen **resultados futuros**, como el comportamiento del cliente o las tendencias de ventas.

04

Proyectos de Optimización:

Optimizan procesos, como la gestión de inventario y la segmentación de clientes.

Resultados:



Mejora en la Toma de Decisiones: El éxito del analista de datos se mide por la mejora en la calidad de las decisiones que toma la empresa basadas en los datos proporcionados.



Ahorro de Costos: Si los análisis conducen a la identificación de áreas donde se pueden reducir costos, esto se considera un resultado positivo.



Incremento en los Ingresos: Si las acciones basadas en análisis conducen a un aumento en los ingresos, se considera un logro importante.



Mejora de la Eficiencia: El analista de datos también puede medir el éxito en función de la eficiencia y la optimización de los procesos empresariales.

En resumen, el analista de datos busca obtener información valiosa a partir de datos, lo que se traduce en mejores decisiones empresariales y resultados positivos para la organización. Su trabajo involucra una serie de tareas, proyectos y entregables, y se mide en función de su impacto en la toma de decisiones y los resultados empresariales.

B. EL ANALISTA DE BUSINESS INTELLIGENCE

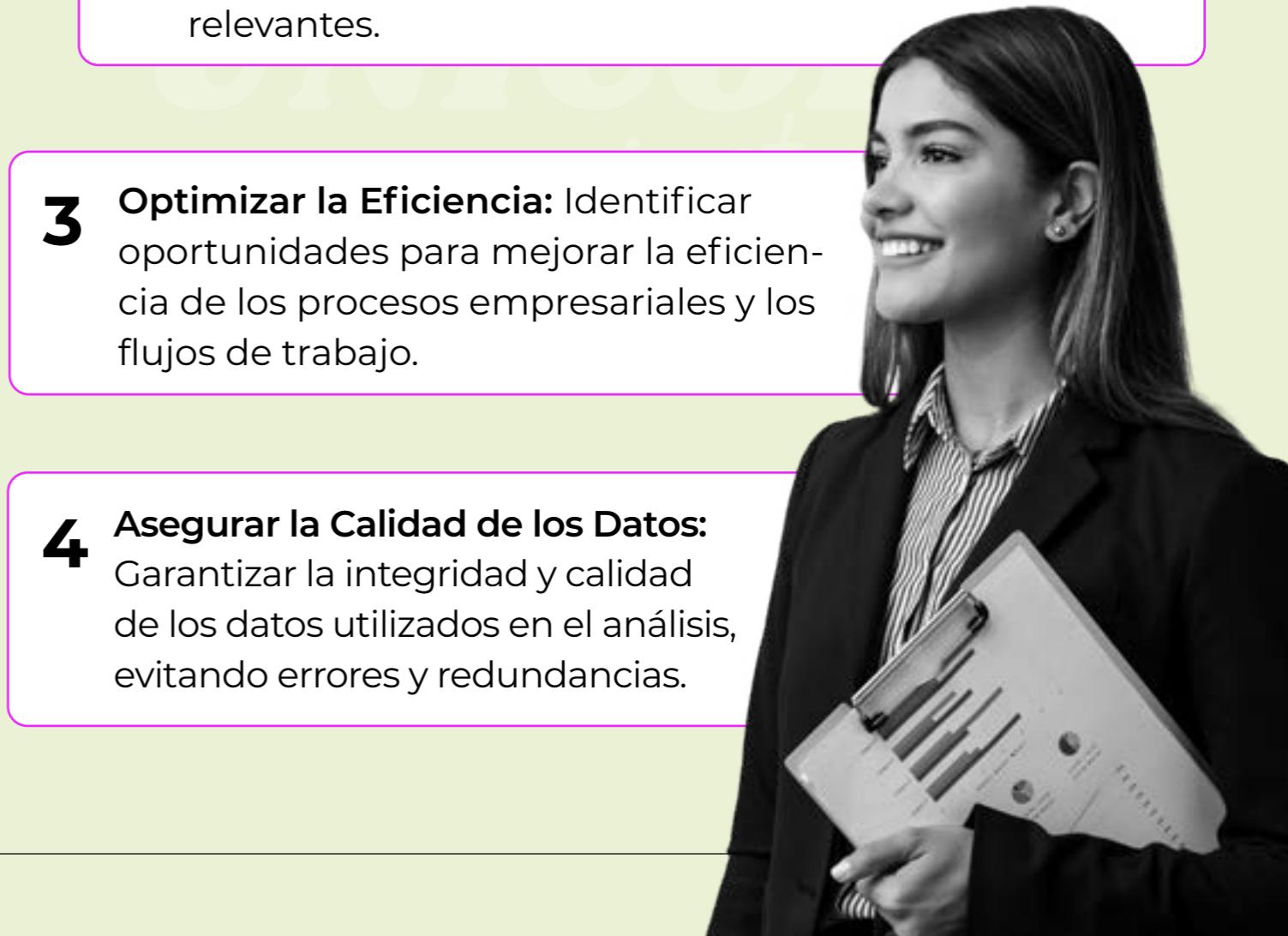
🎯 Objetivos del Profesional de Business Intelligence:

1 Entender las Necesidades Empresariales: El objetivo principal de un profesional de BI es comprender las necesidades y objetivos de la organización en términos de análisis y toma de decisiones basadas en datos.

2 Mejorar la Toma de Decisiones: Facilitar la toma de decisiones más informadas y estratégicas en la empresa, proporcionando datos y análisis oportunos y relevantes.

