

FEPOL  
**Fondo**  
Editorial



**Gestión de la Información para la  
Ingeniería Industrial**

**JULIO CÉSAR ALVAREZ REYES**

*Professionals*  
**on line** S.A.C.

© **Julio César Alvarez Reyes**

<https://orcid.org/0000-0001-8465-8907>

Editada por:

© Professionals On Line SAC. (FEPOL) – Fondo Editorial.

Dirección Av. General José de San Martín N° 790, Dpto. 402, Urb. El Carmen - Pueblo Libre.

[professionalsonline.net@gmail.com](mailto:professionalsonline.net@gmail.com)

Teléf. móvil: +51 981 105 611

Web: <https://professionals.pe/>

Primera edición digital: Abril de 2026

Libro digital disponible en <https://editorialfondo.com/>

Depósito Legal en la Biblioteca Nacional del Perú N° 2026-04527

ISBN: 978-612-99409-0-8

DOI: <https://doi.org/10.47422/ce8dqf65>

Diseño y Diagramación: Gráfica “Imagen”

Professionals On Line

[professionalsonline.net@gmail.com](mailto:professionalsonline.net@gmail.com)

Sello editorial: Fondo Editorial (978-612-48981)



Licencia No Comercial Creative Commons

<https://creativecommons.org/licenses/by-nc/4.0/deed.es>

---

Professionals On Line (FEPOL)

ISSN: 2955-8549

**Este libro está dedicado a:**

A mis estudiantes, por desafiarme a entender el conocimiento como un proceso de mejora permanente.

A mis padres, cuyo ejemplo y valores son los cimientos de mi formación personal y profesional.

A mi esposa e hijos, por ser la inspiración y el propósito detrás de cada página de este libro. Ustedes son mi mayor aprendizaje.

## **Prefacio**

La gestión de la información se ha consolidado como un elemento fundamental en la Ingeniería Industrial contemporánea. En un entorno caracterizado por la transformación digital, la automatización de procesos y el crecimiento constante de los datos organizacionales, el desafío ya no radica únicamente en disponer de información, sino en transformarla en conocimiento útil para la toma de decisiones.

Tradicionalmente, la información fue utilizada principalmente como soporte para el control operativo. Sin embargo, el desarrollo de la analítica de datos, la estadística aplicada y el aprendizaje automático ha ampliado su alcance hacia un enfoque estratégico, orientado a comprender el comportamiento de los procesos, reducir la incertidumbre y fortalecer la mejora continua.

En este contexto, el presente libro aborda la gestión de la información desde una perspectiva propia de la Ingeniería Industrial, integrando fundamentos conceptuales con aplicaciones prácticas. Más allá del uso de herramientas tecnológicas, se enfatiza el criterio analítico del ingeniero industrial como elemento central para interpretar los datos, diseñar indicadores y vincular la información con los procesos organizacionales.

La obra se estructura en cuatro partes progresivas que articulan conceptos, metodología y aplicación práctica:

Parte I. La información como activo estratégico en la Ingeniería Industrial. Se desarrollan los fundamentos de la gestión de la información en las

organizaciones, analizando la relación entre procesos, datos y sistemas. Se introduce el rol del ingeniero industrial como gestor de información, la arquitectura básica de datos, la ingeniería de indicadores y los principios de una gestión basada en evidencia. Asimismo, se presentan casos aplicados que permiten comprender cómo la información refleja el comportamiento real de los procesos.

Parte II. Inteligencia de negocios aplicada a la Ingeniería Industrial. Se aborda el uso de la inteligencia de negocios como herramienta para el control gerencial y la mejora continua. Se desarrollan conceptos de modelado de información, construcción de indicadores y diseño de dashboards, incorporando casos prácticos orientados al análisis de desempeño organizacional mediante herramientas como Microsoft Power BI.

Parte III. Estadística e inferencia estadística para la toma de decisiones. Se presentan los fundamentos de la estadística descriptiva y la inferencia estadística aplicadas al análisis de procesos industriales. Se estudia la variabilidad, la interpretación de datos y la validación de hipótesis como base para la toma de decisiones objetivas, incorporando ejemplos prácticos orientados al análisis de datos reales.

Parte IV. Machine Learning e Inteligencia Artificial en la Ingeniería Industrial. Se introduce el aprendizaje automático como una extensión natural del análisis estadístico, mostrando aplicaciones en contextos industriales como mantenimiento predictivo, seguridad y salud en el

trabajo, y análisis organizacional. Se enfatiza el enfoque conceptual necesario para que el ingeniero industrial comprenda la lógica de los modelos y su aplicación responsable en la toma de decisiones.

Dirigido a estudiantes y profesionales, este libro busca desarrollar la capacidad de analizar información de manera crítica, transformar datos en conocimiento y promover decisiones sustentadas en evidencia. En un entorno donde las organizaciones son cada vez más intensivas en información, estas competencias se convierten en un elemento clave para el desempeño del ingeniero industrial moderno.

## Tabla de contenido

PARTE I. La información como activo estratégico en la Ingeniería Industrial .....	8
Capítulo 1. La gestión de la información en la era industrial y digital .....	8
Capítulo 2. Sistemas de información y procesos organizacionales .....	16
Capítulo 3. Hacia una gestión de la información basada en datos .....	29
Capítulo 4. Arquitectura de información y almacenamiento de datos .....	37
PARTE II. Inteligencia de Negocios aplicada a la Ingeniería Industrial .....	48
Capítulo 5. Inteligencia de Negocios en contextos industriales.....	48
Capítulo 6. Diseño de reportes y dashboards para control gerencial .....	53
Capítulo 7. Caso de estudio en Inteligencia de Negocios. ....	59
PARTE III. Estadística e inferencia estadística para la toma de decisiones .....	77
Capítulo 8. Estadística aplicada a la ingeniería industrial .....	77
Capítulo 9. Inferencia estadística en problemas industriales .....	97
PARTE IV. Machine Learning e Inteligencia Artificial en la Ingeniería Industrial.....	109
Capítulo 10. Machine Learning como extensión del análisis industrial .....	109
Capítulo 11. Aplicaciones de Machine Learning en la ingeniería industrial .....	115
Capítulo 12. Inteligencia Artificial y apoyo a la toma de decisiones .....	121

## **PARTE I. La información como activo estratégico en la Ingeniería Industrial**

### **Capítulo 1. La gestión de la información en la era industrial y digital**

En la actualidad, las organizaciones industriales operan en entornos caracterizados por una alta complejidad, dinamismo y presión competitiva. La automatización de procesos, la digitalización de operaciones y la interconexión de sistemas han multiplicado la cantidad de datos disponibles, generando nuevas oportunidades, pero también nuevos desafíos para la gestión empresarial.

En este contexto, la información se consolida como un activo estratégico fundamental. No se trata únicamente de disponer de grandes volúmenes de datos, sino de contar con la capacidad organizacional para transformarlos en información confiable y, posteriormente, en conocimiento útil para la toma de decisiones. La Ingeniería Industrial, por su naturaleza integradora, ocupa un rol central en este proceso.

#### **1.1 De la información operativa a la información estratégica**

Tradicionalmente, la información en las organizaciones industriales ha sido utilizada principalmente con fines operativos: control de producción, registro de inventarios, seguimiento de costos o verificación de cumplimiento. Si bien estas funciones siguen siendo

esenciales, el entorno actual exige una evolución hacia un uso más estratégico de la información.

La información estratégica permite:

### **1.1.1 Anticipar comportamientos futuros**

La información estratégica permite proyectar el comportamiento futuro de los procesos, recursos y sistemas organizacionales a partir del análisis de datos históricos y actuales. Mediante el estudio de patrones de desempeño, variabilidad y recurrencia de eventos, el ingeniero industrial puede identificar señales tempranas que anticipen escenarios futuros, como incrementos en la demanda, fallas operativas, riesgos de seguridad o cuellos de botella productivos. Esta capacidad de anticipación resulta fundamental para la planificación, la asignación eficiente de recursos y la prevención de eventos no deseados.

Ejemplo: En una planta de producción de alimentos, el análisis histórico de volúmenes de producción, tiempos de parada y demanda estacional permite anticipar incrementos en la carga operativa durante determinados meses del año. Con esta información, el ingeniero industrial puede planificar turnos adicionales, ajustar niveles de inventario y programar mantenimientos preventivos antes de los periodos críticos, evitando sobrecostos y retrasos en la entrega.

### **1.1.2 Identificar patrones y tendencias**

El análisis sistemático de la información facilita la identificación de patrones y tendencias que no son evidentes en la observación aislada de los datos. La información estratégica permite detectar comportamientos recurrentes, relaciones entre variables y cambios graduales en el desempeño de los procesos. Desde la perspectiva de la Ingeniería Industrial, el reconocimiento de estas tendencias constituye la base para la mejora continua, ya que posibilita comprender las causas subyacentes del comportamiento del sistema y orientar acciones correctivas y preventivas con sustento objetivo.

Ejemplo: En una empresa manufacturera, el análisis de registros de fallas en maquinaria revela que determinados equipos presentan un aumento progresivo en la frecuencia de fallas después de cierto número de horas de operación. Este patrón permite identificar una tendencia de deterioro del activo y justificar la implementación de planes de mantenimiento predictivo, reduciendo paradas no programadas y pérdidas de producción.

### **1.1.3 Evaluar escenarios alternativos**

La información estratégica habilita la evaluación de distintos escenarios de decisión antes de su implementación. A través de simulaciones, análisis comparativos y modelos de proyección, es posible estimar el impacto de diversas alternativas sobre los indicadores clave de desempeño. Esta capacidad resulta especialmente

relevante en decisiones que involucran cambios en procesos, inversiones, políticas operativas o asignación de recursos. Evaluar escenarios alternativos reduce el riesgo asociado a la toma de decisiones y fortalece el criterio técnico del ingeniero industrial.

Ejemplo: Una organización logística evalúa dos alternativas para mejorar su nivel de servicio: ampliar su flota de transporte o reorganizar sus rutas de distribución. A partir de información histórica sobre tiempos de entrega, costos operativos y volumen de pedidos, el ingeniero industrial simula ambos escenarios y determina cuál ofrece mayor impacto en el indicador de cumplimiento con menor inversión, antes de ejecutar la decisión.

### **1.1.4 Reducir la incertidumbre en la toma de decisiones.**

Uno de los principales aportes de la información estratégica es la reducción de la incertidumbre inherente a la gestión organizacional. Si bien ninguna decisión puede eliminar completamente el riesgo, el uso de información confiable y oportuna permite disminuir la dependencia de la intuición y la experiencia aislada. En su lugar, las decisiones se sustentan en evidencia objetiva, indicadores consistentes y análisis estructurados, lo que incrementa la probabilidad de obtener resultados favorables y alineados con los objetivos estratégicos de la organización.

Ejemplo: En un proceso de gestión de seguridad y salud en el trabajo, la información consolidada sobre accidentes, actos inseguros y

condiciones peligrosas permite priorizar acciones preventivas en áreas críticas. En lugar de intervenir de forma generalizada, la organización dirige sus esfuerzos hacia los factores de mayor incidencia, reduciendo la incertidumbre y mejorando la efectividad de las decisiones preventivas.



*Fig. 1 De los datos al conocimiento*

### 1.2 El ingeniero industrial como gestor de información

El ingeniero industrial no solo analiza procesos físicos o administrativos, sino también los flujos de información que los sustentan. Desde esta perspectiva, su rol se amplía hacia la gestión integral de la información, actuando como nexo entre los datos generados por los sistemas y las decisiones que orientan la operación y la estrategia.

Gestionar información implica:

### **1.2.1 Comprender el origen y significado de los datos**

El primer paso en la gestión de la información consiste en comprender el origen de los datos y el contexto en el que son generados. Cada dato está asociado a un proceso, una actividad y una finalidad específica. El ingeniero industrial debe identificar qué proceso produce el dato, bajo qué condiciones se registra y qué representa realmente dentro del sistema organizacional. Sin esta comprensión, los datos pueden ser interpretados de forma errónea, generando conclusiones imprecisas y decisiones inadecuadas.

Ejemplo: En una planta de manufactura, el registro del tiempo de ciclo de una línea de producción puede provenir de sensores automáticos o de registros manuales del operario. Comprender el origen del dato permite al ingeniero industrial distinguir entre tiempos reales de operación y tiempos afectados por paradas o reprocesos, evitando interpretaciones erróneas sobre la eficiencia del proceso.

### **1.2.2 Evaluar su calidad y relevancia**

No todos los datos disponibles son útiles para la gestión. Evaluar la calidad implica verificar que los datos sean completos, consistentes, oportunos y confiables. Asimismo, evaluar su relevancia supone determinar si los datos aportan valor para el análisis del proceso y la toma de decisiones. La acumulación de datos irrelevantes incrementa la complejidad del análisis sin generar beneficios reales, mientras que

la selección adecuada de información permite focalizar los esfuerzos en los factores críticos del desempeño organizacional.

Ejemplo: En un sistema de control de calidad, se dispone de múltiples registros de inspección. Sin embargo, solo aquellos datos que cumplen con criterios de completitud y consistencia son considerados para el análisis de defectos. Al filtrar registros incompletos o duplicados, el ingeniero industrial mejora la confiabilidad del diagnóstico y enfoca la gestión en los defectos con mayor impacto.

### **1.2.3 Estructurarlos adecuadamente**

La información debe organizarse de manera que facilite su análisis y utilización. Esto implica definir estructuras de datos coherentes con los procesos que representan, establecer relaciones claras entre variables y asegurar la trazabilidad de la información a lo largo del tiempo. Una adecuada estructuración permite consolidar información proveniente de diversas fuentes, reducir redundancias y habilitar análisis multidimensionales, fundamentales para la Ingeniería Industrial.

Ejemplo: Una empresa de servicios registra información de atención al cliente en diferentes áreas. Al estructurar estos datos en una sola tabla analítica que integra fecha, tipo de servicio, área responsable y tiempo de atención, el ingeniero industrial habilita un análisis multidimensional que permite identificar cuellos de botella y oportunidades de mejora en el proceso de atención.

### **1.2.4 Utilizarlos como soporte para la mejora continua**

El objetivo final de la gestión de la información es apoyar la mejora continua de los procesos organizacionales. La información bien gestionada permite identificar desviaciones, analizar causas, evaluar el impacto de las acciones correctivas y medir los resultados obtenidos. De esta manera, la información se convierte en un insumo activo del ciclo de mejora continua, fortaleciendo la capacidad de la organización para adaptarse, aprender y optimizar su desempeño de forma sostenida.

Ejemplo: En un proceso de seguridad y salud en el trabajo, la información histórica de incidentes y accidentes se utiliza para evaluar la efectividad de las acciones preventivas implementadas. Al comparar los indicadores antes y después de una intervención, la organización puede verificar si las medidas adoptadas redujeron los riesgos y ajustar su estrategia de mejora continua.

Este enfoque constituye la base sobre la cual se desarrolla el resto del libro.

## **Capítulo 2. Sistemas de información y procesos organizacionales**

Los sistemas de información son el soporte tecnológico que permite capturar, almacenar, procesar y distribuir información dentro de la organización. Sin embargo, su verdadero valor no reside en la tecnología utilizada, sino en su alineación con los procesos organizacionales y con las necesidades de gestión.

### **2.1 Información y procesos de negocio**

Todo proceso organizacional —productivo, logístico, comercial o administrativo— genera y consume información. Cada actividad produce registros que, adecuadamente gestionados, permiten controlar el desempeño del proceso y evaluar oportunidades de mejora.

Desde la Ingeniería Industrial, resulta fundamental comprender que:

#### **2.1.1 Los procesos definen qué información es relevante**

Todo proceso organizacional —productivo, logístico, administrativo o de servicios— persigue un objetivo específico y transforma entradas en salidas bajo determinadas condiciones de tiempo, costo, calidad y riesgo.

Desde la perspectiva de la Ingeniería Industrial, no toda la información es valiosa, sino únicamente aquella que permite controlar, evaluar y mejorar el desempeño del proceso.

Por ello, la identificación de la información relevante debe partir siempre del análisis del proceso:

- ¿Qué actividades lo componen?
- ¿Cuáles son sus puntos críticos?
- ¿Qué variables influyen directamente en su desempeño?
- ¿Qué decisiones deben tomarse durante su ejecución?

Indicadores como tiempos de ciclo, niveles de inventario, tasas de defectos, capacidad utilizada o cumplimiento de plazos surgen del proceso mismo y determinan qué datos deben ser capturados, almacenados y analizados. Cuando la información se define sin esta base, se corre el riesgo de generar datos abundantes pero irrelevantes, que no aportan valor a la toma de decisiones.

### **2.1.2 La información refleja el comportamiento real del proceso**

La información es el espejo del proceso. A través de ella se observa cómo el proceso opera realmente y no cómo fue diseñado o cómo se supone que debería funcionar.

Registros históricos, indicadores de desempeño y series de datos permiten identificar patrones, variabilidad, cuellos de botella, fallas recurrentes y oportunidades de mejora.

Desde este enfoque, la información cumple una doble función:

- Descriptiva, al mostrar el estado actual y pasado del proceso.
- Analítica, al permitir explicar causas, evaluar escenarios y predecir comportamientos futuros.

Cuando la información es confiable y oportuna, se convierte en la base para aplicar técnicas de inferencia estadística, inteligencia de negocios y modelos predictivos. En cambio, información incompleta, sesgada o mal estructurada conduce a diagnósticos erróneos y decisiones ineficientes, aun cuando se utilicen herramientas analíticas avanzadas.

### **2.1.3 Los sistemas deben diseñarse en función de esa relación**

Los sistemas de información no deben concebirse como soluciones tecnológicas aisladas, sino como instrumentos al servicio de los procesos y de la información que estos generan.

Un sistema bien diseñado es aquel que:

- Captura los datos críticos del proceso en el momento adecuado.
- Estructura la información de forma coherente con la lógica operativa.
- Facilita el análisis y la toma de decisiones en los distintos niveles de la organización.

