

guayerd

# Fundamentos IA

## Introducción IA y datos

Clase 2

En colaboración con  
**IBM SkillsBuild**





# ¡Bienvenidos!

¿Nos presentamos?

- ¿Qué recuerdan de la clase anterior?
- ¿Qué esperan aprender?
- ¿Tienen alguna pregunta?

# Contenidos

Por temas

01

- Introducción IA

03

- Introducción Python

02

- Fundamentos del dato
- Pensamiento computacional

04

- Introducción a Python

# Objetivos de la clase

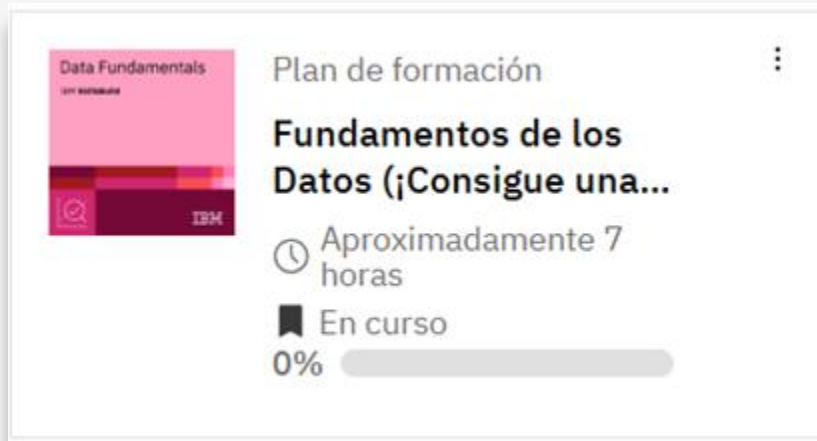


- Dato, información e insights
- Tipos de datos
- Ciclo de vida del dato
- Escalas de medida
- Estructuras
- Secuencias, condiciones y bucles
- Descomposición de problemas
- Relación con la automatización

# Introducción a la IA y los datos

Fundamentos del dato

# Plataforma Skill Build: Fundamentos del dato



The image shows a digital card for the 'Data Fundamentals' course on the IBM SkillsBuild platform. The card has a white background with a pink header area. On the left, there is a pink square with the text 'Data Fundamentals' and 'IBM' at the bottom. To the right of the square, the text 'Plan de formación' is displayed. Below this, the main title 'Fundamentos de los Datos (¡Consigue una...)' is shown in a large, bold font. Underneath the title, there is a clock icon followed by the text 'Aproximadamente 7 horas'. Below that, there is a book icon followed by the text 'En curso'. At the bottom, there is a progress bar showing '0%' completion.

Data Fundamentals

Plan de formación

**Fundamentos de los Datos (¡Consigue una...**

Aproximadamente 7 horas

En curso

0%



# Dato, información e insight

¿Son lo mismo? ¿Porqué?

### Datos:

- **Qué son:** Hechos y cifras sin procesar, sin significado intrínseco.
- **Ejemplo:** Una lista de transacciones de ventas (números brutos), 25, 33, 46, 63 (edades sin contexto).

### Información:

#### •Qué es:

- Datos que han sido procesados, organizados y contextualizados para tener sentido.

#### •Ejemplo:

- Analizar las ventas del último mes por categoría de producto para identificar las más vendidas;



### Insights:

#### •Qué es:

•Una comprensión profunda y un conocimiento práctico derivado del análisis de la información. Va más allá de los hechos, descubriendo las motivaciones y razones ocultas.

#### •Ejemplo:

•Descubrir que las ventas de ciertos productos aumentaron porque se relacionaron con promociones específicas o la estacionalidad; o identificar que la gente compra café de camino a la oficina debido a la necesidad de energía para empezar el día.

# Datos

## Materia prima

- **Describen las características de una entidad**
- Representación simbólica o general de un atributo o variable
- Pueden ser cuantitativos o cualitativos
- Existen de tipo valor numérico, texto y fórmulas



Dato = característica

# Datos

## Materia prima

### Cuantitativo

- **Representan cantidades**
- Referidos a números
- Se pueden transformar en estadísticas utilizables para medir y probar

**Ejemplo:** Edad, altura, precio, etc.

### Cualitativo

- **Representan categorías**
- Referidos a cualidades
- Se pueden identificar con preguntas como ¿cuál? o ¿cuáles?

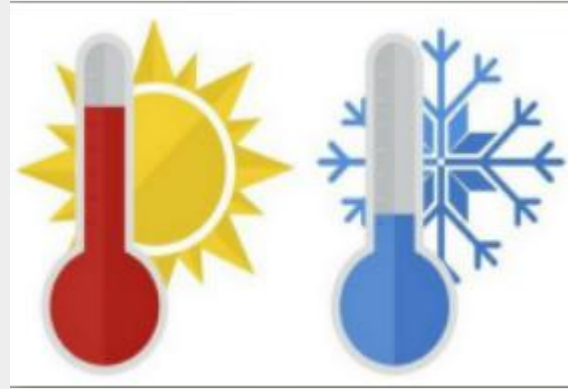
**Ejemplo:** Nombre, país, sexo, etc.