3 Optimizar la Eficiencia: Identificar oportunidades para mejorar la eficiencia de los procesos empresariales y los flujos de trabajo.

4 Asegurar la Calidad de los Datos: Garantizar la integridad y calidad de los datos utilizados en el análisis, evitando errores y redundancias.



Tareas del Profesional de Business Intelligence:



RECOPILACIÓN E INTEGRACIÓN DE DATOS:

Adquirir y consolidar datos de diversas fuentes, como bases de datos, sistemas empresariales, hojas de cálculo y aplicaciones.



MODELADO Y DISEÑO DE ALMACÉN DE DATOS:

Crear modelos de datos que permitan la estructuración adecuada de la información y el diseño de almacenes de datos eficientes.



DESARROLLO DE CUADROS DE MANDO:

Diseñar cuadros de mando y paneles interactivos que presenten datos clave y métricas de rendimiento.



GENERACIÓN DE INFORMES Y ANÁLISIS:

Crear informes y análisis que proporcionen información relevante y útil para los tomadores de decisiones.



GESTIÓN DE PROYECTOS DE BUSINESS INTELLIGENCE:

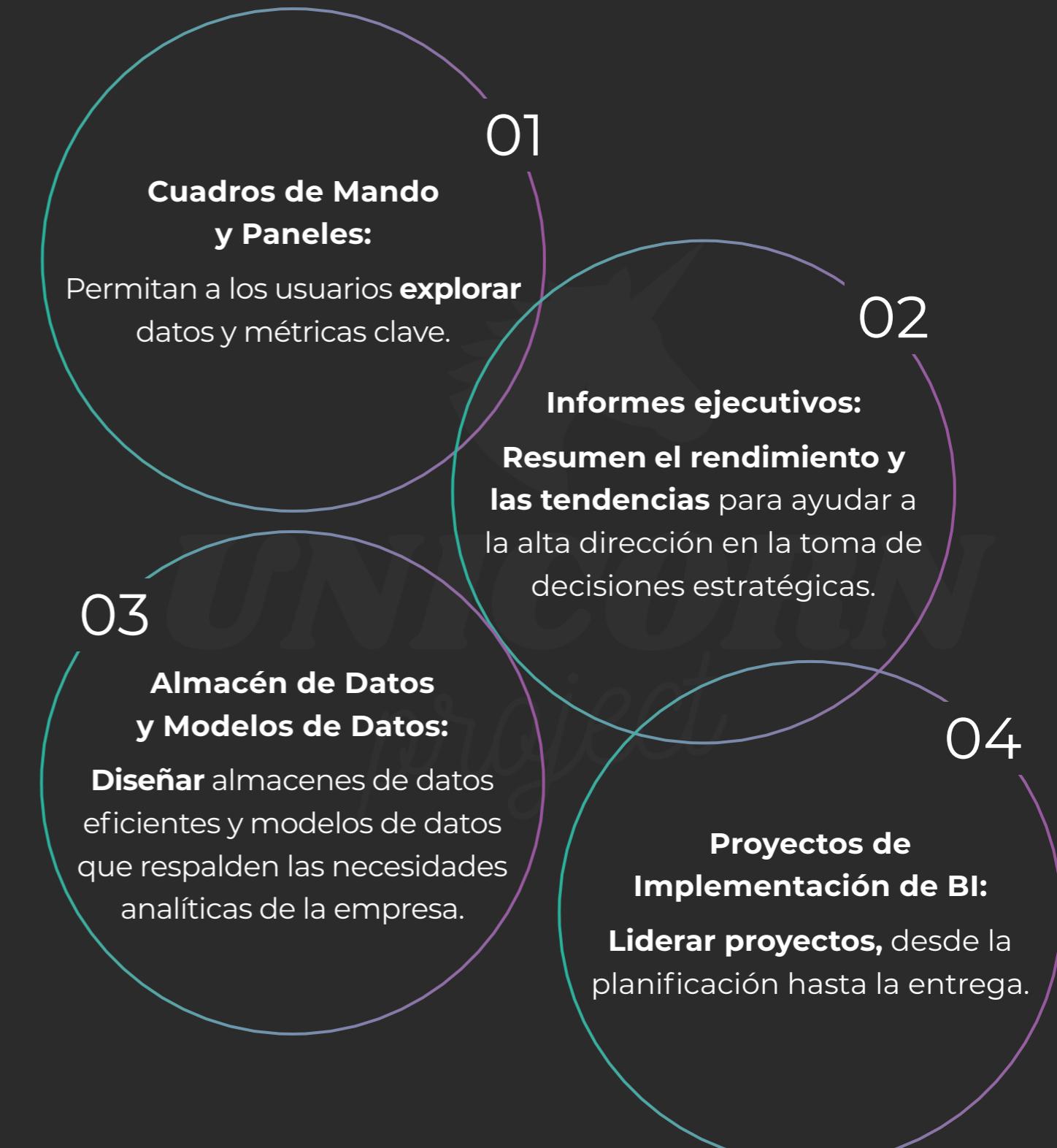
Liderar proyectos relacionados con la implementación de soluciones de BI y garantizar su éxito.



COLABORACIÓN CON LOS EQUIPOS DE NEGOCIOS:

Trabajar en estrecha colaboración con los equipos de negocio para comprender sus necesidades y traducirlas en soluciones de BI.