Desde la Ingeniería Industrial, el diseño de sistemas debe seguir una lógica “de afuera hacia adentro”: primero se comprenden los procesos, luego se define la información necesaria y finalmente se selecciona o construye el sistema que soporte esa relación. Invertir este orden — diseñar sistemas sin entender procesos— suele derivar en soluciones rígidas, subutilizadas o desconectadas de la realidad operativa.

En síntesis, los procesos determinan la información, la información revela el proceso y los sistemas deben integrarlos de manera coherente. Esta visión permite al ingeniero industrial actuar como articulador entre la operación, la analítica y la tecnología, asegurando que la gestión de la información contribuya efectivamente a la mejora continua y a la toma de decisiones estratégicas.

### **2.2 Caso de estudio. Proceso logístico - gestión de inventarios en un almacén de distribución**

**Contexto organizacional.** Una empresa distribuidora de productos de consumo masivo opera un almacén central desde el cual abastece a distintos puntos de venta a nivel regional. El almacén recibe mercadería de múltiples proveedores y atiende pedidos diarios con distintos niveles de prioridad.

En los últimos meses, la gerencia ha identificado problemas recurrentes de quiebres de stock, acumulación excesiva de inventarios en ciertos productos y dificultades para responder oportunamente a la demanda. Estas situaciones han generado sobrecostos operativos, insatisfacción de los clientes internos y pérdida de oportunidades comerciales.

**Descripción del proceso.** El proceso logístico de gestión de inventarios comprende las siguientes etapas:

- a) Abastecimiento
  - Recepción de productos provenientes de proveedores externos.
  - Verificación de cantidades y condiciones de los productos recibidos.
  - Registro de ingresos al almacén.
- b) Almacenamiento
  - Ubicación de los productos en zonas asignadas.
  - Control de niveles de stock por producto.
  - Gestión del espacio disponible y rotación de inventarios.
- c) Despacho
  - Preparación de pedidos según solicitudes recibidas.
  - Verificación de disponibilidad.
  - Registro de salidas de productos.

Cada una de estas etapas genera información operativa que actualmente se registra de forma parcial y con distintos niveles de detalle.

**Situación problemática.** A pesar de contar con registros de entradas y salidas, la empresa presenta las siguientes dificultades:

- Falta de visibilidad en tiempo real de los niveles de inventario.
- Retrasos en la reposición de productos críticos.
- Información dispersa y no estandarizada.

- Dificultad para identificar patrones de consumo y variabilidad de la demanda.

Estas limitaciones impiden una gestión eficiente del inventario y afectan la toma de decisiones tanto operativas como tácticas.

**Propósito del caso de estudio.** El propósito de este caso es analizar la relación entre el proceso logístico, la información generada y los sistemas que la soportan, con el fin de mejorar la gestión de inventarios en el almacén de distribución.

A partir del caso presentado, se desarrollaron las siguientes actividades:

### **Información relevante definida por el proceso**

A partir del análisis del proceso, se identifican como variables críticas:

- Nivel de inventario
- Rotación de productos
- Tiempo de reposición
- Tasa de quiebres de stock
- Costos de almacenamiento

### **Información como reflejo del proceso**

Los registros diarios de entradas y salidas permiten observar:

- Variabilidad en la demanda
- Retrasos recurrentes en proveedores
- Productos con baja rotación
- Sobre stock en determinadas categorías

## Sistema diseñado en función de la relación

El sistema de información debe:

- Registrar transacciones en tiempo real
- Calcular indicadores como stock de seguridad y punto de pedido
- Generar alertas automáticas ante quiebres o excesos



Fig. 2 Relación proceso logístico – indicadores – sistemas de gestión

## 2.3 Ejercicios propuestos

El objetivo de estos ejercicios es comprender y aplicar la relación fundamental entre procesos, información y sistemas, un pilar de la gestión de la información en Ingeniería Industrial. Mediante el análisis de procesos reales de diversos ámbitos organizacionales, se busca desarrollar la capacidad de identificar información clave, interpretar cómo esta refleja el comportamiento del proceso y diseñar sistemas de información alineados con las necesidades operativas y de toma de decisiones.

Más que enfocarse en la resolución numérica inmediata, estos ejercicios promueven el análisis crítico del proceso, el razonamiento ingenieril y la integración de conceptos de gestión, información y tecnología. Su realización fortalece la visión sistémica necesaria para la toma de decisiones y la mejora continua en entornos industriales y organizacionales complejos.

### **Ejercicio 1. Proceso de producción: control de tiempos y eficiencia**

Descripción del proceso. Una planta industrial produce componentes metálicos mediante operaciones secuenciales de corte, mecanizado y ensamblaje. Cada orden de producción atraviesa las mismas etapas, pero se observan diferencias en los tiempos de ejecución entre turnos.

Se pide:

- Identifique las variables del proceso que deben ser medidas.

- Determine qué información es relevante para evaluar la eficiencia.
- Proponga qué tipo de sistema permitiría capturar y analizar dicha información.

### **Ejercicio 2. Proceso de calidad: gestión de productos no conformes**

Descripción del proceso. En una empresa manufacturera, los productos pasan por una inspección final antes de su despacho. Los productos no conformes son registrados y derivados a reproceso o descarte.

Se pide:

- Identifique la información clave que refleja el desempeño del proceso de calidad.
- Analice cómo dicha información permite detectar causas recurrentes de no conformidad.
- Determine los requerimientos básicos de un sistema de información para este proceso.

### **Ejercicio 3. Proceso de mantenimiento: gestión de activos industriales**

Descripción del proceso. Una planta industrial cuenta con equipos críticos que requieren mantenimiento preventivo y correctivo. Las fallas generan paradas no programadas que afectan la producción.

Se pide:

- Identifique las variables críticas del proceso de mantenimiento.
- Defina qué información es necesaria para evaluar la confiabilidad de los equipos.
- Relacione dicha información con decisiones de mantenimiento preventivo.

### **Ejercicio 4. Proceso de seguridad y salud en el trabajo (SSOMA)**

Descripción del proceso. Una empresa industrial gestiona incidentes, accidentes y condiciones inseguras como parte de su sistema de seguridad y salud ocupacional.

Se pide:

- Identifique la información relevante para evaluar el nivel de riesgo.
- Analice cómo la información refleja el comportamiento real del proceso de seguridad.
- Proponga cómo un sistema puede apoyar la prevención de incidentes.

### **Ejercicio 5. Proceso administrativo: gestión de órdenes de servicio**

Descripción del proceso. Una organización presta servicios técnicos y gestiona órdenes que pasan por registro, asignación, ejecución y cierre. Se presentan retrasos y reprocesos administrativos.

Se pide:

- Identifique las variables que afectan el desempeño del proceso.
- Determine qué información permite evaluar tiempos y costos.
- Relacione esta información con decisiones de mejora del proceso.

### **Ejercicio 6. Proceso de gestión pública: atención al ciudadano**

Descripción del proceso. Una entidad pública atiende solicitudes ciudadanas que siguen un flujo definido desde su recepción hasta su resolución final, con plazos normativos establecidos.

Se pide:

- Identifique la información necesaria para controlar el cumplimiento de plazos.
- Analice cómo la información evidencia cuellos de botella.
- Proponga el diseño general de un sistema de seguimiento del proceso.

### **2.4 Tipos de sistemas de información en la organización**

En las organizaciones industriales suelen coexistir distintos tipos de sistemas de información, entre los que destacan:

#### **2.4.1 Sistemas transaccionales**

Orientados al registro estructurado y oportuno de las operaciones diarias de la organización, garantizando la trazabilidad y la integridad de los datos operativos.

Ejemplos: registro de órdenes de producción, movimientos de inventario, partes de mantenimiento, control de asistencia del personal, emisión de comprobantes y atención de solicitudes.

#### **2.4.2 Sistemas de gestión**

Enfocados en el control, seguimiento y evaluación del desempeño de los procesos, mediante indicadores, reportes y mecanismos de monitoreo.

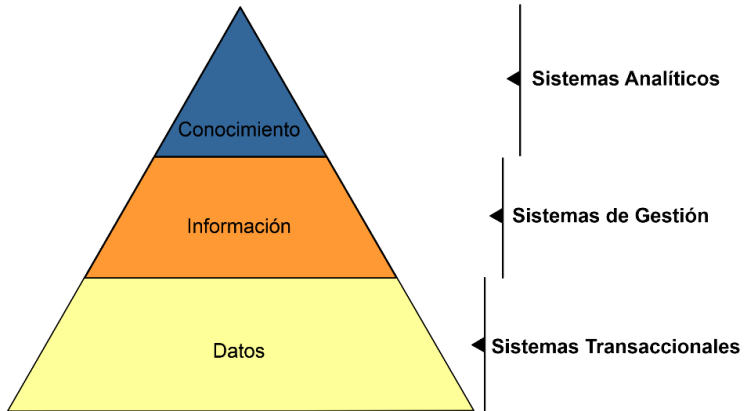
Ejemplos: tableros de control de productividad, seguimiento de indicadores de calidad, control de tiempos de proceso, monitoreo de cumplimiento de planes de mantenimiento o seguridad y salud en el trabajo.

#### **2.4.3 Sistemas analíticos**

Destinados al análisis histórico y avanzado de la información, la identificación de patrones, tendencias y relaciones relevantes, así

como al soporte de la toma de decisiones en los niveles táctico y estratégico.

Ejemplos: análisis de variabilidad de procesos, modelos de pronóstico de demanda, segmentación de clientes o productos.



*Fig. 3 Pirámide del conocimiento y sistemas de información*

La gestión efectiva de la información requiere integrar estos sistemas, evitando silos de datos y asegurando coherencia y trazabilidad.

### **Capítulo 3. Hacia una gestión de la información basada en datos**

La evolución hacia una gestión basada en datos representa un cambio cultural y metodológico en las organizaciones industriales. Este enfoque implica pasar de decisiones fundamentadas en la intuición o la experiencia aislada, hacia decisiones sustentadas en el análisis sistemático de la información.

#### **3.1 La información como base de la toma de decisiones**

La toma de decisiones es una actividad central en la gestión industrial y se manifiesta en distintos niveles de la organización. En todos ellos, la calidad de la decisión depende directamente de la calidad de la información disponible.

##### **3.1.1 Decisiones operativas, tácticas y estratégicas**

Las decisiones operativas se orientan al corto plazo y al control diario de los procesos. Las decisiones tácticas buscan optimizar recursos y planificar actividades en horizontes de mediano plazo. Las decisiones estratégicas definen el rumbo de la organización en el largo plazo.

Cada nivel decisional requiere información con características distintas en términos de:

- Detalle
- Agregación
- Horizonte temporal

- Grado de incertidumbre

Gestionar adecuadamente esta información es una responsabilidad clave del ingeniero industrial.

### **3.1.2 Incertidumbre, riesgo e información**

La información no elimina la incertidumbre, pero permite reducirla y gestionarla. A través del análisis de datos históricos, la identificación de patrones y la evaluación de escenarios, es posible tomar decisiones más informadas y fundamentadas.

En este sentido, la gestión de la información constituye un mecanismo esencial para:

- Evaluar riesgos
- Anticipar desviaciones
- Sustentar decisiones basadas en evidencia

### **3.2 Gestión basada en datos y mejora continua**

La gestión basada en datos se alinea naturalmente con los principios de la mejora continua. El análisis de información permite identificar causas raíz, medir el impacto de las acciones implementadas y evaluar el desempeño de los procesos de manera objetiva.

Este enfoque refuerza el ciclo de mejora:

- Medir
- Analizar

- Actuar
- Controlar



*Fig. 4 Gestión de datos*

### **3.3 Ingeniería de Indicadores**

En las organizaciones modernas, los datos se generan de manera constante como resultado de la ejecución de procesos operativos, administrativos y estratégicos. Sin embargo, la simple acumulación de datos no garantiza una mejora en el desempeño organizacional. Para que la información genere valor, es necesario transformarla en indicadores que permitan medir, analizar y controlar los procesos de manera sistemática.

La ingeniería de indicadores surge como un enfoque estructurado que integra la captura de datos, el modelado de la información, el diseño de indicadores y su representación mediante reportes y dashboards, con el objetivo de apoyar la toma de decisiones en los distintos niveles de la organización. En la Ingeniería Industrial, este enfoque resulta fundamental para alinear la operación diaria con los objetivos estratégicos.

### **3.3.1 Captura y estructuración de los datos**

El primer paso en la construcción de indicadores consiste en la adecuada captura de los datos generados por los procesos organizacionales. Estos datos suelen provenir de diversas fuentes, como sistemas transaccionales, formularios digitales, registros operativos o archivos estructurados.

Desde una perspectiva de ingeniería, la captura de datos debe cumplir con criterios de:

- Consistencia, evitando duplicidades y errores de registro.
- Trazabilidad, permitiendo identificar el origen del dato.
- Oportunidad, asegurando su disponibilidad en el momento requerido.
- Estandarización, mediante catálogos y codificaciones comunes.

Una correcta estructuración de los datos, apoyada en bases de datos bien diseñadas, constituye la base para la generación de información confiable y reutilizable.

### **3.3.2 Del dato al indicador: enfoque de ingeniería**

Transformar datos en indicadores no es un proceso automático, sino una actividad de diseño que requiere criterio ingenieril. Un indicador debe responder a una pregunta concreta relacionada con el desempeño de un proceso, área o sistema.

Desde la ingeniería industrial, un indicador debe cumplir con las siguientes características:

- Ser relevante para la toma de decisiones.
- Ser medible a partir de datos disponibles o factibles de obtener.
- Ser comparable en el tiempo y entre unidades organizacionales.
- Ser comprensible para los usuarios finales.

En este contexto, la ingeniería de indicadores implica definir claramente el objetivo del indicador, su fórmula de cálculo, la periodicidad de medición y los responsables de su análisis.

### **3.3.3 Indicadores operativos, tácticos y estratégicos**

Los indicadores pueden clasificarse según el nivel de decisión al que apoyan:

- **Indicadores operativos**

Se enfocan en el control diario de los procesos y permiten detectar desviaciones en tiempo casi real. Ejemplos incluyen tiempos de proceso, volúmenes producidos o incidencias operativas.

- **Indicadores tácticos**

Apoyan la gestión de áreas o unidades funcionales, permitiendo evaluar tendencias y desempeño en el mediano plazo. Incluyen indicadores de eficiencia, calidad o cumplimiento de metas.

- **Indicadores estratégicos**

Se orientan al logro de los objetivos organizacionales y permiten evaluar el impacto de las decisiones a largo plazo. Suelen estar alineados con la estrategia y los sistemas de control gerencial.

La correcta articulación entre estos niveles permite asegurar coherencia entre la operación y la estrategia.

### **3.3.4 Modelado de indicadores y consistencia analítica**

El modelado de indicadores requiere asegurar la consistencia entre los datos utilizados, las fórmulas aplicadas y el contexto de análisis. En entornos de analítica de datos, esto implica definir correctamente las dimensiones de análisis, los filtros temporales y las relaciones entre variables.

Un indicador mal modelado puede generar interpretaciones erróneas y decisiones inadecuadas. Por ello, el ingeniero industrial debe validar

la lógica de cálculo, analizar la sensibilidad del indicador ante cambios en los datos y verificar su coherencia con la realidad operativa.

### **3.3.5 Visualización y dashboards como herramientas de control**

La visualización de indicadores a través de dashboards transforma los resultados analíticos en herramientas prácticas de gestión. Un dashboard bien diseñado permite:

- Identificar rápidamente desviaciones y tendencias.
- Comparar resultados entre periodos y áreas.
- Priorizar acciones correctivas o preventivas.
- Comunicar información compleja de forma clara y sintética.

Desde el enfoque de la ingeniería industrial, los dashboards deben privilegiar la simplicidad, la claridad visual y la orientación a la acción, evitando la saturación de información.

### **3.3.6 Indicadores y control estratégico**

Los indicadores no deben limitarse al monitoreo operativo, sino integrarse al sistema de control estratégico de la organización. Esto implica vincular los indicadores con objetivos, metas y planes de acción, permitiendo evaluar el avance hacia los resultados esperados.

La ingeniería de indicadores contribuye así a cerrar el ciclo de gestión: medir, analizar, decidir y actuar, fortaleciendo la capacidad de la organización para adaptarse a entornos cambiantes y mejorar su desempeño de manera sostenida.

### **3.3.7 Rol del ingeniero industrial en la ingeniería de indicadores**

El ingeniero industrial cumple un rol central como integrador entre los procesos, la información y la toma de decisiones. Su formación le permite comprender el funcionamiento de los sistemas organizacionales y diseñar indicadores que reflejen la realidad operativa y estratégica.

Más allá de la herramienta tecnológica utilizada, el valor del ingeniero industrial radica en su capacidad para interpretar la información, cuestionar los resultados y proponer mejoras basadas en evidencia.

## Capítulo 4. Arquitectura de información y almacenamiento de datos

La arquitectura de información define cómo se organizan, relacionan y disponibilizan los datos dentro de una organización. Su diseño debe responder a las necesidades de análisis, operación y control propias de la Ingeniería Industrial. Una arquitectura inadecuada limita el uso estratégico de la información, incluso cuando los datos existen.

### 4.1 Arquitecturas de datos en organizaciones industriales

Las organizaciones industriales utilizan diversas arquitecturas de datos, entre ellas las bases de datos operacionales y los repositorios históricos.

