**Engineer**

# analítica de datos

procesar datos con método para hallar patrones y tendencias

## Qué es

La analítica de datos, a menudo llamada simplemente analítica, transforma hechos registrados en respuestas sobre las que alguien pueda actuar. En el marco de big data publicado por NIST, el instituto estadounidense de estándares, consiste en trabajar metódicamente con los datos, transformándolos cuando sea necesario, para revelar patrones, relaciones entre puntos de datos y tendencias a lo largo del tiempo, e intentar predecir lo que sucederá después. NIST sitúa ese paso en un *ciclo de vida de analítica de datos* más amplio que convierte datos brutos en conocimiento utilizable por una organización: recopilar, preparar, analizar, visualizar y facilitar acceso.

Es un tipo de procesamiento distinto del transaccional por lo que lee: una aplicación escribe un pedido cada vez; una consulta analítica lee millones de filas para devolver unas pocas. También difiere de la ingeniería de datos, que construye los flujos de procesamiento y el almacenamiento de los que lee la analítica.

El término comprende trabajos bastante distintos:

- **Informes**: las mismas cifras entregadas con una programación, como PDF o correo.
- **Paneles**: esas cifras actualizadas en una herramienta de inteligencia de negocio como Power BI.
- **Análisis ad hoc**: una consulta SQL para una pregunta antes no planteada, a menudo sobre un almacén como BigQuery.
- **Análisis geoespacial**: el mismo trabajo sobre datos cartográficos en un SIG, preguntando dónde además de cuánto.
- **Aplicaciones analíticas**: un modelo o método estadístico tras un formulario para quienes no escriben código.

- **Analítica web y de producto**: contar visitantes de un sitio o usuarios de un producto y sus acciones.

La última acarrea un coste de privacidad, y a ella se refieren las páginas de privacidad, cookies, créditos y seguridad de este sitio cuando dicen «no analytics»: ninguna etiqueta de analítica, ningún píxel de seguimiento, ninguna cookie ni cuenta con una empresa de analítica. El sitio sí se mide. Su servidor web conserva un registro de acceso y un programa lo transforma en recuentos diarios y una clasificación mensual de páginas leídas, sitios de origen y rastreadores, conservados como cifras y nunca unidos para reconstruir el recorrido de un visitante concreto.

Utilice el concepto en una decisión

Aplique «¿Qué es la analítica de datos?» a una decisión real. Podemos ayudarle a relacionar el concepto con el sistema que está construyendo o comprando. Paneles e informes con las pocas cifras de las que realmente depende una decisión, actualizados sin tener que pedirlo.

**Explore nuestros servicios** Póngase en contacto →

Compruebe los supuestos detrás del término

Utilice la explicación y los trabajos enlazados para formular preguntas sobre su proyecto. Podemos examinar los supuestos con usted antes de elegir una herramienta. Comunicación de marca guiada por la respuesta del público — comunicación de marca guiada por la respuesta del público

**Comunicación de marca guiada por la respuesta del público** Póngase en contacto →

## Por qué importa

Dos decisiones fijan la mayor parte del coste: dónde se ejecutan las consultas y cómo se recopilan los datos de personas.

Analizar sobre la base transaccional es el comienzo más barato: los datos están al día y no se copia nada, pero cada informe compite con la aplicación por memoria y disco. Una réplica de lectura descarga al primario por el precio de otro servidor. Un almacén como BigQuery separa almacenamiento y cómputo, factura cada consulta bajo demanda por los bytes leídos, lo que premia las consultas acotadas, y necesita un flujo que copie los datos.

La dependencia del proveedor se oculta en las definiciones. Una herramienta BI guarda sus medidas en su lenguaje de fórmulas; un año después, la definición de cliente activo vive en un archivo que solo abre esa herramienta. Las definiciones SQL en el repositorio sobreviven al cambio de herramienta.

La analítica web y de producto añade una cuestión legal. La herramienta alojada habitual usa una etiqueta: un script en cada página que fija una cookie u otro identificador en el navegador y comunica las visitas al proveedor. El artículo 5(3) de la Directiva ePrivacy de la UE exige consentimiento para guardar información en el dispositivo del visitante o leerla, con dos excepciones: hacerlo solo para transmitir una comunicación por una red o cuando sea estrictamente necesario para un servicio pedido expresamente. Cuando los eventos pueden vincularse a una persona, incluso mediante una cookie o IP, también se aplica el RGPD: base jurídica, finalidad declarada y plazo de retención. Además, el proveedor es otra empresa que ahora conserva un registro de la actividad de sus visitantes.

La elección real está entre una herramienta alojada rápida de añadir y con paneles, una recopilación propia que construye y opera usted y recuentos a partir de los registros ya escritos por el servidor, que responden a menos preguntas y contienen mucho menos.

## Cómo se aplica en la práctica

Para una plataforma de contactos profesionales, la empresa construyó analítica de producto dentro de PostgreSQL en lugar de una etiqueta de terceros: un esquema de 7 tablas y 47 funciones, tablas de eventos particionadas con 22 definiciones de partición que se depuran automáticamente según una programación, y la identidad del visitante seudonimizada mediante una sal rotatoria. El cliente no envía nada hasta que el visitante ha consentido y los embudos se calculan en la base de datos, junto a los datos, lo que NIST llama localidad de datos. La página nombra el precio: la sal rotatoria obliga a renunciar al análisis de cohortes a largo plazo y no vienen incluidos paneles gratuitos.

En informes, un sistema de informes en Python extraía los datos, ejecutaba el análisis y entregaba el mismo informe PDF con un calendario predecible a quienes tomaban decisiones comerciales. La página atribuye a esos informes un aumento del 400 % de los ingresos trimestrales de software.