Entregables y Proyectos:



Resultados:



Mejora en la Toma de Decisiones: El éxito de un profesional de BI se mide por la mejora en la calidad de las decisiones empresariales basadas en los datos y análisis proporcionados.



Eficiencia Operativa: Si las soluciones de BI contribuyen a la optimización de los procesos y flujos de trabajo, se considera un resultado positivo.



Ahorro de Tiempo y Recursos: La capacidad de proporcionar datos y análisis de manera más eficiente puede resultar en ahorros significativos.



Mejora de la Experiencia del Cliente: Si las iniciativas de BI ayudan a comprender y satisfacer mejor las necesidades de los clientes, se considera un éxito.

En resumen, un profesional de Business Intelligence se centra en proporcionar datos y análisis relevantes para mejorar la toma de decisiones y optimizar los procesos empresariales. Su trabajo implica una variedad de tareas, proyectos y entregables, y se mide en función de su impacto en la calidad de las decisiones, la eficiencia operativa y otros resultados empresariales clave.

C. EL CIENTÍFICO DE DATOS



Comprender los Objetivos Empresariales:

El objetivo primordial de un científico de datos es comprender los objetivos y desafíos de la organización para aplicar análisis avanzados y soluciones basadas en datos.

1

Generar Ideas Estratégicas: Facilitar la generación de ideas estratégicas mediante el análisis de datos complejos, contribuyendo así a la innovación y la ventaja competitiva.

2

Desarrollar Modelos Predictivos: Construir y optimizar modelos predictivos y algoritmos para anticipar tendencias y patrones futuros.

3

Optimizar Procesos de Toma de Decisiones: Mejorar los procesos de toma de decisiones a través de la aplicación de análisis de datos avanzados y técnicas de machine learning.



Tareas del profesional de Ciencia de Datos:



EXPLORACIÓN Y PRE PROCESAMIENTO DE DATOS:

Explorar y pre procesar datos para identificar patrones y prepararlos para el análisis avanzado.



DESARROLLO DE MODELOS PREDICTIVOS:

Utilizar algoritmos de machine learning para construir modelos predictivos y clasificatorios.



ANÁLISIS ESTADÍSTICO AVANZADO:

Aplicar técnicas estadísticas avanzadas para extraer conocimientos significativos de conjuntos de datos complejos.



VISUALIZACIÓN DE DATOS:

Crear visualizaciones efectivas para comunicar hallazgos y patrones de manera comprensible.



OPTIMIZACIÓN DE MODELOS:

Mejorar continuamente la precisión y eficiencia de los modelos a través de la optimización y ajuste.



COLABORACIÓN CON EQUIPOS INTERDISCIPLINARIOS:

Trabajar en estrecha colaboración con equipos interdisciplinarios, como ingenieros de datos y analistas de negocio.

Entregables y Proyectos:

01 Modelos Predictivos Desarrollados:

Que ayuden a anticipar eventos y tomar decisiones informadas.

02

Visualizaciones Persuasivas:

Que faciliten la interpretación de resultados y la toma de decisiones estratégicas.

03

Aplicaciones de Machine Learning:

Para integrar soluciones en los procesos empresariales.

04

Informes de Análisis Avanzado:

Presentar informes de análisis avanzados y proporcionar recomendaciones basadas en resultados.

Resultados:

Mejora en la Precisión Predictiva: El éxito de un científico de datos se mide por la mejora en la precisión de los modelos predictivos y su capacidad para ofrecer insights únicos



Innovación en la Toma de Decisiones: Contribuir a la innovación empresarial al proporcionar información valiosa que impulsa decisiones estratégicas.



Impacto Tangible en la Organización: Lograr un impacto tangible en la eficiencia operativa, la satisfacción del cliente y otros indicadores clave de rendimiento.



Colaboración Efectiva con Equipos Interdisciplinarios: Una colaboración efectiva con otros profesionales para integrar soluciones de datos en toda la organización.

**D. EL INGENIERO DE DATOS****Objetivos del Profesional de Ingeniería de Datos:**

1 Comprender las Necesidades de Datos: El objetivo principal de un ingeniero de datos es comprender las necesidades de datos de la organización y diseñar soluciones para su adquisición, procesamiento y almacenamiento.

2 Garantizar la Disponibilidad de Datos: Asegurar que los datos estén disponibles y accesibles para análisis y toma de decisiones en toda la organización.

3 Optimizar la Eficiencia del Procesamiento: Mejorar la eficiencia de los procesos de manejo de datos, desde la extracción hasta la carga (ETL) y más allá.

4 Implementar Soluciones Escalables: Desarrollar sistemas escalables que puedan manejar grandes volúmenes de datos y adaptarse al crecimiento de la organización.



Tareas del Profesional de Científico de Datos:



DISEÑO DE ARQUITECTURA DE DATOS:

Crear y mantener arquitecturas de sistemas para el manejo eficiente de grandes conjuntos de datos.



DESARROLLO DE PROCESOS ETL:

Diseñar y optimizar procesos de Extracción, Transformación y Carga para la transferencia eficiente de datos entre sistemas.



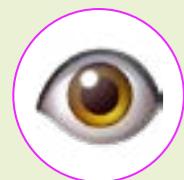
IMPLEMENTACIÓN DE SISTEMAS DE ALMACENAMIENTO DE DATOS:

Seleccionar e implementar sistemas de almacenamiento de datos, como bases de datos relacionales, NoSQL y soluciones en la nube.