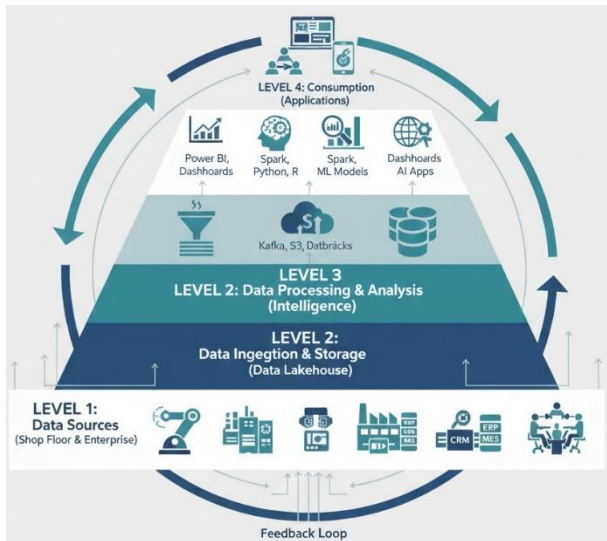


Fig. 5 Arquitectura de datos

### 4.1.1 Bases de datos operacionales

Tal como se desarrolla en el libro “Bases de datos para la Ingeniería Industrial: fundamentos, diseño e implementación para ingenieros industriales”, una base de datos operacional correctamente diseñada trasciende la simple organización de información. Su finalidad es garantizar la consistencia e integridad de los registros, eliminar duplicidades, reducir la dependencia de hojas de cálculo dispersas y facilitar el acceso oportuno a información actualizada.

Desde la perspectiva de la ingeniería industrial, estas bases de datos constituyen el núcleo de los flujos de información operativa de la organización. Su adecuada implementación contribuye al uso eficiente de los recursos —incluida la reducción del uso innecesario de papel— y permite la generación de reportes confiables que ofrecen una visión inmediata del desempeño organizacional. De este modo, se fortalece el control de los procesos, se facilita la detección temprana de desviaciones y se respalda una toma de decisiones fundamentada en información objetiva.

A partir de estas bases de datos operacionales, es posible construir indicadores de desempeño y paneles de control mediante diversas herramientas de analítica y visualización de datos. Entre las más utilizadas en el ámbito organizacional se encuentran hojas de cálculo avanzadas, plataformas de business intelligence, cada una con ventajas

particulares según el contexto, la madurez analítica de la organización y la disponibilidad de recursos.

En el marco de este libro, los ejemplos prácticos y los casos aplicados se desarrollarán utilizando Power BI, por tratarse de una herramienta ampliamente difundida en el entorno empresarial y académico, con capacidades robustas de conexión a múltiples fuentes de datos, modelado de información, creación de indicadores y visualización interactiva. No obstante, los conceptos, metodologías y criterios presentados son independientes de la herramienta, y pueden ser implementados en cualquier plataforma equivalente.

Finalmente, el uso sistemático de indicadores y paneles de control permite evaluar las decisiones adoptadas, generar retroalimentación continua y realizar los ajustes necesarios en los procesos, consolidando un enfoque de mejora continua basado en la gestión eficaz de la información, principio fundamental de la ingeniería industrial.

### **4.1.2 Repositorios históricos**

El resultado inmediato de una base de datos operacional es la construcción de un modelo físico de datos que permite almacenar la información de manera estructurada, consistente y conforme a las reglas del negocio. Sin embargo, desde la perspectiva de la gestión de la información en ingeniería industrial, este no constituye el producto

final del sistema informacional, sino el punto de partida para la generación de conocimiento organizacional.

Los repositorios históricos surgen como una evolución natural de las bases de datos operacionales. Su propósito es conservar la información a lo largo del tiempo, integrando datos provenientes de múltiples fuentes y preservando su contexto histórico, de modo que puedan ser utilizados para análisis comparativos, evaluación de tendencias y soporte a la toma de decisiones estratégicas y tácticas.

A diferencia de las bases de datos operacionales —orientadas al registro eficiente de transacciones—, los repositorios históricos están diseñados para el análisis, priorizando la estabilidad de los datos, la trazabilidad de los cambios y la posibilidad de reconstruir el comportamiento de los procesos en distintos periodos. Esta característica resulta fundamental para el ingeniero industrial, quien requiere comprender no solo el estado actual del sistema, sino también su evolución y desempeño en el tiempo.

En el ámbito organizacional, estos repositorios pueden materializarse en almacenes de datos (data warehouses) o repositorios departamentales (datamarts), cuya estructura responde a los indicadores que se desean analizar y a las decisiones que se busca apoyar. La calidad del repositorio histórico depende directamente de la calidad de las bases de datos operacionales que lo alimentan, así

como de los procesos de integración, limpieza y transformación de los datos.

Desde una perspectiva de mejora continua, los repositorios históricos permiten cerrar el ciclo de gestión de la información: los datos capturados en la operación se transforman en información analítica, esta sustenta decisiones, y los resultados de dichas decisiones retroalimentan el sistema, generando aprendizaje organizacional. En este sentido, los repositorios históricos constituyen un activo estratégico, indispensable para la evaluación del desempeño, la planificación y el rediseño de procesos en la ingeniería industrial.



*Fig. 6 Repositorio de datos históricos*

El enfoque de este libro está orientado al análisis de la información desde la perspectiva del ingeniero industrial, priorizando su aplicación práctica en el soporte a la toma de decisiones. En ese sentido, no se profundizará en el diseño e implementación de arquitecturas

complejas de data warehouse, sino en el uso de estructuras de datos analíticas simplificadas, tales como tablas planas con múltiples dimensiones de análisis, equivalentes en la práctica a repositorios departamentales o datamarts.

Este enfoque permite concentrar el esfuerzo en la comprensión del comportamiento de los procesos, la construcción de indicadores y el análisis del desempeño organizacional, manteniendo la atención en la gestión de la información y no en la complejidad tecnológica. Los conceptos desarrollados son, no obstante, plenamente compatibles con arquitecturas de mayor escala, y pueden ser extendidos o integrados a soluciones de data warehouse en organizaciones con mayor nivel de madurez analítica.

### **4.2 Ingeniería de datos como base de la gestión de la información**

La ingeniería de datos constituye el soporte estructural sobre el cual se construyen la analítica, la ciencia de datos y la inteligencia de decisiones. Sin una adecuada ingeniería de datos, los modelos analíticos carecen de confiabilidad y sostenibilidad en el tiempo.

En la Ingeniería Industrial, la ingeniería de datos debe alinearse con los procesos del negocio, asegurando que la información fluya de manera consistente, trazable y oportuna a lo largo de la organización.

#### 4.2.1 Rol de la ingeniería de datos en entornos industriales

La ingeniería de datos se enfoca en diseñar, construir y mantener las infraestructuras necesarias para capturar, almacenar, transformar y disponibilizar datos de calidad. Su objetivo no es analizar la información, sino garantizar que esta esté preparada para ser analizada.

En contextos industriales, esto implica integrar datos provenientes de:

- **Sistemas transaccionales**, que registran las operaciones diarias de la organización, como producción, inventarios, recursos humanos o ventas, y constituyen la base estructurada para el análisis de procesos.
- **Sensores y dispositivos IoT**, que permiten capturar datos en tiempo real sobre variables físicas y operativas, tales como temperatura, vibración, consumo energético o tiempos de operación, ampliando la capacidad de monitoreo y control de los procesos industriales.



Fig. 7 Sensor de vibración IoT inalámbrico Node 6.

- **Fuentes externas relevantes**, como información normativa, indicadores sectoriales, condiciones ambientales o datos de proveedores y clientes, que permiten contextualizar el desempeño interno y apoyar decisiones estratégicas.

El ingeniero industrial debe comprender estos flujos para asegurar coherencia entre los datos y los procesos que representan.

La integración coherente de estas fuentes habilita una visión integral del sistema productivo, condición necesaria para el diseño de indicadores confiables y el control efectivo de los procesos industriales.

### **4.2.2 Calidad y confiabilidad de los datos**

La calidad de los datos es un requisito fundamental para la toma de decisiones. Datos incompletos, inconsistentes o desactualizados generan diagnósticos erróneos y afectan directamente el desempeño organizacional.

La ingeniería de datos establece mecanismos para:

- Validar datos en origen
- Estandarizar formatos
- Gestionar errores y duplicidades
- Asegurar integridad y consistencia

Una gestión rigurosa de la calidad de los datos reduce riesgos operativos, incrementa la confiabilidad de los sistemas de información y fortalece el proceso de toma de decisiones en la organización.

### **Resumen de la PARTE I. La información como activo estratégico en la Ingeniería Industrial**

La Parte I del libro desarrolla los fundamentos conceptuales de la gestión de la información como un activo estratégico en la Ingeniería Industrial, en un contexto organizacional marcado por la digitalización, la automatización y la creciente complejidad de los procesos. Se plantea que el valor de la información no radica en la acumulación de datos, sino en la capacidad de transformarlos en información confiable que sustente la toma de decisiones.

En este marco, se analiza la evolución desde una información utilizada principalmente con fines operativos hacia una información de carácter estratégico, capaz de anticipar comportamientos, identificar patrones y tendencias, evaluar escenarios alternativos y reducir la incertidumbre en la gestión. Se destaca el rol del ingeniero industrial como gestor de información, responsable de comprender el origen de los datos, evaluar su calidad y relevancia, estructurarlos adecuadamente y utilizarlos como soporte de la mejora continua.

Asimismo, se estudia la relación entre procesos organizacionales y sistemas de información, enfatizando que los procesos definen la información relevante, que la información refleja el comportamiento real de los procesos y que los sistemas deben diseñarse en función de esta relación. A través de un caso de estudio logístico, se evidencia cómo una gestión inadecuada de la información limita el control operativo y la toma de decisiones.

Finalmente, se introduce el enfoque de gestión basada en datos, la ingeniería de indicadores y la arquitectura de información como elementos clave para el análisis del desempeño organizacional. En conjunto, esta parte sienta las bases para una gestión de la información orientada a la toma de decisiones y a la mejora continua, preparando el terreno para el desarrollo de la inteligencia de negocios y el control gerencial en las partes posteriores del libro.

## **PARTE II. Inteligencia de Negocios aplicada a la Ingeniería Industrial**

### **Capítulo 5. Inteligencia de Negocios en contextos industriales**

La creciente complejidad de los entornos industriales exige que las organizaciones transformen grandes volúmenes de datos operativos en información útil para la toma de decisiones. En este contexto, la Inteligencia de Negocios (Business Intelligence, BI) se consolida como una disciplina fundamental para la gestión moderna, al articular procesos, información y tecnología con un enfoque sistemático y orientado al desempeño.

Desde la perspectiva de la Ingeniería Industrial, la Inteligencia de Negocios no se limita al uso de herramientas informáticas, sino que constituye un enfoque de gestión que permite comprender el comportamiento de los procesos, evaluar su desempeño y apoyar la mejora continua.

#### **5.1 BI como disciplina de gestión**

La Inteligencia de Negocios puede definirse como el conjunto de principios, métodos y herramientas orientados a la recopilación, integración, análisis y presentación de información relevante para apoyar la toma de decisiones organizacionales.

En contextos industriales, el valor del BI no reside únicamente en la capacidad de generar reportes o visualizaciones, sino en su contribución a:

- La transparencia de los procesos.
- El alineamiento entre operación y estrategia.
- La reducción de la incertidumbre en la toma de decisiones.
- Desde un enfoque de gestión, el BI se sustenta en tres pilares:
- Procesos bien definidos, que determinan qué información debe ser capturada.
- Indicadores adecuados, que reflejan el desempeño real de dichos procesos.
- Decisiones informadas, basadas en el análisis sistemático de la información.

El ingeniero industrial desempeña un rol clave en este contexto, al actuar como traductor entre la operación, la analítica y la gestión, asegurando que el BI responda a necesidades reales y no a modas tecnológicas.

### **5.2 Relación entre procesos, indicadores y decisiones**

La efectividad de la Inteligencia de Negocios depende directamente de la coherencia entre procesos, indicadores y decisiones. Esta relación puede entenderse como una cadena lógica:

Proceso → Indicador → Información → Decisión



*Fig. 8 Ciclo de valor de la información industrial*

Los procesos industriales generan datos a partir de su ejecución diaria. Sin embargo, estos datos adquieren valor únicamente cuando se transforman en indicadores de desempeño (KPI) que permiten evaluar el cumplimiento de objetivos operativos, tácticos y estratégicos.

Un indicador mal definido o desvinculado del proceso puede conducir a decisiones erróneas, incluso cuando la información esté disponible en tiempo real. Por ello, es responsabilidad del ingeniero industrial:

- Asegurar que cada indicador tenga un propósito claro.
- Verificar que los datos utilizados sean confiables y oportunos.
- Interpretar los resultados considerando el contexto del proceso.

La toma de decisiones apoyada por BI no reemplaza el criterio profesional, sino que lo fortalece, proporcionando evidencia objetiva para priorizar acciones, asignar recursos y evaluar resultados.

### **5.3 BI y mejora continua**

La Inteligencia de Negocios constituye un habilitador clave de la mejora continua, al permitir un ciclo sistemático de análisis y retroalimentación. En contextos industriales, este ciclo puede describirse de la siguiente manera:

- Ejecución del proceso
- Captura y consolidación de datos
- Análisis mediante BI
- Toma de decisiones
- Implementación de mejoras

Este enfoque se alinea naturalmente con metodologías ampliamente utilizadas en la Ingeniería Industrial, como el ciclo PDCA, la gestión por procesos y los enfoques de calidad y excelencia operacional.

El uso adecuado de BI permite:

- Identificar desviaciones y variabilidad en los procesos.
- Detectar tendencias y patrones de comportamiento.
- Evaluar el impacto de las acciones de mejora.
- Promover una cultura organizacional basada en datos.

En este sentido, la Inteligencia de Negocios no debe entenderse como un proyecto aislado, sino como un componente estructural de la gestión industrial, orientado a la sostenibilidad del desempeño y a la toma de decisiones informadas en todos los niveles de la organización.

### **Capítulo 6. Diseño de reportes y dashboards para control gerencial**

En los entornos industriales actuales, la disponibilidad de información no garantiza una mejor gestión. La diferencia entre una organización reactiva y una organización orientada al desempeño radica en su capacidad para presentar, interpretar y utilizar la información de manera efectiva. En este contexto, los reportes y dashboards constituyen instrumentos clave para el control gerencial, al sintetizar el estado de los procesos y facilitar la toma de decisiones oportunas.

Desde la Ingeniería Industrial, el diseño de reportes y dashboards debe responder a una lógica clara: mostrar lo esencial, en el momento adecuado y para el nivel decisonal correcto.

#### **6.1 Reportes operativos vs. reportes gerenciales**

Los reportes pueden clasificarse, de manera general, en operativos y gerenciales, según su propósito y nivel de uso dentro de la organización.

Los reportes operativos están orientados al seguimiento detallado de la ejecución diaria de los procesos. Su principal función es apoyar la supervisión y el control inmediato de las operaciones. Estos reportes suelen:

- Presentar información granular y detallada.
- Utilizar datos en tiempo real o de alta frecuencia.
- Estar dirigidos a supervisores y responsables operativos.

Por su parte, los reportes gerenciales tienen un enfoque sintético y estratégico. Su objetivo es proporcionar una visión consolidada del desempeño, facilitando la evaluación de resultados y la toma de decisiones de carácter táctico o estratégico. Estos reportes se caracterizan por:

- Presentar indicadores clave de desempeño.
- Enfatizar tendencias, comparaciones y desviaciones.
- Estar dirigidos a jefaturas, gerencias y alta dirección.

El ingeniero industrial debe asegurar la coherencia entre ambos tipos de reportes, evitando duplicidad de información y garantizando que los reportes gerenciales se sustenten en datos confiables provenientes de la operación.

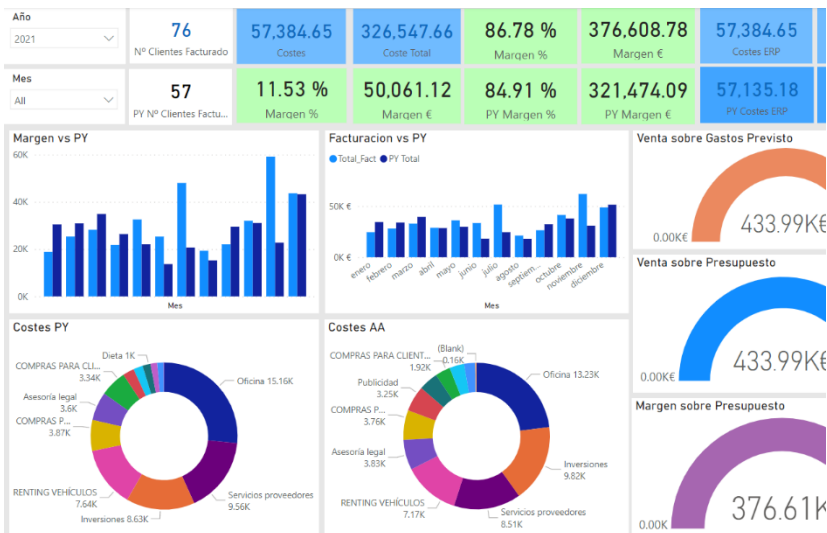


Fig. 9 Dashboard gerencial

### 6.2 Principios de visualización ejecutiva

La visualización de información en contextos gerenciales debe priorizar la claridad, relevancia y simplicidad. Un dashboard mal diseñado puede generar confusión, incluso cuando los datos sean correctos.

Entre los principales principios de visualización ejecutiva destacan:

Enfoque en indicadores clave: mostrar únicamente las métricas que impactan en los objetivos del proceso.

- Jerarquía visual: resaltar lo más importante mediante tamaño, posición y contraste.
- Consistencia: utilizar escalas, colores y formatos uniformes para facilitar la interpretación.
- Comparabilidad: permitir contrastar resultados con metas, periodos anteriores o referencias.
- Simplicidad: evitar sobrecargar el dashboard con gráficos innecesarios.

El objetivo de la visualización ejecutiva no es mostrar todos los datos disponibles, sino facilitar una lectura rápida y significativa del desempeño.