Este sitio es el tercer ejemplo. Su página de privacidad detalla la medición: un registro de acceso que sustituye por ceros el último tramo de cada dirección IP antes de escribir una línea, el sitio de origen conservado solo como nombre y líneas que duran unos treinta días. Un programa lee el registro para generar un recuento diario y una clasificación mensual, ambos conservados indefinidamente como cifras sin nada que permita volver a unirlas en una visita. Sin etiqueta, sin cookie y sin terceros, la página no muestra un banner de consentimiento y da

como motivo «nada que consentir». También declara su límite: mientras existe una línea, un navegador poco común en una red pequeña permite acotar a un visitante más de lo que el sitio desearía. La contrapartida son menos respuestas y ningún embudo, a cambio de no almacenar nada en el dispositivo del lector ni enviar nada a nadie más.

# Dónde surgen los problemas

## Dos herramientas dan dos cifras de visitantes

La herramienta alojada muestra una cifra del mes pasado y el registro el doble; ambas aciertan sobre cosas diferentes. La etiqueta cuenta solo navegadores que ejecutaron el script y, con banner, visitantes que aceptaron. El registro cuenta cada solicitud, incluidos rastreadores, hasta filtrarlos. Escriba una sola vez en SQL la definición de cada métrica, con inclusiones y exclusiones.

## Los informes ralentizan la aplicación

Cada lunes la aplicación se ralentiza mientras el informe semanal recorre tablas completas en la base transaccional. Trasládelo a una réplica o almacén, o agregue las filas brutas cada noche en una pequeña tabla resumen y aplique funciones de ventana:

```
SELECT day,
       signups,
       sum(signups) OVER (ORDER BY day
                          ROWS BETWEEN 6 PRECEDING AND CURRENT ROW)
         AS last_7_days
FROM daily_counts
ORDER BY day;
```

## Desaparece el banner y permanece el seguimiento

Un equipo sustituye las cookies por un producto «sin cookies», retira el banner de consentimiento y sigue rastreando a los visitantes. El artículo 5(3) se refiere al acceso a información del dispositivo, sea cual sea el mecanismo, y las directrices del EDPB sobre su ámbito técnico indican que también pueden quedar comprendidos el fingerprinting mediante cabeceras HTTP o el seguimiento de la navegación únicamente por dirección IP. A la vez, recuerdan que el hecho de que se aplique la norma no implica por sí solo que sea necesario el consentimiento. Juzgue una herramienta por si vincula una solicitud con la siguiente, no por si instala una cookie.

### Un consentimiento que rechazarían los reguladores

Un banner bloquea la página hasta aceptar, interpreta desplazarse como acuerdo o dificulta retirar más que dar el consentimiento. Las directrices del EDPB dicen que el muro de cookies impide un consentimiento libre y que desplazarse nunca es un acto afirmativo claro. El artículo 7(3) del RGPD exige que retirarlo sea tan fácil como darlo. Dé el mismo peso al rechazo y a la aceptación y no envíe nada antes de la respuesta.

### Eventos seudónimos conservados para siempre

La tabla de eventos crece en millones de filas al mes y los identificadores de usuario sometidos a hash se tratan como anónimos. El considerando 26 del RGPD trata como datos personales los datos seudonimizados que podrían atribuirse a alguien mediante información adicional, y el considerando 30 cita las direcciones IP y los identificadores de cookies entre los identificadores en línea cuyas huellas, combinadas con otra información, pueden utilizarse para elaborar perfiles e identificar a alguien. Establezca el plazo de retención antes de que llegue el primer evento, particione por tiempo para eliminar de una vez un mes vencido y conserve las cifras a largo plazo como agregados sin identificadores.

## Lecturas adicionales

- NIST SP 1500-1r2, Big Data Interoperability Framework: Definitions — las definiciones de analítica, ciclo de vida de analítica de datos y localidad de datos de las que parte este artículo.
- Directiva 2002/58/CE consolidada — artículo 5(3), la regla de consentimiento para todo lo almacenado en el dispositivo de un visitante o leído de él, con sus dos excepciones.
- Reglamento (UE) 2016/679, el RGPD — la definición de consentimiento, el artículo 7 sobre su retirada y los considerandos 26 y 30 sobre datos seudonimizados e identificadores en línea.
- Directrices 05/2020 del EDPB sobre consentimiento — muros de cookies, desplazamiento y retirada, cada uno con ejemplos desarrollados.
- Directrices 2/2023 del EDPB sobre el artículo 5(3) — por qué el seguimiento basado en IP, el fingerprinting y los píxeles de seguimiento pueden quedar sujetos a la misma regla que las cookies.
- Tutorial de PostgreSQL: funciones de ventana — clasificaciones y totales acumulados sobre filas relacionadas, calculados dentro de la base de datos.

Glosario

Elija según sus restricciones reales

Traiga un conjunto de datos representativo, las preguntas que debe responder y el cuello de botella actual. Podemos acordar cómo medir la calidad, la capacidad de procesamiento y la recuperación.

**Vea cuánto cuesta** Póngase en contacto →

# Haga práctica la terminología

Pregunte cómo se aplica «¿Qué es la analítica de datos?» a su organización. Basta un ejemplo concreto de su sistema actual para iniciar la conversación.

**Explore nuestros servicios**

welcome@engineer.company +45 50 10 61 84 Copenhague, Dinamarca

Llévese esta página

Tarjeta de contacto

**Creado con esmero y cariño por D.**

**CVR 45969703 · Tagensvej 105C, 3. th, 2200 København N**

**Derechos de autor © 2025 – actualidad · Engineer ApS · Todos los derechos reservados.**

**Impreso desde https://engineer.company/es/glossary/analytics/**