GARANTIZAR LA SEGURIDAD DE LOS DATOS:

Implementar medidas de seguridad para proteger la integridad y confidencialidad de los datos.



MONITOREO Y MANTENIMIENTO:

Supervisar el rendimiento de los sistemas y realizar mantenimiento preventivo para garantizar la disponibilidad continua de los datos.



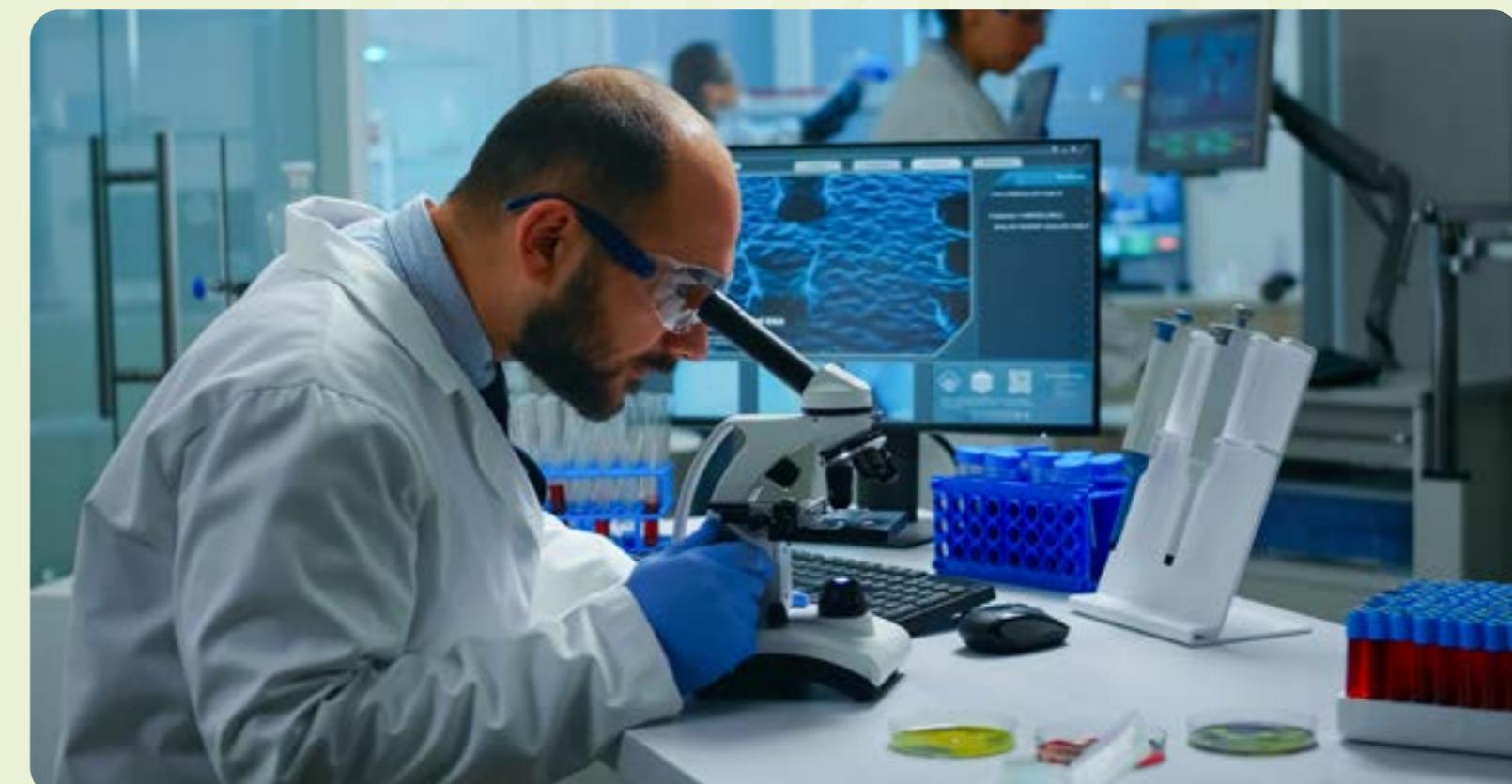
CADOPCIÓN DE TECNOLOGÍAS EMERGENTES:

Mantenerse al tanto de las últimas tecnologías en gestión de datos, como herramientas de big data y soluciones en la nube, e integrarlas de manera apropiada.



COLABORACIÓN CON EQUIPOS DE DATOS Y ANALISTAS:

Trabajar en estrecha colaboración con científicos de datos, analistas y otros profesionales para comprender y abordar las necesidades de datos de la organización.



Entregables y Proyectos:



Resultados:



Disponibilidad y Accesibilidad de Datos Mejoradas:
El éxito de un ingeniero de datos se mide por la mejora en la disponibilidad y accesibilidad de datos para la organización.



Eficiencia en el Procesamiento de Datos: Lograr una mayor eficiencia en los procesos de manejo de datos, reduciendo tiempos de carga y optimizando pipelines.



Adaptación Exitosa a Cambios en la Escala: La capacidad para diseñar sistemas escalables que se adapten al crecimiento de la organización.



Colaboración Efectiva con Equipos de Datos: Una colaboración efectiva que asegure que los sistemas de datos satisfagan las necesidades de analistas y científicos de datos.



> CONOCIMIENTOS TÉCNICOS

PARTE V

project

A. ALMACENAMIENTO DE DATOS

Las bases de datos tienen como principal objetivo almacenar datos. Un almacenamiento de datos es un **repositorio central de información** que se puede **analizar** para tomar decisiones mejor informadas. Los datos fluyen hacia un almacenamiento de datos desde sistemas transaccionales, bases de datos relacionales y otros orígenes.