### 6.3 Semáforos de desempeño

Los semáforos de desempeño constituyen uno de los mecanismos más utilizados en dashboards gerenciales por su capacidad de comunicar el estado de un indicador de manera inmediata.

Este enfoque clasifica el desempeño en categorías generalmente representadas por colores:

- Verde: desempeño dentro de los parámetros esperados
- Amarillo: desempeño en situación de alerta o riesgo
- Rojo: desempeño fuera de los límites aceptables

El uso de semáforos debe estar sustentado en:

- Metas claramente definidas
- Umbrales realistas y alineados a la estrategia
- Indicadores directamente vinculados a procesos críticos

Un uso inadecuado de semáforos —por ejemplo, sin criterios claros o con excesivos indicadores en estado rojo— puede generar desensibilización y pérdida de efectividad en el control gerencial.



Fig. 10 Semáforos de desempeño

#### 6.4 Interpretación y acción

El valor de un reporte o dashboard no reside en su diseño visual, sino en su capacidad para provocar interpretación y acción. La información presentada debe responder a preguntas clave como:

- ¿Qué está ocurriendo en el proceso?
- ¿Por qué está ocurriendo?
- ¿Qué acciones deben tomarse?

Desde la perspectiva de la Ingeniería Industrial, la interpretación de la información debe considerar:

- El contexto operativo del proceso
- La variabilidad inherente a las operaciones
- La relación entre indicadores y causas subyacentes

Un dashboard efectivo no solo muestra resultados, sino que orienta la gestión, facilitando la priorización de acciones, el seguimiento de mejoras y la evaluación del impacto de las decisiones tomadas.

## **Capítulo 7. Caso de estudio en Inteligencia de Negocios.**

### **7.1 Contexto del caso**

La seguridad y salud en el trabajo constituye uno de los ejes fundamentales de la gestión industrial, particularmente en el sector minero, donde los riesgos operativos son elevados y las consecuencias de los accidentes laborales pueden ser significativas tanto en términos humanos como económicos. En este contexto, la Minera ZET (caso aplicado de carácter ilustrativo) registra información relacionada con accidentes laborales como parte de sus obligaciones operativas y normativas.

Sin embargo, como ocurre en muchas organizaciones, la información generada durante la operación diaria se encuentra inicialmente dispersa en registros transaccionales, formularios, reportes individuales y archivos electrónicos, lo que dificulta su análisis sistemático. Esta situación limita la capacidad de identificar patrones, causas recurrentes y tendencias que permitan diseñar acciones preventivas efectivas.

Desde la perspectiva de la ingeniería industrial, el problema no radica en la ausencia de datos, sino en la gestión inadecuada de la información para apoyar la toma de decisiones.

### **7.2 Bases de datos operacionales en SSOMA**

La Minera ZET cuenta con bases de datos operacionales orientadas al registro de accidentes laborales. Estas bases de datos almacenan información como fechas, trabajadores involucrados, áreas, descripciones del evento y consecuencias inmediatas. Su diseño responde principalmente a necesidades operativas y de cumplimiento normativo, priorizando el registro oportuno de los hechos.

No obstante, estas bases de datos presentan características propias de los sistemas transaccionales:

- Están orientadas al registro del evento, no al análisis.
- Presentan información altamente normalizada.
- No conservan de forma explícita la evolución histórica de ciertos atributos.
- No están diseñadas para el cálculo directo de indicadores.

Si bien estas bases de datos cumplen adecuadamente su función operativa, no permiten por sí solas responder preguntas clave, tales como:

- ¿Cuáles son las principales causas de los accidentes?
- ¿En qué áreas se concentran?
- ¿Qué condiciones o actos inseguros son recurrentes?

### **7.3 Repositorios históricos para el análisis de accidentes**

Para superar estas limitaciones, se plantea la construcción de un repositorio histórico de accidentes laborales, alimentado a partir de las

bases de datos operacionales. Este repositorio consolida la información de múltiples periodos y fuentes, preservando el contexto temporal y operativo de cada evento.

El repositorio histórico permite:

- Analizar la evolución de los accidentes en el tiempo.
- Comparar periodos, áreas y condiciones de trabajo.
- Evaluar el impacto de decisiones preventivas.

Generar conocimiento organizacional a partir de la experiencia acumulada.

Desde el enfoque de este libro, dicho repositorio adopta la forma de una tabla plana analítica, suficiente para soportar indicadores y paneles de control, sin requerir una arquitectura compleja de data warehouse.

### **7.4 Análisis de la base de datos**

Se analizó el modelo de datos transaccional de accidentes laborales presentado (Fig. 11), concluyéndose que se trata de un diseño sólido, coherente y apropiado para un sistema SSOMA de carácter operativo, especialmente en contextos industriales de alto riesgo como el sector minero. El modelo refleja claramente una orientación al registro exhaustivo y sistemático de los eventos, alineada con las necesidades de control interno y con los requerimientos de cumplimiento normativo.

Asimismo, la centralidad de la tabla accidente como eje del modelo resulta correcta y natural. Esta entidad captura de manera estructurada el evento operativo, permitiendo un registro detallado, consistente y confiable de cada incidente ocurrido.

No obstante, como es habitual en los sistemas transaccionales bien diseñados, el modelo no está orientado al análisis gerencial ni a la toma de decisiones estratégicas. La información se encuentra distribuida en múltiples tablas relacionadas, lo que dificulta:

- El análisis agregado de datos históricos
- La construcción directa de indicadores
- La identificación rápida de patrones y tendencias
- La visualización ejecutiva del desempeño en SSOMA

Es importante destacar que esta limitación no constituye una debilidad del modelo, sino que confirma que cumple adecuadamente su propósito transaccional. Pretender que un sistema operativo resuelva necesidades analíticas complejas iría en contra de las buenas prácticas de diseño de sistemas de información.

En síntesis, el modelo de datos presentado puede calificarse como:

- Adecuado y bien estructurado para la operación diaria
- Consistente con los procesos reales de SSOMA

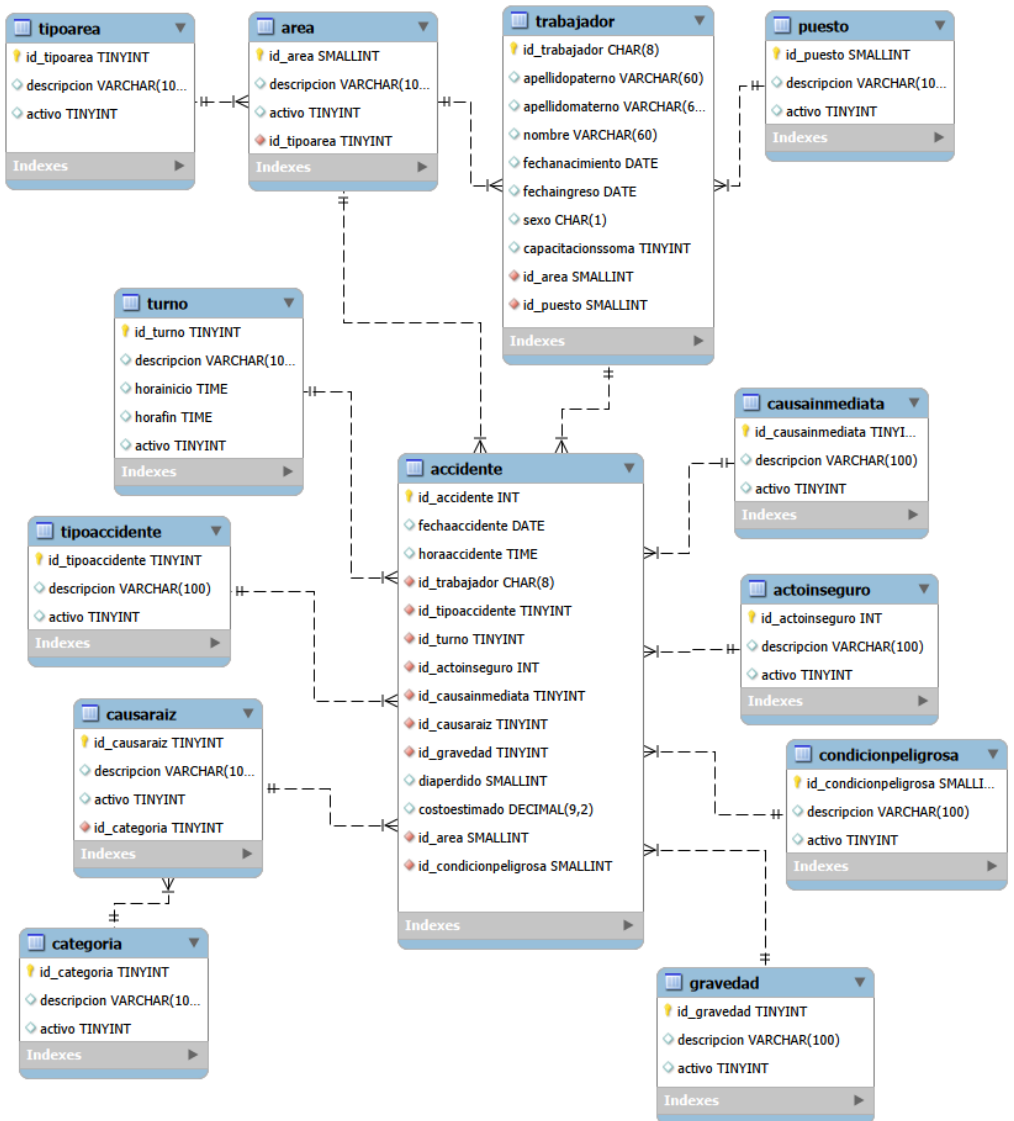


Fig. 11 Base de datos de gestión de accidentes de la minera

### 7.5 Tabla plana propuesta

La tabla plana de datos propuesta se estructura como una única tabla analítica, en la que se integran las dimensiones previamente seleccionadas junto con las métricas relevantes del proceso. Esta estructura se encuentra optimizada para el análisis en herramientas de inteligencia de negocios, tales como Power BI, facilitando la creación de indicadores y paneles de control de manera eficiente.

La generación de esta tabla puede realizarse mediante un procedimiento almacenado, desarrollado por el área de tecnologías de la información a partir de los requerimientos definidos por el ingeniero industrial. El resultado de dicho procedimiento almacenado es posteriormente exportado a un archivo en formato Excel, el cual actúa como repositorio intermedio de información analítica.

Este archivo es utilizado como fuente de datos en Power BI para el desarrollo de los análisis, permitiendo una separación clara entre la capa de generación de datos y la capa de análisis y visualización. Al consolidar en una sola estructura las distintas dimensiones del accidente laboral, la tabla plana posibilita un análisis multidimensional del evento, facilitando la exploración de relaciones entre causas, condiciones, áreas, puestos y periodos de tiempo, aspecto fundamental para la ingeniería industrial orientada a la mejora de los procesos de seguridad y salud en el trabajo.



The image shows a screenshot of a database table definition for a table named 'accidente'. The table has 21 columns with the following data types: id\_accidente (INT), trabajador (VARCHAR(100)), sexo (VARCHAR(20)), edad (TINYINT), tiempoempresa (DECIMAL(9,2)), capacitacionssoma (CHAR(2)), areaaccidente (VARCHAR(100)), areatrabajador (VARCHAR(100)), puesto (VARCHAR(100)), turno (VARCHAR(100)), causainmediata (VARCHAR(100)), tipoaccidente (VARCHAR(100)), actoinseguro (VARCHAR(100)), causaraiz (VARCHAR(100)), categoriacausa (VARCHAR(100)), condicionpeligrosa (VARCHAR(100)), gravedad (VARCHAR(100)), fechaaccidente (DATE), horaaccidente (TIME), diaperdido (SMALLINT), costoestimado (DECIMAL(9,2)), anio (CHAR(4)), and mes (VARCHAR(20)). At the bottom of the screenshot, there is a section labeled 'Indexes' with a right-pointing arrow.

Column Name	Data Type
id_accidente	INT
trabajador	VARCHAR(100)
sexo	VARCHAR(20)
edad	TINYINT
tiempoempresa	DECIMAL(9,2)
capitacionssoma	CHAR(2)
areaaccidente	VARCHAR(100)
areatrabajador	VARCHAR(100)
puesto	VARCHAR(100)
turno	VARCHAR(100)
causainmediata	VARCHAR(100)
tipoaccidente	VARCHAR(100)
actoinseguro	VARCHAR(100)
causaraiz	VARCHAR(100)
categoriacausa	VARCHAR(100)
condicionpeligrosa	VARCHAR(100)
gravedad	VARCHAR(100)
fechaaccidente	DATE
horaaccidente	TIME
diaperdido	SMALLINT
costoestimado	DECIMAL(9,2)
anio	CHAR(4)
mes	VARCHAR(20)

Fig. 12 Tabla plana propuesta para la analítica de datos

El resultado del proceso de generación de datos es exportado a un archivo en formato Excel, que funciona como repositorio de información para su posterior análisis, tal como se aprecia en la siguiente figura.

id	nombre	sexo	edad	capacitacion	sumo	tiempo	empresa	area	accidente	area	trabajador	puesto	turno	causal	mediata	tipo	acciden
1	Vargas Quispe, Pedro	Masculino	42	SI				12.3	Planta de Chancado	Mina Subterránea	Operador	Noche	Falta de procedimiento				Golpe
2	Quispe Flores, María	Femenino	45	NO				4.8	Taller Mecánico	Planta de Chancado	Técnico	Noche	Falta de procedimiento				Atrapamien
3	Gómez Pérez, Miguel	Masculino	53	SI				13.4	Almacén	Planta Concentradora	Técnico	Día	Falta de capacitación				Caída
4	Pérez Rojas, Carmen	Femenino	28	SI				17.4	Planta Concentradora	Taller Mecánico	Supervisor	Rotativo	Procedimiento inadecuado				Exposición i
5	Ramírez Pérez, Rosa	Femenino	50	SI				21.4	Planta de Chancado	Planta Concentradora	Supervisor	Rotativo	Exceso de confianza				Caída
6	Ramírez Ramírez, Carmen	Femenino	41	SI				24.8	Planta de Chancado	Planta Concentradora	Técnico	Rotativo	Falta de procedimiento				Corte
7	Quispe Vargas, Miguel	Masculino	40	SI				18.3	Almacén	Mina Subterránea	Supervisor	Noche	Procedimiento inadecuado				Caída
8	Quispe Vargas, Juan	Masculino	30	SI				11.7	Planta Concentradora	Planta Concentradora	Técnico	Rotativo	Falta de capacitación				Atrapamien
9	Torres Vargas, Ana	Femenino	34	NO				15.4	Planta Concentradora	Taller Mecánico	Operador	Noche	Procedimiento inadecuado				Golpe
10	Ramírez Pérez, Pedro	Masculino	55	NO				6.7	Planta de Chancado	Taller Mecánico	Supervisor	Noche	Falta de procedimiento				Exposición i
11	Gómez Flores, María	Femenino	43	SI				4.4	Planta Concentradora	Taller Mecánico	Mecánico	Día	Falta de capacitación				Corte
12	Pérez Flores, Rosa	Femenino	58	SI				19.4	Mina Subterránea	Almacén	Operador	Día	Falta de procedimiento				Corte
13	Flores Flores, Juan	Masculino	29	SI				1.7	Planta Concentradora	Taller Mecánico	Técnico	Noche	Falta de capacitación				Caída
14	Gómez Flores, Lucía	Femenino	30	NO				18.9	Almacén	Mina Subterránea	Operador	Rotativo	Procedimiento inadecuado				Corte
15	Torres Gómez, Carmen	Femenino	35	SI				17.2	Taller Mecánico	Mina Subterránea	Supervisor	Día	Exceso de confianza				Exposición i
16	Ramírez Vargas, María	Femenino	40	SI				1.2	Planta de Chancado	Almacén	Mecánico	Día	Exceso de confianza				Golpe
17	Quispe Flores, Ana	Femenino	46	SI				7.9	Almacén	Planta Concentradora	Mecánico	Noche	Procedimiento inadecuado				Caída
18	Vargas Flores, Carlos	Masculino	38	SI				6	Mina Subterránea	Almacén	Operador	Día	Falta de capacitación				Atrapamien
19	Torres Flores, Miguel	Masculino	25	NO				6.5	Planta Concentradora	Mina Subterránea	Supervisor	Día	Procedimiento inadecuado				Exposición i

Fig. 13 Excel con 2000 datos para el análisis del Power BI

## 7.6 Construcción de indicadores y reportes de alto impacto

A partir del repositorio histórico de datos, se construyen indicadores clave de desempeño (KPI) que permiten evaluar de manera objetiva el comportamiento de la organización en materia de seguridad y salud en el trabajo. Estos indicadores constituyen la base analítica para el control operativo, la gestión preventiva y la toma de decisiones a nivel táctico y estratégico.

Entre los indicadores más relevantes se consideran:

- Índice de Frecuencia (IF). Mide qué tan seguido ocurren accidentes respecto al tiempo trabajado.

$$IF = \frac{\text{N}^\circ \text{ de accidentes} \times 1,000,000}{\text{Horas} - \text{hombre trabajadas}}$$

- Índice de Severidad (IS). Mide qué tan graves son los accidentes.

$$IS = \frac{\text{Días perdidos} \times 1,000,000}{\text{Horas} - \text{hombre trabajadas}}$$

- Índice de Gravedad (IG). Combina ambos:

$$IG = \frac{IS}{IF}$$

Sin embargo, el valor de la información no se limita a la generación aislada de indicadores. Estos KPI deben integrarse en reportes de alto impacto y dashboards interactivos, los cuales permiten:

- Visualizar tendencias y patrones de comportamiento en el tiempo.
- Identificar áreas críticas y procesos con mayor exposición al riesgo.
- Comparar el desempeño entre periodos, áreas o grupos de trabajadores.
- Evaluar la efectividad de las acciones preventivas y programas de capacitación en SSOMA.
- Facilitar la comunicación de resultados a los distintos niveles de la organización, desde el nivel operativo hasta la alta dirección.