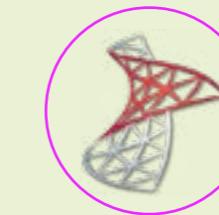
El almacenamiento de datos puede contener varias bases de datos. Dentro de cada base de datos, los datos se organizan en **tablas y en columnas**. Dentro de cada columna, se puede definir una descripción de los datos, como un número entero, un campo de datos o una cadena. Las tablas se pueden organizar dentro de esquemas, que se pueden concebir como carpetas. Cuando los datos se incorporan, se almacenan en varias tablas descritas por el esquema. Las herramientas de consulta usan el esquema para determinar a qué tablas de datos obtendrán acceso y cuáles de ellas analizarán.

Es importante tener en cuenta que la base de datos se encuentra almacenada en un servidor que la empresa puede tener propio o puedes contratar a un tercero con tecnología en la nube por ejemplo.

Tenemos distintos gestores de bases de datos que son los software que usamos como analistas para acceder a la información:



MySQL



SQL SERVER



POSTGRESQL



ORACLE

Otro tema importante a tener en cuenta es que SQL es un LENGUAJE. No es la base de datos en sí.

Los software que nos permiten acceder al servidor que tiene la base de datos (los anteriores que mencionamos) tienen el nombre de SQL incorporado, pero en sí esas siglas significan “structured query language” .

¿Por qué necesitas saber SQL?

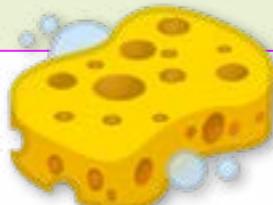
SQL es un lenguaje que cualquier rol en datos debería aprender.



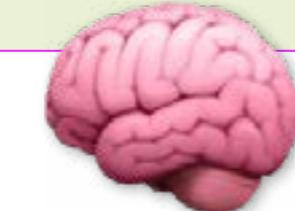
Te permite acceder a las bases de datos y **consultar** información.



Te permite **manipular** información.



Te permite **limpiar** datos.



Es de los lenguajes más **fáciles de aprender**.

B. ETL | EXTRACCIÓN, TRANSFORMACIÓN Y CARGA

El proceso ETL (Extracción, Transformación y Carga) es fundamental en el análisis de datos, ya que facilita la **preparación y organización de los datos** para su posterior análisis. **Aquí te presento cómo se utiliza el ETL en el mundo del análisis de datos:**

EXTRACCIÓN DE DATOS:

- **Descripción:** La fase de extracción implica la recopilación de datos desde diversas fuentes, como bases de datos, archivos planos, API web y sistemas empresariales. Los datos pueden provenir de fuentes internas o externas a la organización.
- **Uso en el Análisis de Datos:** Permite consolidar información dispersa en una ubicación centralizada, creando un punto de acceso único para su análisis. Los analistas pueden extraer datos relevantes para sus investigaciones sin tener que acceder a múltiples sistemas.

TRANSFORMACIÓN DE DATOS:

- **Descripción:** La transformación implica la limpieza, estructuración y modificación de los datos para que se ajusten a los requisitos del análisis. Esto puede incluir la corrección de errores, la normalización de formatos y la creación de nuevas variables.
- **Uso en el Análisis de Datos:** Prepara los datos para el análisis al asegurar que estén en un formato consistente y listos para ser utilizados por herramientas de análisis. Los analistas pueden transformarlos para adaptar los datos a sus necesidades.



CARGA DE DATOS:

- **Descripción:** La fase de carga implica la inserción de datos transformados en un repositorio destinado al análisis, como una base de datos o un almacén de datos. Puede incluir la actualización de datos existentes o la creación de nuevas tablas.
- **Uso en el Análisis de Datos:** Proporciona a los analistas un entorno de trabajo estructurado y optimizado para realizar análisis. Los datos están listos para su consulta y pueden ser utilizados eficientemente en diversas herramientas de análisis y visualización.

LIMPIEZA Y VALIDACIÓN:

- **Descripción:** Durante el proceso ETL, se implementan rutinas para limpiar y validar los datos. Esto incluye la identificación y manejo de datos faltantes, la detección y corrección de errores, y la aplicación de reglas de validación.
- **Uso en el Análisis de Datos:** Garantiza la calidad de los datos utilizados en el análisis, mejorando la confiabilidad de los resultados. Los analistas pueden confiar en que los datos están libres de inconsistencias y errores.

INTEGRACIÓN DE DATOS:

- **Descripción:** En entornos empresariales complejos, a menudo implica la integración de datos provenientes de diversas fuentes y sistemas, permitiendo una visión más completa y unificada.
- **Uso en el Análisis de Datos:** Proporciona una visión holística al combinar datos de diferentes áreas de la organización. Los analistas pueden realizar análisis más amplios y comprensivos al tener acceso a datos integrados.

AUTOMATIZACIÓN DE PROCESOS:

- **Descripción:** La automatización en el ETL permite ejecutar procesos de manera programada y regular, asegurando la actualización continua de los datos y la coherencia en el análisis.
- **Uso en el Análisis de Datos:** Asegura que los análisis se realicen con datos actualizados, evitando la dependencia de datos estáticos. Los analistas pueden confiar en la consistencia y actualidad de la información.

En resumen, el proceso ETL en el análisis de datos proporciona la base esencial para que los analistas trabajen con datos limpios, integrados y preparados para el análisis. Facilita la transformación de datos crudos en información valiosa, contribuyendo a la toma de decisiones informadas y al descubrimiento de patrones y tendencias significativas.

Herramientas:



PYTHON



PENTAHO



C. VISUALIZACIÓN DE DATOS

En esta parte es cuando traemos los datos a la vida. Es tomar todos esos datos y análisis que has ido realizando y comunicarnos de forma efectiva.

El cerebro humano no está construido para poder interpretar datos en crudo entonces necesitamos **patrones claros y visualizaciones** para hacer sentido de esta información compleja. Esto es un hecho científico y está atado a cómo funciona nuestro cerebro.

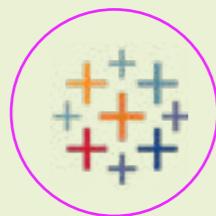
La visualización es un proceso por el cual, a través de diferentes técnicas y herramientas, conseguimos **extraer información** de los datos y **mostrarla de forma eficiente**.

Ante la creciente cantidad de datos, las nuevas y potentes formas de procesarlos y su implementación en los modelos de negocio como base del conocimiento y la estrategia de la empresa, se crea la necesidad de extraer información de los propios datos a través de la **visualización**.

Herramientas:



POWER BI



TABLEAU

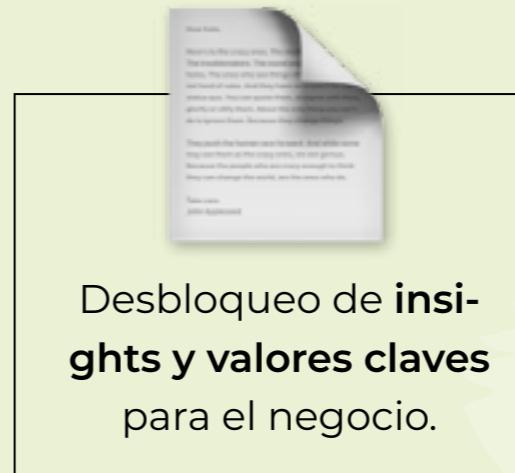


QLIKVIEW

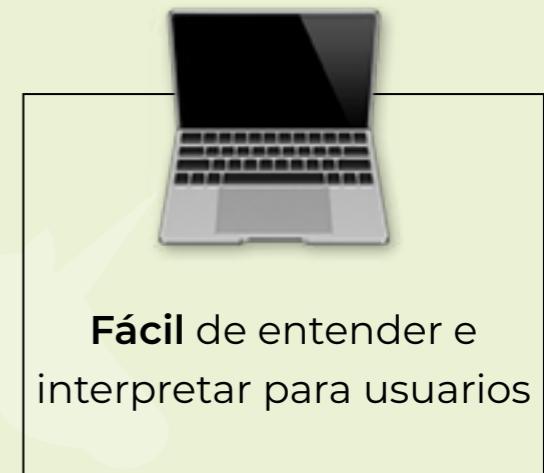


LOOKER

¿Por qué necesitamos tener conocimientos en visualización de datos?



Desbloqueo de **insights y valores claves** para el negocio.



Fácil de entender e interpretar para usuarios



Más atractivo y fácil de enganchar al usuario



Transmitir información y mostrar relaciones complejas de forma **sencilla**.

En resumen, la visualización de datos es esencial en el análisis de datos, ya que permite a los analistas y tomadores de decisiones comprender rápidamente patrones, identificar insights y comunicar resultados de manera efectiva. La elección de visualizaciones depende del tipo de datos y los objetivos específicos del análisis.



D. LENGUAJES DE PROGRAMACIÓN

La programación desempeña un papel esencial en el análisis de datos al proporcionar a los profesionales las herramientas necesarias para **manipular, procesar y extraer información** valiosa a partir de **conjuntos de datos**.

Aquí hay varias formas en las que se utiliza la programación en el análisis de datos:



Manipulación y Limpieza de Datos

- La programación permite a los analistas de datos manipular y limpiar datos de manera eficiente. Esto incluye la identificación y corrección de valores atípicos, manejo de datos faltantes y transformaciones necesarias para preparar los datos para el análisis.



Extracción de Datos

- La programación facilita la extracción de datos desde diversas fuentes, como bases de datos, archivos CSV, hojas de cálculo, API web, entre otros. Los analistas pueden automatizar este proceso para manejar grandes volúmenes de datos de manera rápida y precisa.



Optimización de Procesos

- Los analistas de datos pueden programar procesos para mejorar la eficiencia y velocidad del análisis. Esto es crucial cuando se trabaja con grandes conjuntos de datos y se requiere rendimiento optimizado.



Análisis Exploratorio de Datos (EDA)

- La programación permite realizar análisis exploratorios de datos de manera eficiente. Utilizando bibliotecas específicas, los analistas pueden crear visualizaciones, estadísticas descriptivas y explorar relaciones entre variables para comprender mejor la estructura de los datos.



Pre procesamiento de Datos

- Antes de aplicar modelos predictivos o algoritmos de aprendizaje automático, es común realizar pre procesamiento de datos. Esto incluye la normalización, la codificación de variables categóricas y la ingeniería de características, tareas que se pueden automatizar mediante la programación.