En este contexto, los dashboards se convierten en herramientas fundamentales para la gestión visual de la seguridad, ya que transforman grandes volúmenes de datos históricos en información clara, comprensible y accionable, apoyando una gestión preventiva basada en evidencia y fortaleciendo la cultura de seguridad dentro de la organización.

### **7.7 Uso de Power BI como herramienta de análisis**

Para los ejemplos desarrollados en este libro, se utiliza Power BI como herramienta de analítica y visualización. Power BI permite:

- Conectar el repositorio histórico de datos.
- Modelar las relaciones entre tablas.
- Crear indicadores mediante expresiones analíticas.
- Diseñar paneles interactivos para la exploración de la información.

Es importante destacar que el enfoque metodológico presentado es independiente de la herramienta, y puede ser implementado en cualquier plataforma equivalente. Power BI se utiliza por su amplia adopción y facilidad de uso en contextos académicos y organizacionales.

Como referencia, enfoques similares pueden desarrollarse utilizando otras herramientas de inteligencia de negocios, tales como Tableau, Qlik Sense, Looker, SAP Analytics Cloud u otras plataformas BI

disponibles en el mercado, manteniendo los mismos principios de gestión de la información, análisis de datos y soporte a la toma de decisiones.

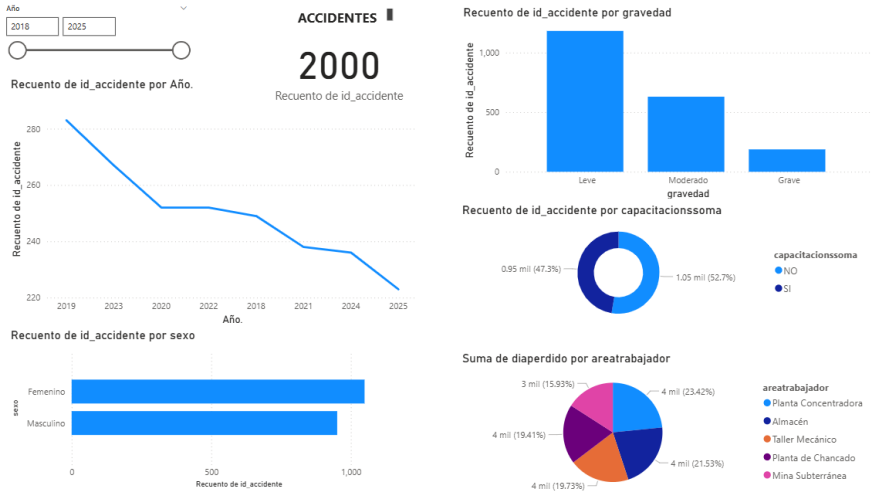


Fig. 14 Dashboard de impacto para el análisis de accidentes

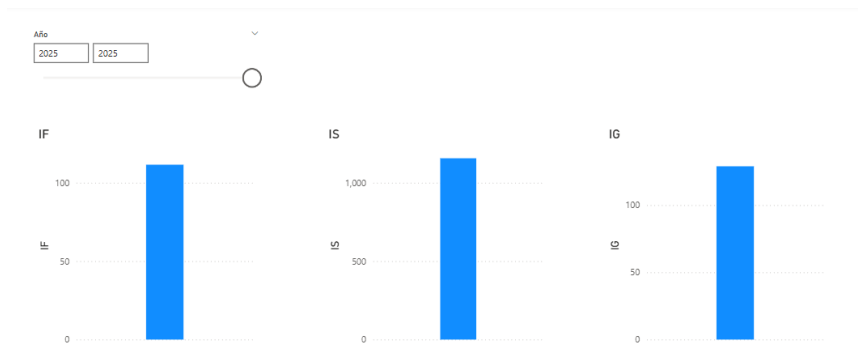


Fig. 15 Indicadores para el análisis de accidentes

### **7.8 Análisis de causas de los accidentes laborales**

Mediante los paneles de control construidos, es posible identificar que un porcentaje significativo de los accidentes se concentra en determinadas áreas operativas y está asociado principalmente a causas básicas de tipo organizacional y humano, tales como deficiencias en la capacitación, supervisión insuficiente y fatiga laboral.

Este análisis permite al ingeniero industrial ir más allá del evento puntual, enfocándose en las causas raíz y no únicamente en las consecuencias visibles del accidente.

### **7.9 Retroalimentación y mejora continua**

El uso sistemático del repositorio histórico y de los indicadores generados permite cerrar el ciclo de gestión de la información. Las decisiones adoptadas a partir del análisis pueden ser evaluadas en periodos posteriores, verificando su impacto real en la reducción de accidentes y en la mejora de las condiciones de trabajo.

De este modo, la información deja de ser un simple registro histórico para convertirse en un insumo estratégico para la mejora continua, alineado con los principios de la ingeniería industrial.

### Ejercicios propuestos

#### **Caso de estudio 1: Evaluación de la efectividad de un programa de capacitación**

Contexto. Una empresa industrial ha implementado un programa anual de capacitación técnica y operativa para su personal, pero no cuenta con información clara sobre su impacto en el desempeño laboral.

Problema. La gerencia desconoce si las horas de capacitación se traducen en mejoras reales en productividad, reducción de errores o disminución de incidentes operativos.

Objetivo del caso. Analizar la información histórica de capacitación y desempeño para evaluar la efectividad del programa.

- Información disponible
- Horas de capacitación por trabajador
- Tipo de capacitación
- Indicadores de desempeño antes y después
- Área de trabajo y puesto

Entregable

- Dashboard de análisis de capacitación

### **Caso de Estudio 2: Análisis de rotación de personal y sus causas**

Contexto. Una empresa manufacturera presenta altos niveles de rotación de personal, lo que genera costos operativos y pérdida de conocimiento organizacional.

Problema. No se identifican claramente las causas principales de la rotación ni los perfiles más afectados.

Objetivo del caso. Identificar patrones y factores asociados a la rotación del personal mediante análisis de información.

Información disponible

- Fecha de ingreso y salida
- Área y puesto
- Motivo de salida
- Antigüedad y nivel de capacitación

Entregable

- Reporte analítico de rotación

### **Caso de Estudio 3: Control de calidad y reducción de reprocesos**

Contexto. Una planta de producción registra altos niveles de reprocesos y productos defectuosos, afectando los costos y los tiempos de entrega.

Problema. La información de calidad se encuentra dispersa y no permite identificar las causas raíz.

Objetivo del caso. Analizar datos de producción y calidad para reducir reprocesos y defectos.

Información disponible

- Tipo de defecto
- Línea de producción
- Turno de trabajo
- Volumen producido y rechazado

Entregable

- Dashboard de calidad

### **Caso de Estudio 4: Análisis de eficiencia operativa y tiempos de proceso**

Contexto. Una empresa de servicios desea mejorar la eficiencia de sus procesos internos, pero no cuenta con indicadores claros de tiempos y cuellos de botella.

Problema. Las decisiones de mejora se toman de manera intuitiva, sin respaldo de información.

Objetivo del caso. Analizar tiempos de proceso y niveles de servicio para identificar oportunidades de mejora.

Información disponible

- Tiempo estándar vs. tiempo real
- Tipo de proceso

- Área responsable
- Volumen de atención

Entregable

- Indicadores de eficiencia y recomendaciones

### **Caso de Estudio 5: Análisis del ausentismo laboral y su impacto en la operación**

Contexto. El ausentismo laboral ha aumentado de manera sostenida en una empresa industrial, afectando la productividad y la planificación operativa.

Problema. No se conoce con claridad qué factores influyen en el ausentismo ni su impacto real.

Objetivo del caso. Evaluar el ausentismo laboral mediante análisis de información histórica.

Información disponible

- Días de ausencia
- Tipo de ausencia
- Área y turno
- Relación con carga laboral y capacitación

Entregable

- Dashboard de ausentismo

### **Resumen de la Parte II. Sistemas de información en la ingeniería industrial**

La Parte II del libro desarrolla la aplicación de la Inteligencia de Negocios como un enfoque de gestión orientado al análisis del desempeño y a la toma de decisiones en contextos industriales. Se plantea que, frente a la creciente complejidad de los procesos y al aumento del volumen de datos operativos, el desafío principal de las organizaciones no es la generación de información, sino su adecuada gestión, interpretación y uso estratégico.

Desde la perspectiva de la Ingeniería Industrial, la Inteligencia de Negocios se aborda como una disciplina integradora que articula procesos, indicadores y decisiones. Se enfatiza que el valor del BI no reside únicamente en las herramientas tecnológicas, sino en la capacidad de transformar datos operativos en información relevante que permita comprender el comportamiento de los procesos, reducir la incertidumbre y apoyar la mejora continua. En este marco, se destaca el rol del ingeniero industrial como nexo entre la operación, la analítica y la gestión, asegurando la coherencia entre los indicadores definidos y los objetivos organizacionales.

Asimismo, se desarrolla el diseño de reportes y dashboards como instrumentos clave para el control gerencial, diferenciando claramente entre reportes operativos y gerenciales. Se presentan principios de visualización ejecutiva orientados a la claridad, la síntesis y la acción,

incorporando el uso de semáforos de desempeño como mecanismos de comunicación rápida del estado de los procesos.

La parte culmina con un caso de estudio aplicado al ámbito de la seguridad y salud en el trabajo, en el cual se ilustra el tránsito desde bases de datos transaccionales hacia repositorios analíticos, la construcción de indicadores y el desarrollo de dashboards para el análisis de accidentes laborales. Este caso evidencia cómo la Inteligencia de Negocios permite convertir la información histórica en un insumo estratégico para la prevención, el control gerencial y la mejora continua, consolidando una gestión industrial basada en datos y evidencia objetiva.

## **PARTE III. Estadística e inferencia estadística para la toma de decisiones**

### **Capítulo 8. Estadística aplicada a la ingeniería industrial**

La estadística no es solo una rama de las matemáticas para el ingeniero industrial; es el lenguaje de la calidad y la eficiencia. Permite transformar el ruido de los datos operativos en señales claras para la gestión. En la era de la Industria 4.0, donde los sensores generan flujos masivos de datos, la capacidad de identificar la variabilidad y respaldar decisiones con evidencia científica es lo que diferencia una gestión reactiva de una proactiva.

Este capítulo aborda los fundamentos de la estadística descriptiva, el análisis de la variabilidad y la construcción de indicadores, integrando ejemplos prácticos en R, el estándar de oro para la ciencia de datos. El objetivo no es solo aprender a calcular medidas, sino a interpretarlas en el contexto de los procesos industriales y utilizarlas como base para la mejora continua.

#### **8.1 Estadística descriptiva**

La estadística descriptiva busca resumir y caracterizar un conjunto de datos para comprender su estructura. En ingeniería, esto es vital para auditar tiempos de ciclo, niveles de inventario, consumos energéticos, o cualquier variable que refleje el comportamiento de un proceso.

### 8.1.1 Medidas de tendencia central

Localizan el "centro" de la distribución de los datos. Las más utilizadas son:

- Media muestral ( $\bar{x}$ ): Es el promedio aritmético de los datos. Se calcula sumando todos los valores y dividiendo entre el número de observaciones. Es sensible a valores extremos (atípicos).

Fórmula:

$$\bar{x} = \frac{\sum x_i}{n}$$

- Mediana ( $\tilde{x}$ ): Es el valor que divide el conjunto de datos ordenados en dos partes iguales (50% por debajo, 50% por encima). Es robusta frente a valores atípicos.
- Moda: Es el valor que aparece con mayor frecuencia. Útil para identificar la falla más recurrente, el producto más vendido, etc.

#### Ejemplo en R: Tiempos de atención

Supongamos que se registran los tiempos de atención (en minutos) de 10 servicios técnicos:

##### Código en R

```
tiempos <- c(18, 20, 22, 19, 21, 23, 20, 19, 22, 24)
```

```
#Calculamos las medidas:
media <- mean(tiempos)
mediana <- median(tiempos)
#Para la moda, creamos una tabla de frecuencias
tabla <- table(tiempos)
moda <- as.numeric(names(tabla[tabla == max(tabla)]))
```

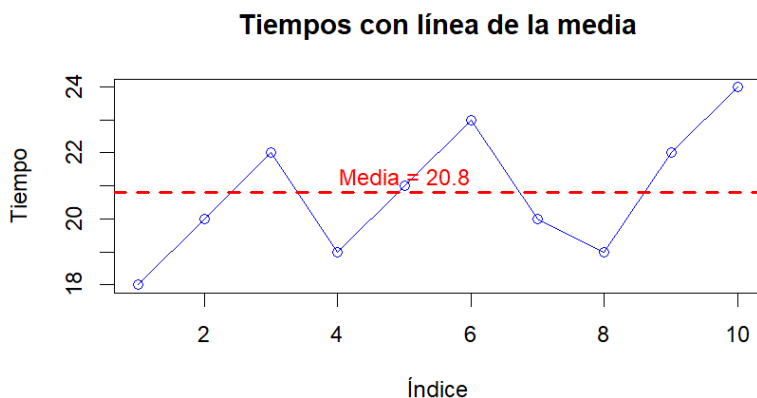
## Resultados:

- Media = 20.8 minutos
- Mediana = 20.5 minutos
- Moda = 19, 20 y 22 (tres valores con frecuencia 2)

Interpretación: La media de 20.8 minutos indica el tiempo promedio de atención. La mediana muy cercana sugiere que la distribución no está fuertemente sesgada. La presencia de múltiples modas puede indicar que existen varios tiempos típicos, quizás asociados a diferentes tipos de servicios.

## Código en R

```
#Graficar la media
plot(tiempos, type = "o", col = "blue",
     main = "Tiempos con línea de la media",
     xlab = "Índice", ylab = "Tiempo")
# Añadir la línea de la media
abline(h = media, col = "red", lwd = 2, lty = 2)
# Mostrar el valor de la media en el gráfico
text(x = length(tiempos)/2, y = media + 0.5,
     labels = paste("Media =", round(media, 2)),
     col = "red")
```



*Fig. 16 Media del ejemplo de tiempos de atención*

**Código en R**

```
#Graficar la mediana
# Histograma
hist(tiempos,
      main = "Histograma de tiempos",
      xlab = "Tiempos",
      col = "lightblue",
      border = "black")

# Línea de la mediana
abline(v = median(tiempos), col = "red", lwd = 2)

# Texto indicando la mediana
text(median(tiempos),      2,      paste("Mediana =",
median(tiempos)),
      col = "red", pos = 4)
```

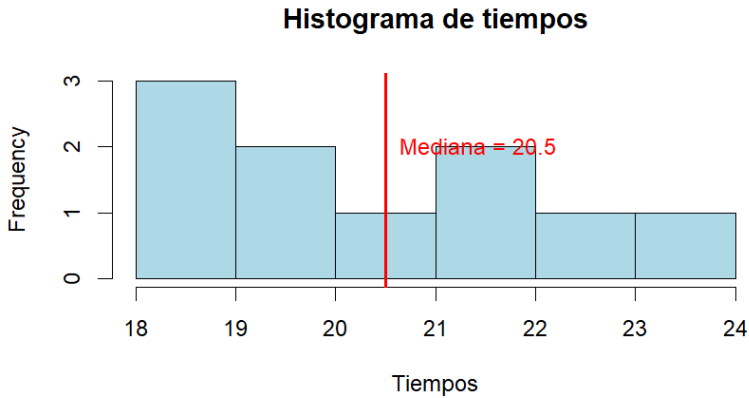


Fig. 17 Mediana del ejemplo de tiempos de atención

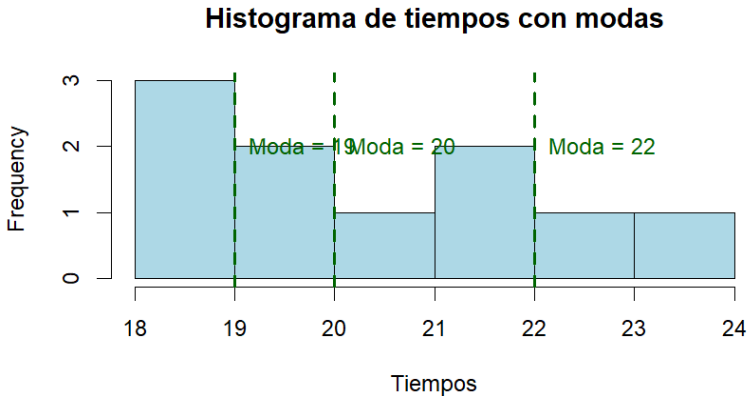
**Código en R**

```
#Graficar la moda
# Histograma
hist(tiempos,
     main = "Histograma de tiempos con modas",
     xlab = "Tiempos",
     col = "lightblue",
     border = "black")

# Calcular modas
tabla <- table(tiempos)
moda <- as.numeric(names(tabla[tabla == max(tabla)]))

# Añadir líneas en las modas
abline(v = moda, col = "darkgreen", lwd = 2, lty = 2)

# Texto indicando las modas
text(moda, 2, paste("Moda =", moda),
     col = "darkgreen", pos = 4)
```



*Fig. 18 Moda del ejemplo de tiempos de atención*

### 8.1.2 Medidas de dispersión

Indican qué tan "estirada" o concentrada está la distribución. Las principales son:

- **Rango:** Diferencia entre el valor máximo y el mínimo. Da una idea rápida de la extensión de los datos, pero es muy sensible a valores extremos.
- **Varianza muestral ( $s^2$ ):** Promedio de las desviaciones al cuadrado respecto a la media. Mide la dispersión en unidades al cuadrado, lo que dificulta la interpretación directa.

Fórmula:

$$s^2 = \sum(x_i - \bar{x})^2 / (n - 1)$$

- Desviación estándar muestral (s): Raíz cuadrada de la varianza. Se expresa en las mismas unidades que los datos, por lo que es la medida de dispersión más intuitiva.
- Coeficiente de variación (CV): Relación entre la desviación estándar y la media ( $CV = s / \bar{x}$ ). Permite comparar la variabilidad relativa entre procesos con diferentes unidades o magnitudes.