Desarrollo de Modelos de Aprendizaje Automático

- La programación es esencial en la construcción, entrenamiento y evaluación de modelos de aprendizaje automático. Los analistas utilizan bibliotecas especializadas en lenguajes como Python y R para implementar algoritmos y ajustar modelos a los datos.



Automatización de Informes y Presentaciones

- Al programar scripts, los analistas pueden automatizar la generación de informes y presentaciones. Esto facilita la actualización regular de informes con nuevos datos y mejora la eficiencia en la comunicación de resultados a partes interesadas.





Integración de Herramientas y Tecnologías

- La programación permite la integración de diversas herramientas y tecnologías en el flujo de trabajo de análisis de datos. Esto incluye la conexión a bases de datos, el uso de servicios en la nube y la implementación de soluciones de Big Data.

En resumen, la programación en análisis de datos es esencial para abordar tareas desde la preparación inicial de datos hasta la implementación de modelos avanzados. Proporciona flexibilidad, automatización y la capacidad de trabajar eficientemente con grandes volúmenes de datos, contribuyendo al proceso de toma de decisiones basado en información precisa y valiosa.

Lenguajes:



PYTHON



R



D. ESTADÍSTICA

La estadística juega un papel fundamental en el mundo de los datos, proporcionando métodos y técnicas para describir, analizar y entender patrones y tendencias en conjuntos de datos.

A Continuación un detalle de como se usa la estadística en análisis de datos:

1. Descripción de Datos

La estadística descriptiva se utiliza para resumir y describir características clave de un conjunto de datos, como la media, la mediana, la moda, la desviación estándar y los cuartiles. Estos resúmenes permiten una comprensión inicial de la distribución y la variabilidad de los datos.

2. Regresión Estadística:

La regresión se utiliza para modelar la relación entre una variable dependiente y una o más variables independientes. Esto ayuda a entender la naturaleza de la asociación y predecir valores futuros.

3. Análisis de Varianza (ANOVA):

ANOVA se emplea para comparar las medias de tres o más grupos y determinar si existen diferencias significativas entre ellos. Es útil para evaluar el impacto de variables categóricas en una variable numérica.

4. Inferencia Estadística:

La inferencia estadística se ocupa de hacer afirmaciones o inferencias sobre una población basada en una muestra de datos. Esto incluye la estimación de parámetros y la realización de pruebas de hipótesis para tomar decisiones informadas sobre características desconocidas de la población.

5. Análisis Exploratorio de Datos (EDA):

La EDA utiliza técnicas estadísticas y gráficos para explorar visualmente las relaciones entre variables, identificar patrones y detectar posibles valores atípicos. Esto ayuda a formular preguntas y generar hipótesis antes de realizar análisis más avanzados.

6. Diseño Experimental:

En experimentos controlados, la estadística se emplea para diseñar la estructura del experimento y analizar los resultados. Esto incluye la asignación aleatoria de tratamientos y la aplicación de diseños estadísticos para maximizar la validez de las conclusiones.

7. Análisis de Series Temporales:

La estadística es fundamental en el análisis de series temporales para modelar y pronosticar patrones temporales en los datos, como tendencias, estacionalidades y ciclos.

8. Pruebas de Hipótesis:

Las pruebas de hipótesis estadísticas se utilizan para tomar decisiones basadas en evidencia sobre afirmaciones relacionadas con los datos. Esto incluye pruebas t como la prueba t de Student y pruebas no paramétricas como la prueba de Wilcoxon.

9. Análisis de Correlación:

El análisis de correlación evalúa la relación entre dos variables. La correlación de Pearson mide la fuerza y la dirección de una relación lineal, mientras que métodos como la correlación de Spearman pueden manejar relaciones no lineales.

10. Distribuciones de Probabilidad:

Se utilizan distribuciones de probabilidad para modelar y entender la variabilidad en los datos. La distribución normal es comúnmente utilizada, pero existen otras distribuciones específicas para ciertos tipos de datos.

En resumen, la estadística proporciona las herramientas para explorar, resumir, inferir y tomar decisiones basadas en datos. Desde la descripción inicial de datos hasta la construcción de modelos y la validación de resultados, la estadística es esencial para el análisis efectivo y la obtención de información significativa a partir de conjuntos de datos.

PARTE VI

> ¿COMO SIGO? 

A. ¿QUÉ TRABAJOS PODRÍA CONSEGUIR SI ESTUDIO DATOS?

Estudiar análisis de datos o business intelligence puede abrirte las puertas a una variedad de roles y trabajos en el campo de la analítica y la toma de decisiones empresariales. Aquí hay una lista de algunos de los roles comunes a los que podrías aspirar al especializarte en estas áreas:

Roles relacionados con el Análisis de Datos:



Científico de Datos: Los científicos de datos son expertos en el análisis de datos, utilizan técnicas estadísticas y de aprendizaje automático para descubrir patrones y tendencias en los datos y desarrollar modelos predictivos.



Analista de Datos: Los analistas de datos se centran en la recopilación, limpieza y análisis de datos para proporcionar información valiosa a las empresas. Pueden trabajar en diversas industrias, como marketing, finanzas o salud.





Ingeniero de Datos: Los ingenieros de datos diseñan y mantienen infraestructuras de datos, creando pipelines de datos para la recopilación y procesamiento de información de diversas fuentes.



Analista de Negocios o Business Analyst: Los analistas de negocios utilizan datos para identificar oportunidades de mejora en las operaciones de una empresa y proponer soluciones basadas en datos.