Ejemplo en R (continuación):

### Código en R

```
rango <- max(tiempos) - min(tiempos)
varianza <- var(tiempos)
desviacion <- sd(tiempos)
cv <- desviacion / media
```

Resultados:

Rango = 24 - 18 = 6 minutos

Varianza = 3.73 min<sup>2</sup>

Desviación estándar = 1.93 minutos

Coeficiente de variación = 1.93 / 20.8  $\approx$  0.093 (9.3%)

Interpretación: Una desviación estándar de 1.93 minutos significa que, en promedio, los tiempos se desvían casi 2 minutos respecto a la media. En un proceso de atención al cliente, esta variabilidad puede

traducirse en incertidumbre en la planificación de recursos. El CV del 9.3% indica una variabilidad moderada; si superara el 15% podríamos considerarlo como un proceso poco estable.

### 8.1.3 Visualización de la distribución

Antes de cualquier análisis numérico, es fundamental explorar los datos gráficamente. Las herramientas más comunes son:

- Histograma: Muestra la forma de la distribución (simétrica, sesgada, multimodal).
- Diagrama de caja (boxplot): Resume la mediana, cuartiles y valores atípicos.
- Gráfico de puntos o de líneas: Útil para series temporales.

Ejemplo en R:

#### Código en R

```
summary(tiempos)
```

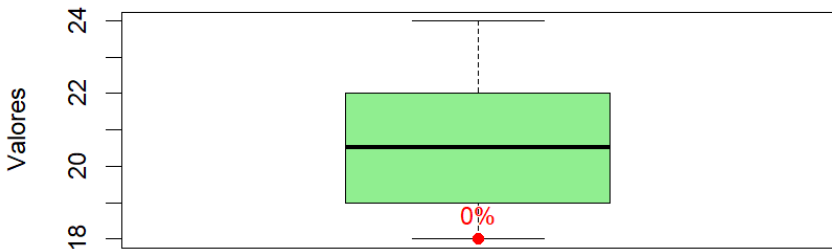
Resultados

- Mínimo: 18
- Q1 (25%): 19.25
- Mediana (Q2): 20.5
- Q3 (75%): 22.25
- Máximo: 24
- Sin valores atípicos (los bigotes abarcan todo el rango)

**Código en R**

```
# Boxplot
boxplot(tiempos,
        main = "Boxplot de tiempos",
        ylab = "Valores",
        col = "lightgreen")
# Añadir puntos para los cuartiles, mínimo y máximo
points(quantile(tiempos), col = "red", pch = 19)
text(quantile(tiempos), labels = "0%",
     names(quantile(tiempos)),
     pos = 3, col = "red")
```

**Boxplot de tiempos**



*Fig. 19 Boxplot del ejemplo de tiempos de atención*

Interpretación: La caja es relativamente estrecha ( $Q3 - Q1 = 3$  minutos), lo que indica que el 50% central de los datos está concentrado. Esto sugiere cierta estabilidad, aunque la presencia de colas (valores en los extremos) indica que hay servicios que se salen del rango habitual.

## 8.2 Análisis de la variabilidad en procesos

La variabilidad es inherente a todo proceso industrial. Ningún proceso produce resultados idénticos de manera continua. La labor del ingeniero industrial no es eliminar la variabilidad, sino entenderla, medirla y controlarla.

### 8.2.1 Tipos de variabilidad

Variabilidad común (aleatoria): Es la variabilidad natural del proceso, debida a causas inherentes al sistema (pequeñas fluctuaciones en materiales, temperatura, operarios, etc.). Es predecible dentro de ciertos límites.

Variabilidad especial (asignable): Causada por eventos anómalos y evitables: fallas de máquina, errores humanos, cambios bruscos en materia prima, etc. Se manifiesta como puntos fuera de los límites de control.

Comprender la variabilidad es esencial para:

- Control estadístico de procesos (CEP)
- Mejora continua
- Reducción de desperdicios
- Toma de decisiones basadas en datos

### 8.2.2 Ejemplo: Control de peso en envasado

Una máquina llena bolsas de café con un peso nominal de 500 gramos. Se toma una muestra de 10 bolsas y se registran los pesos:

**Código en R**

```
pesos <- c(498, 502, 501, 499, 500, 503, 497, 501, 500,
499)
#Calculamos estadísticos:
media_peso <- mean(pesos)
desviacion_peso <- sd(pesos)
```

**Resultados**

media\_peso = 500.0 g

desviación\_peso = 1.89 g

Interpretación: La media cumple exactamente el objetivo (500 g), pero la desviación estándar de 1.89 g indica que los pesos varían casi 2 gramos. Si la especificación del cliente es  $500 \pm 5$  g, el proceso parece capaz. Sin embargo, si la variabilidad aumentara, podrían generarse productos fuera de especificación.

Para visualizar la variabilidad, podemos construir un histograma o un boxplot. Además, podríamos calcular la proporción de bolsas que caerían fuera de los límites si asumimos normalidad.

**Código en R**

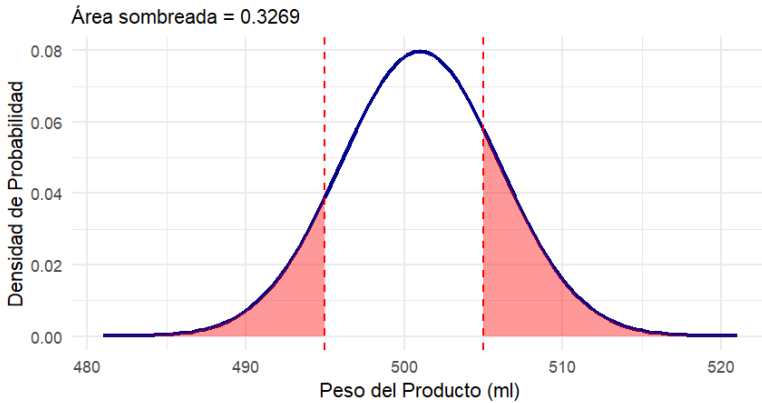
```
# Probabilidad de peso < 495 o > 505
p_menor <- pnorm(495, mean = media_peso, sd =
desviacion_peso)
p_mayor <- 1 - pnorm(505, mean = media_peso, sd =
desviacion_peso)
p_fuera <- p_menor + p_mayor
```

```

## graficar Parámetros (asumiendo valores del
ejercicio anterior)
library(ggplot2)
# Crear el data frame para la curva
df <- data.frame(x = c(media_peso - 4*desviacion_peso,
media_peso + 4*desviacion_peso))

ggplot(df, aes(x)) +
  # Curva normal
  stat_function(fun = dnorm, args = list(mean =
media_peso, sd = desviacion_peso),
               color = "darkblue", size = 1) +
  # Sombreado inferior
  stat_function(fun = dnorm, args = list(mean =
media_peso, sd = desviacion_peso),
               xlim = c(media_peso -
4*desviacion_peso, 495),
               geom = "area", fill = "red", alpha =
0.4) +
  # Sombreado superior
  stat_function(fun = dnorm, args = list(mean =
media_peso, sd = desviacion_peso),
               xlim = c(505, media_peso +
4*desviacion_peso),
               geom = "area", fill = "red", alpha =
0.4) +
  # Líneas verticales de límites
  geom_vline(xintercept = c(495, 505), linetype =
"dashed", color = "red") +
  labs(title = "Análisis de Límites de Aceptación",
        subtitle = paste("Área sombreada =",
round(p_fuera, 4)),
        x = "Peso del Producto (ml)",
        y = "Densidad de Probabilidad") +
  theme_minimal()

```



*Fig. 20 Análisis de límites de aceptación del ejemplo de control de peso en envasado*

Este tipo de análisis es la base para el control estadístico de procesos, que se profundizará en capítulos posteriores.

### 8.3 Indicadores estadísticos para la gestión

Los indicadores estadísticos permiten monitorear el desempeño de los procesos, convirtiéndose en herramientas clave para la gestión operativa y estratégica. A continuación, se presentan los más utilizados en ingeniería industrial, con énfasis en su interpretación.

#### 8.3.1 Media del proceso ( $\bar{x}$ )

Indica el nivel promedio de la variable de interés. Es fundamental para comparar con estándares o metas. Por ejemplo, la media de producción por turno, el tiempo medio de ciclo, etc.

### 8.3.2 Desviación estándar (s)

Mide la dispersión absoluta. Un proceso con baja desviación es más predecible y fácil de controlar. La reducción de la desviación estándar suele ser más estratégica que la mejora de la media, porque aumenta la capacidad del proceso.

### 8.3.3 Coeficiente de variación (CV)

Permite comparar la variabilidad relativa entre diferentes procesos, incluso si tienen diferentes unidades o magnitudes. Por ejemplo, comparar la variabilidad del tiempo de ciclo de dos líneas de producción con diferentes velocidades.

### 8.3.4 Percentiles

Los percentiles indican el valor por debajo del cual cae un porcentaje dado de observaciones. Son útiles para establecer límites de especificación o niveles de servicio. Por ejemplo, el percentil 90 del tiempo de espera indica que el 90% de los clientes esperan menos de ese tiempo.

### 8.3.5 Ejemplo comparativo de dos líneas de producción

Una empresa tiene dos líneas de ensamblaje. Se registran los tiempos de producción (en minutos) para 7 lotes en cada línea:

#### Código en R

```
linea_A <- c(30, 32, 31, 29, 33, 30, 31)
linea_B <- c(28, 35, 27, 36, 29, 34, 26)
```

```
###Calculamos estadísticos:
media_A <- mean(linea_A)      # 30.86
media_B <- mean(linea_B)      # 30.71
sd_A <- sd(linea_A)          # 1.35
sd_B <- sd(linea_B)          # 3.95
cv_A <- sd_A / media_A       # 0.044
cv_B <- sd_B / media_B       # 0.129

###Graficamos
library(ggplot2)
library(tidyr)

# Preparar los datos en formato largo (Tidy Data)
datos <- data.frame(linea_A, linea_B)
datos_long <- pivot_longer(datos, cols = everything(),
names_to = "Linea", values_to = "Tiempo")

ggplot(datos_long, aes(x = Linea, y = Tiempo, fill =
Linea)) +
  geom_boxplot(alpha = 0.7) +
  geom_jitter(width = 0.1, size = 2) +
  scale_fill_manual(values = c("linea_A" = "#56B4E9",
"linea_B" = "#E69F00")) +
  labs(title = "Análisis de Variabilidad: Línea A vs
Línea B",
        subtitle = "Obsérvese la mayor dispersión en la
Línea B",
        x = "Línea de Ensamblaje",
        y = "Tiempo de Producción (min)") +
  theme_minimal() +
  theme(legend.position = "none")
```

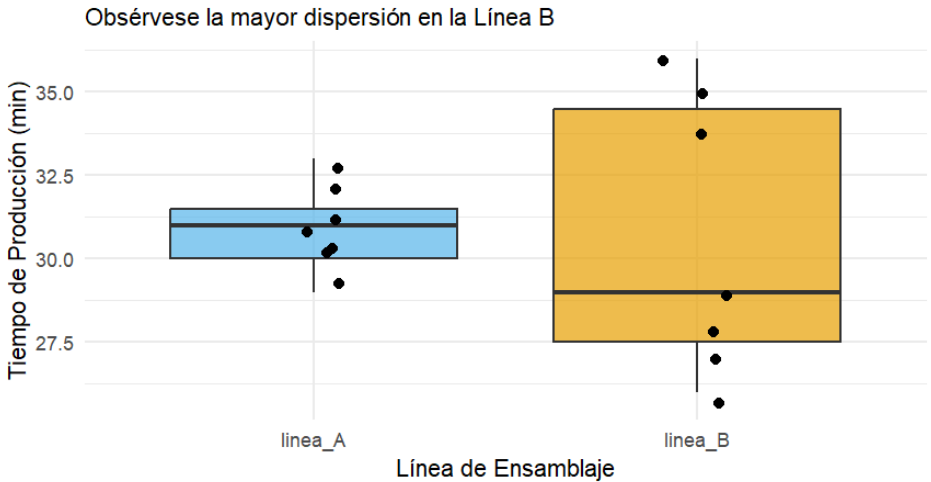


Fig. 21 Análisis de variabilidad: Línea A vs Línea B

Interpretación: Ambas líneas tienen medias similares (alrededor de 30.8 minutos). Sin embargo, la desviación estándar de la línea B es casi tres veces mayor, lo que se refleja en un CV del 12.9% frente al 4.4% de la línea A. La línea B es mucho más variable, lo que dificulta la planificación y puede generar cuellos de botella impredecibles. La acción de mejora debería centrarse en reducir la variabilidad de la línea B.

### 8.4 Ejercicios propuestos con R

Los siguientes ejercicios tienen como objetivo reforzar la aplicación práctica de la estadística descriptiva, el análisis de la variabilidad y el uso de indicadores estadísticos en contextos propios de la Ingeniería

Industrial. Todos los ejercicios deben resolverse utilizando R y acompañarse de una interpretación técnica de los resultados.

### **Ejercicio 1. Análisis descriptivo de tiempos de proceso**

Una empresa registra los tiempos de ensamblaje (en minutos) de un producto durante un turno de trabajo:

#### **Código en R**

```
tiempos <- c(42, 45, 40, 43, 44, 46, 41, 45, 47, 42)
```

Se pide:

1. Calcular la media, mediana y moda.
2. Determinar el rango, varianza y desviación estándar.
3. Elaborar un histograma de los tiempos.
4. Interpretar los resultados desde el punto de vista del desempeño del proceso.

### **Ejercicio 2. Evaluación de la variabilidad de un proceso productivo**

Se registran los diámetros (en mm) de piezas producidas por una máquina:

#### **Código en R**

```
diametros <- c(50.2, 49.8, 50.1, 50.0, 49.9, 50.3, 49.7,  
50.2, 50.1, 49.8)
```

Se pide:

1. Calcular la media y desviación estándar.
2. Analizar la variabilidad del proceso.
3. Construir un histograma.
4. Indicar si el proceso parece estable y justificar la respuesta.

### **Ejercicio 3. Comparación de la variabilidad entre dos procesos**

Dos líneas de producción fabrican el mismo producto. Los tiempos de producción (en minutos) son:

#### **Código en R**

```
linea_1 <- c(60, 62, 61, 59, 63, 60, 61)
linea_2 <- c(58, 66, 57, 65, 59, 67, 56)
```

Se pide:

1. Calcular la media y desviación estándar de cada línea.
2. Calcular el coeficiente de variación.
3. Determinar qué línea presenta mayor estabilidad.
4. Explicar las implicancias operativas de la variabilidad observada.

### **Ejercicio 4. Indicadores estadísticos en un proceso de servicio**

Un centro de atención registra los tiempos de espera (en minutos) de los clientes:

### Código en R

```
espera <- c(12, 15, 14, 13, 16, 18, 11, 14, 17, 15)
```

Se pide:

1. Calcular la media y desviación estándar.
2. Calcular el coeficiente de variación.
3. Determinar el percentil 90 del tiempo de espera.
4. Interpretar el percentil como indicador de calidad del servicio.

### Ejercicio 5. Detección de valores atípicos en un proceso

Se tienen los siguientes consumos de energía (kWh) por turno:

### Código en R

```
energia <- c(120, 125, 118, 122, 130, 119, 121, 180,  
124, 123)
```

Se pide:

1. Calcular los principales estadísticos descriptivos.
2. Construir un diagrama de caja (boxplot).
3. Identificar posibles valores atípicos.
4. Analizar las causas operativas que podrían generar estos valores.

### Ejercicio 6. Análisis integral de un proceso industrial

Una empresa mide la producción diaria (unidades) durante dos semanas:

### Código en R

```
produccion <- c(980, 1005, 995, 990, 1010, 1000, 985,  
                990, 1020, 1015, 995, 1005, 1000, 990)
```

Se pide:

1. Calcular la media, desviación estándar y coeficiente de variación.
2. Elaborar un histograma.
3. Evaluar la estabilidad del proceso.
4. Proponer acciones de mejora basadas en el análisis estadístico.

## Capítulo 9. Inferencia estadística en problemas industriales

La estadística descriptiva nos permite resumir y visualizar datos, pero en la gestión industrial necesitamos ir más allá: debemos tomar decisiones basadas en evidencia, generalizar conclusiones a partir de muestras y cuantificar la incertidumbre. La inferencia estadística es el motor que transforma datos aislados en evidencia organizacional. En este capítulo aprenderemos a no confiar ciegamente en un promedio, sino a cuestionar la variabilidad y la confianza de cada indicador que gestionamos.

### 9.1 Estimación y muestreo

En ingeniería industrial, el muestreo no es solo una técnica matemática, es una decisión de costo-beneficio. Medir todo (censo) es costoso; no medir nada es arriesgado. La clave está en seleccionar muestras representativas que nos permitan estimar parámetros poblacionales con un nivel aceptable de precisión.

#### 9.1.1 Conceptos básicos de muestreo

- Población: Conjunto total de elementos de interés (ej. todas las piezas producidas en un mes).
- Muestra: Subconjunto de la población seleccionado para el análisis.
- Parámetro: Valor numérico que describe una característica de la población (ej. media poblacional  $\mu$ ).

- Estadístico: Valor numérico calculado a partir de la muestra.  
(ej. media muestral, que denotamos como:  $\bar{x}$ )

### 9.1.2 Distribución muestral y error estándar

Si tomamos muchas muestras de la misma población, las medias muestrales seguirán una distribución (teorema central del límite) con media igual a la media poblacional y desviación estándar llamada error estándar:

$$\text{Error estándar} = \frac{\sigma}{\sqrt{n}}$$

donde  $\sigma$  es la desviación estándar poblacional (o muestral si no se conoce) y  $n$  el tamaño de muestra. El error estándar mide la precisión de nuestra estimación: cuanto mayor es  $n$ , menor es el error.