Administrador de Bases de Datos: Se encarga de la gestión y administración de bases de datos, asegurando su disponibilidad, seguridad y rendimiento.

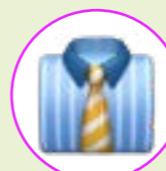


Roles relacionados con Business Intelligence:



Consultor de Business Intelligence: Los consultores de BI trabajan con empresas para desarrollar estrategias de inteligencia de negocios, implementar herramientas de BI y optimizar procesos de informes.

Desarrollador de BI: Los desarrolladores de BI diseñan y mantienen sistemas de inteligencia de negocios, creando informes, paneles de control y cuadros de mando utilizando herramientas como Tableau, Power BI o QlikView.



Gerente de BI: Los gerentes de BI supervisan los proyectos de BI de una organización, dirigen equipos de analistas y aseguran que los esfuerzos de BI estén alineados con los objetivos empresariales.



Analista de Datos Financieros: Estos profesionales se especializan en la interpretación de datos financieros y utilizan herramientas de BI para ayudar en la toma de decisiones financieras.



Especialista en Marketing Digital (Analista digital): Los especialistas en marketing digital utilizan datos para optimizar estrategias de marketing en línea, realizando análisis de rendimiento y segmentación de audiencias.



Analista de Riesgos: En instituciones financieras, los analistas de riesgos utilizan análisis de datos para evaluar y gestionar riesgos financieros.

Estos son solo algunos ejemplos de los roles que podrías obtener al estudiar análisis de datos o business intelligence. El campo de la analítica de datos es amplio y en constante evolución, lo que significa que hay muchas oportunidades de carrera disponibles en una variedad de industrias y sectores. Tu elección dependerá de tus intereses, habilidades y objetivos profesionales.



B. APLICACIÓN SEGÚN TU PERFIL

Si estás aquí seguramente caigas en alguno de estos perfiles:

- 01 Trabajar para una empresa como analista de datos
- 02 Para un emprendedor o CEO
- 03 Complementar mis habilidades actuales
- 04 Aplicarlo en mi trabajo
- 05 Ser freelance



Te cuento porque te va a servir aprender análisis de datos:



1. Trabajar para una empresa como analista de datos

Aprender análisis de datos te abrirá las puertas a un mundo de oportunidades laborales. Como analista de datos, serás un activo invaluable para las empresas, ya que podrás descubrir insights que impulsen la toma de decisiones estratégicas. Este rol te permitirá sentirte realizado al aplicar tus habilidades para resolver problemas y contribuir al éxito de la empresa.



2. Ser freelance:

Ser freelance te brinda la flexibilidad de ofrecer servicios completos e integrales. Aprender análisis de datos te permitirá destacarte al ofrecer a tus clientes la capacidad de analizar y presentar datos de manera efectiva. Esto te convertirá en un profesional independiente más valioso y competitivo en el mercado.



3. Para un emprendedor o CEO:

Como emprendedor o CEO, el análisis de datos te empoderará para tomar decisiones empresariales más acertadas. Te brindará la capacidad de comprender en profundidad el rendimiento de tu negocio, identificar oportunidades de crecimiento y adaptarte rápidamente a las demandas del mercado. Serás capaz de liderar con visión el equipo de datos porque no solo entenderás de tecnología sino que serás el que más conoce el negocio y sus necesidades.



4. Complementar mis habilidades actuales:

Independientemente de tu rol actual, el análisis de datos es una habilidad complementaria poderosa. Mejorará tu capacidad para entender y utilizar datos, fortaleciendo tus habilidades profesionales y haciéndote más versátil en cualquier entorno laboral.



5. Aplicarlo en mi trabajo:

Integrar el análisis de datos en tu trabajo diario te permitirá ser más eficiente y efectivo. Automatizar tareas, comprender a fondo los procesos y tomar decisiones basadas en datos se convertirán en parte integral de tu enfoque laboral, brindándote una ventaja en la calidad de tus resultados y convirtiéndote en un profesional versátil y 360 para un negocio.



G. PALABRAS FINALES

Así como no todo es color de rosa a mi me gusta ser transparente. Te muestro el lado A, pero también el B. **Aprender análisis de datos no es fácil, tampoco es difícil.. pero requiere un compromiso y una disciplina de tu parte.** No se aprende análisis de datos consumiendo vídeos de forma pasiva , sino que requiere práctica. Si lo vas a usar como un conocimiento complementario genial. Pero no es algo que se aprende de la noche a la mañana, estamos en constante aprendizaje y trabajando con **tecnologías nuevas**.

Vas a experimentar mucha frustración al inicio sobre todo si no venís de carreras técnicas. No porque sea difícil, sino porque no nos enseñaron en la escuela a pensar o trabajar de la forma que se hace en tecnología.

Si estas comprometido a lograrlo requiere esfuerzo, pero al otro lado del camino te espera un futuro brillante. A mi el análisis de datos me ha abierto demasiadas puertas, siempre voy a motivar a todos a hacerlo porque **mi experiencia ha sido increíble**.



¡Felicitaciones por dar este paso tan importante! Este e-book te ha proporcionado insights valiosos. Ahora, prepárate para el próximo paso con nosotros. Estamos aquí para guiarte y apoyarte en cada etapa de tu viaje hacia el análisis de datos. Gracias por confiar en nosotros para guiarte en este emocionante viaje.

Espero verte pronto. Caterina





UNICORN project

Tu futuro tiene una nueva oportunidad,
Unicorn Project.