### 9.1.3 Determinación del tamaño de muestra

El ingeniero debe determinar cuántas observaciones son necesarias para que el error de estimación no afecte la rentabilidad. Para estimar la media poblacional con un margen de error  $E$  y un nivel de confianza  $(1 - \alpha)$ , el tamaño mínimo de muestra (población infinita) es:

$$n = \left( \frac{Z_{\alpha/2} \cdot \sigma}{E} \right)^2$$

donde

$Z_{\alpha/2}$  es el valor crítico de la distribución normal estándar (por ejemplo, para un nivel de confianza del 95%,  $\alpha = 0.05$ , entonces  $\alpha/2 = 0.025$  y

$z(0.025) \approx 1.96$ ). En la práctica,  $\sigma$  se estima a partir de estudios previos o una muestra piloto.

Ejemplo: Deseamos estimar el tiempo promedio de atención en un servicio técnico con un error máximo de  $\pm 2$  minutos y un nivel de confianza del 95%. Si por estudios previos sabemos que la desviación típica es de 8 minutos, ¿cuántas observaciones necesitamos?

$$n = \left( \frac{1.96 \times 8}{2} \right)^2 = (7.84)^2 = 61.5 = 62 \text{ observaciones}$$

#### 9.1.4 Intervalos de confianza

Un intervalo de confianza (IC) proporciona un rango de valores plausibles para el parámetro poblacional, con un cierto nivel de confianza (generalmente 95%). Para la media:

Si  $\sigma$  es conocida:

$$IC = \tilde{x} \pm z \frac{\alpha}{2} \frac{\sigma}{\sqrt{n}}$$

Si  $\sigma$  es desconocida (lo más frecuente), se usa la distribución t de Student:

$$IC = \tilde{x} \pm t(n-1, \alpha/2) \frac{s}{\sqrt{n}}$$

donde  $t(n-1, \alpha/2)$  es el valor crítico de la distribución t con n-1 grados de libertad y nivel de significancia  $\alpha/2$ .

Un IC estrecho indica un proceso bajo control y una estimación precisa; un IC amplio refleja incertidumbre operativa y puede alertar sobre la necesidad de aumentar el tamaño de muestra o reducir la variabilidad.

Ejemplo en R: Tiempos de descarga de camiones (datos simulados).

### Código en R

```
# Datos de muestra
set.seed(123)
tiempos_descarga <- rnorm(50, mean = 45, sd = 8)

# Intervalo de confianza al 95%
t.test(tiempos_descarga, conf.level = 0.95)$conf.int
```

## 9.2 Pruebas de hipótesis

Una prueba de hipótesis es un filtro contra el autoengaño. Nos permite distinguir entre una mejora real y una fluctuación aleatoria. En el entorno industrial, las pruebas de hipótesis se utilizan para validar cambios en procesos, comparar proveedores, evaluar la efectividad de capacitaciones, etc.

### 9.2.1 Elementos de una prueba de hipótesis

- Hipótesis nula ( $H_0$ ): Afirmación que se asume cierta hasta que se demuestre lo contrario (ej. "el nuevo proveedor no mejora la resistencia").
- Hipótesis alternativa ( $H_1$ ): Lo que queremos demostrar (ej. "el nuevo proveedor tiene menor resistencia").

- Estadístico de prueba: Valor calculado a partir de los datos que se compara con una distribución teórica.
- Región de rechazo: Conjunto de valores del estadístico que llevan a rechazar  $H_0$ .
- Nivel de significancia ( $\alpha$ ): Probabilidad de rechazar  $H_0$  cuando es verdadera (Error Tipo I). Comúnmente  $\alpha = 0.05$ .
- Valor p (p-value): Probabilidad de obtener un resultado tan extremo o más que el observado, suponiendo que  $H_0$  es cierta. Si el valor p es menor que  $\alpha$ , rechazamos  $H_0$ .

### 9.2.2 Errores Tipo I y Tipo II en la industria

- **Error Tipo I (falsa alarma):** Rechazar  $H_0$  cuando es verdadera. Ejemplo: Detener una línea de producción porque se cree que hay un problema cuando en realidad todo está bien. Esto genera pérdidas de tiempo y dinero.
- **Error Tipo II (falso negativo):** No rechazar  $H_0$  cuando es falsa. Ejemplo: No detectar que un lote de productos es defectuoso, lo que puede llevar a enviar productos no conformes al cliente, con el consiguiente riesgo de reclamos y daño a la reputación.

El ingeniero debe equilibrar estos errores, eligiendo un  $\alpha$  adecuado y asegurando un tamaño de muestra suficiente para tener una **potencia** adecuada (probabilidad de detectar un efecto real).

### 9.2.3 Tipos de pruebas comunes

- Prueba t para una muestra: Compara la media de una muestra con un valor de referencia.
- Prueba t para dos muestras independientes: Compara las medias de dos grupos (ej. turnos, proveedores).
- Prueba t pareada: Compara mediciones antes y después en los mismos elementos (ej. mejora de proceso).
- Prueba de proporciones: Compara proporciones (ej. tasa de defectos).

### 9.2.4 Interpretación del valor p

El valor p es la métrica de credibilidad. Si  $p < 0.05$  (usualmente), los datos "gritan" que algo ha cambiado en el proceso. Sin embargo, el valor p no mide la magnitud del efecto; un efecto muy pequeño puede ser estadísticamente significativo con una muestra grande. Por eso es crucial complementar con intervalos de confianza y análisis práctico.

### 9.3 Casos de estudio aplicados

#### Caso A: Validación de un nuevo proveedor de materia prima

Contexto: Una planta de inyección de plástico evalúa cambiar de proveedor de resina. El nuevo proveedor promete la misma resistencia mecánica a un costo 15% menor. El estándar histórico de resistencia es de 80 MPa (megapascales) con una desviación de 5 MPa.

El problema: El ingeniero toma una muestra de 40 piezas del nuevo proveedor y encuentra una resistencia media de 78.2 MPa con una

desviación de 4.8 MPa. ¿Es esta diferencia estadísticamente significativa?

Análisis de inferencia: Se plantea una prueba de hipótesis de una cola (inferior) porque nos interesa detectar si la resistencia es menor que el estándar.

$H_0: \mu = 80$  (la resistencia es igual)

$H_1: \mu < 80$  (la resistencia es menor)

Suponiendo que la desviación poblacional es conocida (5 MPa), usamos la prueba z:

$$z = \frac{(\bar{x} - \mu_0)}{\left(\frac{\sigma}{\sqrt{n}}\right)} = \frac{(78.2 - 80)}{\left(\frac{5}{\sqrt{40}}\right)} = -2.28$$

El valor p correspondiente es la probabilidad de obtener un z menor que -2.28 en una distribución normal estándar: aproximadamente 0.0113. Como  $p < 0.05$ , rechazamos  $H_0$ . Concluimos que la resistencia media del nuevo proveedor es significativamente menor. El ahorro nominal se traduciría en fallos estructurales, por lo que se recomienda no cambiar de proveedor o negociar mejoras.

#### Código en R

```
# Datos
xbar <- 78.2
mu0 <- 80
sigma <- 5
n <- 40
z <- (xbar - mu0) / (sigma / sqrt(n))
p_valor <- pnorm(z) # para una cola inferior
p_valor
```

**Caso B: Reducción de tiempos de set-up (SMED)**

Contexto: Tras implementar una metodología SMED, el equipo de mejora afirma haber reducido el tiempo de cambio de molde de 120 a 95 minutos. La gerencia sospecha que los 95 minutos fueron un "golpe de suerte" de un solo turno.

El problema: Se registran 15 cambios antes y 15 cambios después. Los datos son:

**Antes:** 122, 118, 125, 119, 121, 123, 120, 124, 117, 126, 120, 121, 119, 122, 118

**Después:** 95, 98, 92, 97, 96, 94, 99, 93, 95, 97, 96, 98, 94, 95, 97

Análisis de inferencia: Se comparan las medias mediante una prueba t para dos muestras independientes (asumiendo varianzas iguales o diferentes). Se realiza una prueba de una cola (disminución).

$H_0: \mu_{\text{antes}} = \mu_{\text{después}}$

$H_1: \mu_{\text{antes}} > \mu_{\text{después}}$

**Código en R**

```

antes <-
c(122,118,125,119,121,123,120,124,117,126,120,121,119,122,118
)
despues <- c(95,98,92,97,96,94,99,93,95,97,96,98,94,95,97)

# Prueba t de una cola (alternativa "greater" porque antes >
después)
t.test(antes, despues, alternative = "greater",
var.equal = FALSE)

```

El resultado arroja un valor p extremadamente pequeño ( $< 0.001$ ), lo que indica que la reducción es estadísticamente significativa. Por tanto, se estandariza el nuevo procedimiento en el sistema de gestión de información.

### **Caso C: Análisis de efectividad de un programa de capacitación en seguridad**

Contexto: En la minera ZET (capítulo 7), se implementó un programa de capacitación en seguridad. Se desea evaluar si la tasa de accidentes (por cada 100 trabajadores) disminuyó después de la capacitación.

Datos: Se tienen las tasas mensuales antes y después durante 12 meses.

Antes: 3.2, 3.5, 3.1, 3.8, 3.4, 3.6, 3.3, 3.7, 3.2, 3.9, 3.5, 3.4

Después: 2.8, 2.9, 2.7, 3.0, 2.6, 2.9, 2.5, 2.8, 2.7, 2.9, 2.6, 2.8

Análisis: Se utiliza una prueba t pareada porque las mediciones están emparejadas por mes.

#### **Código en R**

```
antes <-  
c(3.2, 3.5, 3.1, 3.8, 3.4, 3.6, 3.3, 3.7, 3.2, 3.9, 3.5, 3.4)  
despues <-  
c(2.8, 2.9, 2.7, 3.0, 2.6, 2.9, 2.5, 2.8, 2.7, 2.9, 2.6, 2.8)  
t.test(antes, despues, paired = TRUE, alternative =  
"greater")
```

El valor  $p$  es menor a 0.05, confirmando que la capacitación tuvo un efecto significativo en la reducción de accidentes. Esta evidencia respalda la continuidad y expansión del programa.

## 9.4 Ejercicios propuestos

Los siguientes ejercicios tienen como objetivo aplicar los conceptos de inferencia estadística en contextos industriales. Resuélvalos utilizando R y acompañe cada resultado con una interpretación ingenieril.

### Ejercicio 1. Estimación de capacidad de almacén

Una empresa logística registra los tiempos de descarga de camiones. En una muestra de 50 camiones, la media fue de 45 minutos con una desviación estándar de 8 minutos.

Se pide:

Construya un intervalo de confianza del 95% para el tiempo medio de descarga.

Si el sistema de turnos está diseñado para bloques de 40 minutos, ¿qué nos dice el límite superior del intervalo sobre la saturación del muelle?

Datos: Use los datos simulados o los proporcionados. En R:

#### Código en R

```
tiempos <- rnorm(50, mean=45, sd=8) # o usar los
datos dados
ic <- t.test(tiempos, conf.level=0.95)$conf.int
ic
```

**Ejercicio 2. Comparación de turnos (prueba de hipótesis)**

El turno mañana produce un promedio de 200 unidades/hora con una desviación de 15 unidades (muestra de 30 días). El turno noche produce 192 unidades/hora con desviación de 18 unidades (muestra de 25 días). Utilice una prueba de hipótesis con  $\alpha = 0.05$  para determinar si la diferencia de productividad es significativa (suponga varianzas diferentes). Interprete el resultado y sugiera posibles causas basadas en la gestión de la información.

**Código en R**

```
# Simular datos o usar estadísticos resumen
# Prueba t de dos muestras
t.test(x = rnorm(30, 200, 15), y = rnorm(25, 192,
18), alternative = "two.sided")
```

**Ejercicio 3. Análisis de variabilidad en llenado**

Una embotelladora debe llenar envases de 500 ml. Una muestra de 100 botellas muestra un promedio de 501 ml, pero una desviación estándar de 5 ml.

¿Está el proceso cumpliendo con la normativa si esta exige que el 95% de las botellas tengan al menos 495 ml?

¿Es preferible ajustar la media a 505 ml o trabajar en reducir la varianza? Calcule el impacto en el desperdicio de producto (giveaway).

Ayuda: Para evaluar el cumplimiento, calcule la proporción esperada de botellas por debajo de 495 ml asumiendo normalidad. Use la función `pnorm(495, mean=501, sd=5)`. Luego discuta las opciones.

#### **Ejercicio 4. Prueba de hipótesis para mejora en mantenimiento**

Un equipo de mantenimiento implementa un nuevo procedimiento y afirma que reduce el tiempo medio de reparación de 60 minutos a menos de 55 minutos. Se registran 20 reparaciones después del cambio, con media 53 minutos y desviación 6 minutos. Realice una prueba de hipótesis con  $\alpha = 0.05$  y concluya si la mejora es estadísticamente significativa. ¿Qué tamaño de muestra sería necesario para detectar una reducción de 5 minutos con una potencia del 80%?

#### **Ejercicio 5. Comparación de proporciones de defectos**

Dos proveedores suministran el mismo componente. Se inspeccionan 200 piezas de cada proveedor. El proveedor A presenta 12 defectuosas; el proveedor B, 20. ¿Existe diferencia significativa en la proporción de defectos? Use una prueba de proporciones.

**Código en R**

```
prop.test(x = c(12, 20), n = c(200, 200))
```

## **PARTE IV. Machine Learning e Inteligencia Artificial en la Ingeniería Industrial**

### **Capítulo 10. Machine Learning como extensión del análisis industrial**

La estadística inferencial, presentada en el capítulo anterior, nos permite confirmar hipótesis y validar si un cambio en el proceso es significativo. Sin embargo, en la era de la Industria 4.0, los volúmenes de datos y la complejidad de las relaciones entre variables exigen herramientas que no solo confirmen, sino que también descubran patrones ocultos, automaticen la detección de anomalías y aprendan de la experiencia acumulada. Aquí es donde el Machine Learning (ML) se convierte en una extensión natural del análisis industrial.

#### **10.1 Relación entre estadística, BI y ML**

Es un error común pensar en la estadística, la inteligencia de negocios y el machine learning como disciplinas aisladas o excluyentes. En realidad, forman un continuo analítico que el ingeniero industrial debe dominar.

- **Estadística:** Nos proporciona el marco de pensamiento crítico. Nos enseña a no confiar ciegamente en los datos, a cuantificar la incertidumbre y a validar si un hallazgo es real o producto del azar. Es la base sobre la que se construye todo lo demás.
- **Inteligencia de Negocios (BI):** Se enfoca en el pasado y el presente. Responde a preguntas como "¿qué ocurrió?" y "¿por

qué ocurrió?". A través de dashboards y reportes, el BI nos da la visibilidad para entender el desempeño histórico de los procesos, tal como se desarrolló en la Parte II de este libro.

- **Machine Learning (ML):** Se proyecta hacia el futuro y se adentra en la complejidad. Responde a preguntas como "¿qué ocurrirá?" (predicción) y "¿qué deberíamos hacer?" (prescripción). El ML aprende de los datos históricos para identificar patrones que serían imposibles de codificar con reglas fijas.

La transición del BI al ML no es un salto, sino una evolución. Primero, el BI nos ayuda a saber que la temperatura de un horno se desvía a menudo de su valor objetivo. Luego, la estadística nos ayuda a confirmar que esa desviación no es aleatoria. Finalmente, el ML nos permite predecir con antelación cuándo ocurrirá esa desviación en función de múltiples variables (humedad, desgaste de la resistencia, lote de materia prima, etc.) y sugerir un ajuste preventivo.

### **10.2 Tipos de problemas abordables en entornos industriales**

No todos los problemas industriales requieren ML. De hecho, aplicarlo donde una regla simple o un análisis estadístico básico sería suficiente es una receta para la ineficiencia. El ML es valioso cuando la relación entre variables es compleja, no lineal, o cuando el volumen de datos es tan grande que un humano no podría procesarlo. Los principales tipos de problemas que aborda el ML en la industria se clasifican en:

- **Aprendizaje Supervisado:** Entrenamos un modelo con datos históricos que ya contienen la respuesta. El objetivo es aprender la relación entre las variables de entrada (características) y la variable de salida (etiqueta) para predecirla en nuevos datos.
- **Regresión:** La variable de salida es un número continuo. Ejemplos: predecir la demanda de un producto para la próxima semana, estimar la vida útil restante de un equipo, pronosticar el consumo energético de una planta.
- **Clasificación:** La variable de salida es una categoría. Ejemplos: determinar si un lote será defectuoso o no (defectuoso/ok), clasificar el tipo de falla de una máquina (eléctrica/mecánica/operacional), identificar si una transacción es fraudulenta.
- **Aprendizaje No Supervisado:** Buscamos patrones y estructuras ocultas en los datos sin tener una variable de salida predefinida.
- **Segmentación (Clustering):** Agrupar elementos similares. Ejemplos: segmentar clientes según su patrón de compra, agrupar máquinas según su perfil de vibración, categorizar órdenes de trabajo según su causa raíz.
- **Detección de Anomalías:** Identificar puntos que se desvían drásticamente del comportamiento normal. Ejemplos: detectar picos de consumo eléctrico anómalos, identificar

piezas con defectos de fabricación no vistos antes, alertar sobre transacciones inusuales en la cadena de suministro.

El ingeniero industrial debe ser capaz de identificar qué tipo de problema enfrenta y, por tanto, qué familia de herramientas de ML es más adecuada.

### **10.3 Datos como insumo crítico: del BI al ML**

El éxito de cualquier proyecto de ML depende, en un 80%, de la calidad y preparación de los datos. Un modelo sofisticado con datos basura produce resultados basura. Aquí es donde la ingeniería de datos y el BI juegan un papel fundamental.

Lo que funciona para un dashboard de BI no siempre está listo para un modelo de ML. El ML requiere:

- Datos históricos etiquetados y limpios: Para un modelo de mantenimiento predictivo, no basta con saber que una máquina falló el martes. Necesitamos saber qué valores de temperatura, vibración, presión, etc., tenía la máquina en las horas y días previos a la falla. La falla es la "etiqueta", y las lecturas de los sensores son las "características".
- Granularidad y contexto: Mientras que un dashboard puede mostrar la producción diaria, un modelo de ML puede necesitar datos por minuto, integrados con información del turno, el operador, el lote de materia prima y las condiciones

ambientales. Es la unificación de datos de sensores (IoT) con datos de sistemas transaccionales (ERP, MES).

- Series temporales bien formadas: Muchos problemas industriales son esencialmente de evolución en el tiempo. La secuencia de los datos es crucial. No es lo mismo un valor de vibración alto hoy que una tendencia creciente en los últimos 10 días.

El flujo de trabajo en un proyecto de Machine Learning no es lineal, sino un ciclo de mejora continua que transforma señales físicas en decisiones estratégicas:

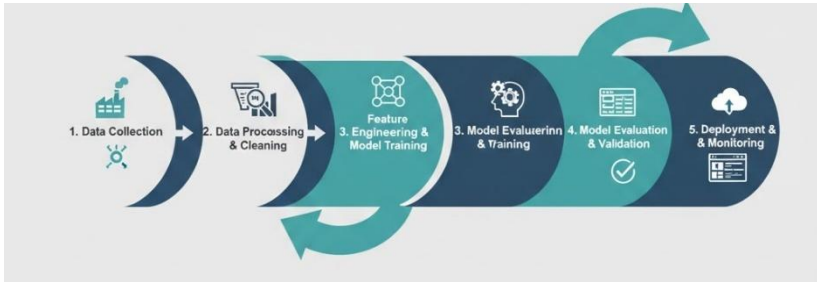
**Captura (Ingesta):** Se recolectan datos brutos de sensores (IoT), PLCs y sistemas de gestión (ERP/MES) para digitalizar la realidad de la planta.

**Saneamiento (Limpieza):** Se elimina el "ruido" industrial y los errores de lectura, asegurando que el modelo trabaje con información veraz y sincronizada.

**Ingeniería de Valor (Features):** El ingeniero traduce su conocimiento del proceso en variables clave que el algoritmo pueda entender para detectar patrones.

**Entrenamiento y Validación:** El algoritmo aprende de los datos históricos y se somete a pruebas de precisión (como la matriz de confusión) para asegurar que sus predicciones sean confiables.

Despliegue y Evolución: El modelo se aplica en la operación real para asistir en la toma de decisiones, manteniéndose en constante monitoreo para ajustarse a los cambios naturales del entorno productivo.



*Fig. 22 Flujo de datos en un proyecto de Machine Learning*

## Capítulo 11. Aplicaciones de Machine Learning en la ingeniería industrial

Si el capítulo anterior sentó las bases teóricas, este despliega su potencial en el terreno operativo. El Machine Learning se integra hoy como un estándar indispensable en el maletín de herramientas del ingeniero industrial, siendo aplicable a cualquier proceso que *genere datos*. A través de casos concretos, veremos cómo esta disciplina permite evolucionar de una gestión reactiva a una proactiva en áreas tan diversas como el mantenimiento predictivo, la logística inteligente y el control de calidad automatizado.

Es fundamental precisar que el propósito de este texto no es la escritura exhaustiva de código, sino el desarrollo del pensamiento analítico industrial: la capacidad de comprender el proceso de raíz, identificar las variables críticas que lo gobiernan y estructurar problemas que puedan ser resueltos mediante datos. Si bien se recomienda al lector desarrollar competencias en lenguajes como Python para ganar *autonomía técnica*, el verdadero valor del ingeniero industrial radica en su criterio para conectar la algoritmia con la realidad operativa. El objetivo final es asegurar que la tecnología esté siempre al servicio de la eficiencia y la toma de decisiones estratégicas.

### 11.1 Predicción de demanda y optimización de inventarios

Uno de los problemas más clásicos es la predicción de la demanda. Tradicionalmente, se han utilizado promedios móviles o suavizado exponencial, que son métodos estadísticos robustos pero que capturan relaciones lineales simples.

El enfoque con ML: Un modelo de regresión (como Random Forest o XGBoost) puede integrar cientos de variables que impactan la demanda, yendo mucho más allá de la simple serie de tiempo.

Variables de entrada: Historial de ventas, precios, promociones, datos macroeconómicos, clima, tendencias en redes sociales, festividades, lanzamientos de productos de la competencia.

Resultado: Un pronóstico de demanda no solo más preciso, sino que además puede ofrecer intervalos de confianza más ajustados.

### **Caso de estudio aplicado: Reducción de quiebres de stock**

Una empresa de consumo masivo (como la del Capítulo 2) implementa un modelo de ML para predecir la demanda de sus 500 productos en 50 centros de distribución. El modelo, entrenado con 5 años de datos, no solo considera las ventas históricas, sino también el clima (las ventas de helados se disparan con calor) y las promociones planificadas.

- Resultado: El modelo supera en precisión al método tradicional en un 25%. Esto se traduce en una reducción del 15% en los quiebres de stock y una disminución del 10% en el inventario de seguridad, liberando capital de trabajo.

- Rol del Ingeniero Industrial: No es quien programa el modelo, sino quien define las variables relevantes (conocimiento del proceso), valida que los datos históricos sean confiables (calidad del dato) y, lo más importante, traduce la predicción en una decisión de negocio: ajustar los niveles de inventario y renegociar volúmenes con proveedores.

### **11.2 Mantenimiento predictivo y confiabilidad de activos**

El mantenimiento industrial ha evolucionado del mantenimiento correctivo (reparar cuando falla) al preventivo (reparar cada cierto tiempo), y ahora al predictivo (reparar justo antes de que falle) .

El enfoque con ML: Se utilizan modelos de clasificación o regresión para predecir la probabilidad de falla o la vida útil remanente de un activo. Los sensores IoT (vibración, temperatura, corriente, etc.) proporcionan un flujo continuo de datos que alimentan el modelo.

Problema: Un modelo de clasificación se entrena con datos históricos donde cada registro de sensor está etiquetado como "funcionamiento normal" o "falla inminente" (en las horas previas a una parada). El modelo aprende el patrón de firma de los sensores que anticipa una falla.

Problema de regresión: Se predice el número de horas restantes antes de que un equipo necesite mantenimiento, basándose en su perfil de uso y desgaste.

### **11.3 Caso de estudio aplicado: Planta de manufactura de autopartes**

Una planta de mecanizado (como el contexto del Capítulo 1) instala sensores de vibración en sus centros de mecanizado. El objetivo es predecir la rotura de la herramienta de corte (fresa).

Solución: Se entrena un modelo de ML con datos de vibración etiquetados (condiciones: herramienta nueva, herramienta desgastada, herramienta rota). El modelo aprende a identificar la "huella digital" de la vibración que precede a la rotura.

Resultado: El sistema alerta al operario con 30 minutos de antelación. Esto evita paradas no programadas, reduce el scrap de piezas producidas con herramienta defectuosa y optimiza la vida útil de la herramienta, cambiándola solo cuando es necesario. Empresas como Nemak han dado un paso más allá, creando agentes de IA que permiten a los líderes de línea "chatear" con las máquinas para preguntar por su estado y obtener respuestas en tiempo real .

### **11.4 Gestión del talento humano en entornos industriales**

El ML también está transformando la gestión de las personas, un área crucial para el ingeniero industrial que busca optimizar el sistema completo, no solo las máquinas.

El enfoque con ML:

Predicción de rotación de personal: Entrenar un modelo de clasificación para identificar qué empleados tienen mayor riesgo de dejar la empresa. Las variables pueden incluir: antigüedad, ausentismo, evaluaciones de desempeño, distancia del hogar a la planta, turno, etc. Esto permite a RRHH y a los jefes de planta tomar acciones preventivas (entrevistas de retención, ajustes de turno, etc.).

- **Asignación de turnos y optimización de la fuerza laboral:** Utilizar ML para predecir la carga de trabajo y optimizar la asignación de personal a las líneas de producción en tiempo real, considerando las habilidades de cada operario y la complejidad de las tareas.
- **Preservación del conocimiento:** Como en el caso de Nematik, se pueden crear agentes conversacionales (Knowledge Centers) que permiten a los ingenieros noveles acceder a décadas de conocimiento técnico acumulado en reportes, lecciones aprendidas y manuales, simplemente haciendo preguntas en lenguaje natural. Esto democratiza el conocimiento y acelera la curva de aprendizaje.

### **11.5 Seguridad, riesgo y prevención (SSOMA)**

La seguridad laboral, vista en el Capítulo 7, es un área de alto impacto para el ML. Pasar de un análisis reactivo (investigar accidentes que ya

ocurrieron) a uno predictivo (predecir dónde y cuándo es más probable un accidente) es el siguiente paso natural.

El enfoque con ML:

- **Predicción de incidentes:** Un modelo de clasificación puede ser entrenado con datos históricos de accidentes, incidentes, actos y condiciones inseguras. Las variables de entrada incluirían: área de trabajo, turno, antigüedad del personal, horas extras trabajadas en la última semana, incumplimientos de procedimientos, condiciones ambientales (ruido, temperatura), y reportes de casi-accidentes. El modelo podría identificar que, en ciertas condiciones, la probabilidad de un incidente se dispara, permitiendo una intervención proactiva .
- **Visión por computadora para seguridad:** Utilizar cámaras y modelos de visión artificial para detectar en tiempo real si los operarios están usando el equipo de protección personal (casco, guantes) o si están realizando movimientos que podrían causar una lesión musculoesquelética (fatiga, posturas forzadas). El sistema puede enviar una alerta inmediata al supervisor.

## **Capítulo 12. Inteligencia Artificial y apoyo a la toma de decisiones**

El Machine Learning nos da predicciones. La Inteligencia Artificial, en su concepción más amplia, busca ir un paso más allá: apoyar, y en algunos casos ejecutar, la toma de decisiones para lograr un objetivo. Este es el dominio de la Inteligencia de Decisiones y, más recientemente, de los sistemas multi-agente.

### **12.1 IA como soporte al ingeniero: decisiones asistidas y aumentadas**

El objetivo de la IA en la industria no es reemplazar al ingeniero, sino amplificar sus capacidades. La IA es excelente procesando ingentes cantidades de datos y encontrando patrones, pero el ingeniero es quien aporta el contexto, el criterio y la responsabilidad ética.

En el nivel más básico, la IA actúa como un asistente inteligente:

- **Filtrado de alarmas:** En una planta química, pueden saltar cientos de alarmas al día. Un sistema de IA puede priorizarlas, correlacionarlas y presentar al operador solo las 5 alarmas críticas y su posible causa raíz, en lugar de una lista abrumadora.
- **Recomendaciones de ajuste:** Un modelo de control predictivo basado en modelo (MPC) puede analizar continuamente las variables de un proceso (temperatura, presión, flujo) y recomendar al operador el ajuste óptimo para maximizar la eficiencia o la calidad, incluso en condiciones cambiantes.

- Generación de procedimientos: Ante una falla, un agente de IA puede generar automáticamente una guía de resolución de problemas paso a paso para el técnico, basándose en manuales e históricos de reparaciones exitosas.

### **12.2 Automatización del análisis y de la decisión: Agentes de IA y sistemas multi-agente**

El siguiente nivel es la automatización no solo del análisis, sino también de la decisión. Esto se materializa en los agentes de IA. Un agente de IA es un sistema autónomo que puede percibir su entorno, razonar, tomar decisiones y ejecutar acciones para lograr un objetivo específico.

A diferencia de un chatbot que solo responde, un agente de IA actúa.

- En mantenimiento: Un agente de IA no solo predice una falla, sino que, de forma autónoma, crea una orden de trabajo en el sistema de gestión de mantenimiento, reserva las piezas de repuesto necesarias en el inventario y coordina con la agenda del técnico de guardia la hora óptima para la intervención, todo sin intervención humana.
- En logística: Un agente monitorea la cadena de suministro en tiempo real. Si detecta que un proveedor clave sufre un retraso, el agente puede replanificar la producción, buscar proveedores alternativos y ajustar las rutas de transporte para minimizar el impacto.

El futuro apunta a los sistemas multi-agente, donde múltiples agentes especializados (un agente de producción, un agente de logística, un agente de calidad, un agente de energía) colaboran y negocian entre sí para optimizar el funcionamiento global de la planta . El ingeniero industrial pasa entonces a ser el "orquestador" de este ecosistema inteligente, definiendo los objetivos, las reglas y los límites dentro de los cuales operan estos agentes.

### **12.3 Ética y responsabilidad profesional en la era de la IA**

La adopción de la IA no es un tema puramente técnico; conlleva profundas implicaciones éticas y de responsabilidad profesional que el ingeniero industrial debe liderar.

- **Sesgo algorítmico:** Si un modelo de ML para la contratación de personal se entrena con datos históricos que reflejan sesgos (por ejemplo, ciertos puestos siempre ocupados por hombres), el modelo perpetuará y amplificará ese sesgo. El ingeniero debe auditar los datos y los modelos para asegurar la equidad.
- **Transparencia y Explicabilidad (XAI):** No basta con que un modelo prediga una falla. El ingeniero debe poder entender por qué el modelo hizo esa predicción para poder confiar en ella y actuar en consecuencia. Esto es especialmente crítico en entornos de alto riesgo. Un modelo de "caja negra" no es aceptable.
- **Responsabilidad:** Si un agente de IA toma una decisión que causa un accidente o una pérdida económica, ¿quién es el

responsable? ¿El desarrollador, el operador o la empresa? La definición de marcos claros de responsabilidad y supervisión humana es crucial.

- Privacidad de los datos: El uso de datos de operarios (para análisis de fatiga o productividad) debe hacerse con respeto a su privacidad, transparencia y dentro del marco legal aplicable. La confianza de la fuerza laboral es un activo invaluable.

El ingeniero industrial, por su formación integral y su posición en la organización, está llamado a ser el guardián de estos principios, asegurando que la IA se use para aumentar el bienestar y la eficiencia, no para crear riesgos ocultos o perpetuar injusticias.

### **Resumen de la Parte IV**

En esta parte hemos recorrido el camino desde el Machine Learning como herramienta predictiva hasta la Inteligencia Artificial como soporte para la toma de decisiones autónoma.

- El Machine Learning no reemplaza a la estadística o al BI, sino que los extiende, permitiendo abordar problemas de mayor complejidad y volumen.
- Las aplicaciones industriales del ML son amplias y de alto impacto: predicción de demanda, mantenimiento predictivo, gestión del talento y seguridad laboral, entre otras.

- La calidad y preparación de los datos sigue siendo el factor crítico de éxito. Un modelo de ML es tan bueno como los datos con los que se entrena.

La evolución hacia los agentes de IA promete automatizar no solo el análisis, sino también la ejecución de decisiones, optimizando la operación en tiempo real.

Este poder conlleva una gran responsabilidad. El ingeniero industrial debe liderar la implementación de la IA con un fuerte sentido ético, asegurando la transparencia, la equidad y la supervisión humana.

**Referencias bibliográficas**

- Alvarez, J. (2026). Bases de datos para la Ingeniería Industrial (1.<sup>a</sup> ed.). Fondo Editorial.
- Chaffey, D., y White, G. (2019). Business information management: Improving performance using information systems (3.<sup>a</sup> ed.). Pearson.
- Davenport, T. H., y Harris, J. G. (2017). Competing on analytics: The new science of winning (2.<sup>a</sup> ed.). Harvard Business Review Press.
- Han, J., Kamber, M., y Pei, J. (2011). Data mining: Concepts and techniques (3.<sup>a</sup> ed.). Morgan Kaufmann.
- James, G., Witten, D., Hastie, T., y Tibshirani, R. (2021). An introduction to statistical learning: With applications in R (2.<sup>a</sup> ed.). Springer. <https://doi.org/10.1007/978-1-0716-1418-1>
- Kimball, R., y Ross, M. (2013). The data warehouse toolkit: The definitive guide to dimensional modeling (3.<sup>a</sup> ed.). Wiley.
- Kleppmann, M. (2017). Designing data-intensive applications: The big ideas behind reliable, scalable, and maintainable systems. O'Reilly Media.
- Laudon, K. C., y Laudon, J. P. (2021). Management information systems: Managing the digital firm (17.<sup>a</sup> ed.). Pearson.
- Lorraine, C. (2020). Decision intelligence: A toolkit for managers. CRC Press.
- Montgomery, D. C., y Runger, G. C. (2018). Applied statistics and probability for engineers (7.<sup>a</sup> ed.). Wiley.
- O'Neil, C. (2016). Weapons of math destruction: How big data increases inequality and threatens democracy. Crown.
- Provost, F., y Fawcett, T. (2013). Data science for business: What you need to know about data mining and data-analytic thinking. O'Reilly Media.

- Reis, J., y Housley, M. (2022). *Fundamentals of data engineering: Plan and build robust data systems*. O'Reilly Media.
- Shmueli, G., Patel, N. R., y Bruce, P. C. (2017). *Data mining for business analytics: Concepts, techniques, and applications in R*. Wiley.
- Goodfellow, I., Bengio, Y., y Courville, A. (2016). *Deep learning*. MIT Press.
- Géron, A. (2022). *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow (3.ª ed.)*. O'Reilly Media.
- Molnar, C. (2022). *Interpretable machine learning: A guide for making black box models explainable*. Leanpub.
- Russell, S., y Norvig, P. (2020). *Artificial intelligence: A modern approach (4.ª ed.)*. Pearson.
- Sutton, R. S., y Barto, A. G. (2018). *Reinforcement learning: An introduction (2.ª ed.)*. MIT Press.