## DATOS CUANTITATIVOS

Discretos



Continuos



## DATOS CUALITATIVOS

Categoricos



Ordinales



# Datos

## Materia prima

### Según estructura

- **Estructurados:** tabulares, organizados en filas y columnas. Ej: Excel, SQL.
- **No estructurados:** texto libre, imágenes, audio, video. Ej: comentarios, fotos, grabaciones.
- **Semiestructurados:** tienen estructura flexible. Ej: JSON, XML, formularios web.

### Según origen

- **Primarios:** recolectados directamente por el analista.
- **Secundarios:** recopilados por otros (fuentes abiertas, sistemas previos).

# ¿Cómo determinar la calidad de los datos?

Mediante las siguientes **dimensiones**:

- **Exactitud:** Datos correctos
- **Compleitud:** Sin valores faltantes
- **Consistencia:** Mismo formato y codificación
- **Actualidad:** Datos vigentes
- **Relevancia:** Adecuados al objetivo

# Información



**Conjunto de datos procesados y organizados que proporcionan significado y contexto.** Son fundamental para la toma de decisiones, el análisis y la comunicación.



# Insight

**Descubrimientos o comprensiones profundas y reveladoras** obtenidas a partir del análisis de datos e información. Estos revelan patrones, tendencias o verdades ocultas que no son inmediatamente obvios.



# Ciclo de vida del dato



# Ciclo de vida del dato

Etapa	Objetivo	Ejemplos
Captura	Recoger datos	Formularios, sistemas, redes
Almacenamiento	Guardar datos	Bases de datos, archivos
Preparación	Limpiar, transformar y organizar	Python, Azure Data
Análisis	Explorar patrones y relaciones	
Comunicación	Presentar hallazgos	Power BI, Looker Studio
Decisión	Accionar en base a información	
Retroalimentación	Medir y ajustar resultados	

# Escala de medición

Escala	Características	Ejemplos
Nominal	Categorica sin orden	Género, tipo de producto
Ordinal	Categorica con orden	Nivel de satisfacción
Intervalo	Numérica sin cero real	Temperatura, fechas
Razón	Numérica con cero absoluto	Ingresos, cantidad de ventas



Saber la **escala** de cada dato es clave para aplicar el **análisis correcto** y evitar errores de interpretación

# Estructuras básicas de datos

- **Campos (columnas):** Representan características individuales de una entidad
- **Registros (filas):** Instancias únicas u observaciones de la entidad
- **Tabla:** Estructura que agrupa múltiples registros bajo un mismo conjunto de campos

# Identificadores

- **Clave primaria (Primary Key - PK):** Campo/s que garantizan la unicidad de cada registro

UNICA E INEQUIVOCA A UN ELEMENTO (REGISTRO) DE LA ENTIDAD

- 

- **Clave foránea (Foreign Key - FK):** Campo que establece una relación lógica con la clave primaria

ES UNA PRIMARY KEY EN OTRA TABLA HACE REFERENCIA A LOS VALORES RELACIONADOS DE OTRA ENTIDAD

# Modelo lógico de datos

Define **cómo se organizan y conectan los datos a nivel de diseño**.

Establece:

- 

- Qué **entidades (tablas)** existen
- Qué **atributos (columnas)** tiene cada una
- Qué **relaciones** hay entre ellas
- Qué **claves** las vinculan
- Qué **tipo de relación (cardinalidad)** corresponde

Es la base técnica del análisis y el paso previo a cualquier implementación física

# Tipos de tablas

- **Hechos:** contienen eventos medibles (Ventas, Transacciones)
- **Dimensiones:** describen los elementos del hecho (Cliente, Producto)
- **Puente:** resuelven relaciones muchos a muchos o jerarquías complejas entre dimensiones (Productos x Proveedores)



# ¿Qué tabla es correcta?

¿Porqué?



A

Producto	Enero	Febrero	Marzo
Avena	50	45	60
Granola	30	40	50

B

Producto	Mes	Ventas
Avena	Enero	50
Avena	Febrero	45
Avena	Marzo	60
Granola	Enero	30
Granola	Febrero	40
Granola	Marzo	50

**La tabla B es la correcta** desde el punto de vista de **estructura de datos** para análisis, especialmente en contextos como:

- Ciencia de datos
- Bases de datos relacionales

- Sirve para procesamiento, filtrado, agrupación, análisis y visualización.
- Cada fila representa una sola observación: un producto, en un mes, con su valor de ventas.
- Este formato facilita:
  - Agrupar por mes o producto
  - Realizar comparaciones o tendencias
  - Cargar en software de análisis sin necesidad de transformación previa

# Introducción a la IA y los datos

Pensamiento computacional

Conjunto de **habilidades para plantear problemas y definir soluciones** de manera que puedan ser comprendidas y ejecutadas tanto por personas como por sistemas informáticos.



# Secuencias, condiciones y bucles

**Secuencias:** Pasos que se ejecutan en orden para alcanzar un objetivo.

- Cada paso depende del anterior
- El orden afecta el resultado

**Condiciones:** Reglas que activan acciones según un criterio.

- Relación tipo “si... entonces...”
- Permiten decisiones en el flujo
- Deben cubrir todos los casos posibles

**Bucles:** Repiten instrucciones mientras se cumpla una condición.

- Requieren de inicio, condición de parada y actualización
- Usados para repetir procesos automáticamente
- Bucles infinitos si no están bien definidos

# Descomposición de problemas

¿Qué es?

Descomponer es **separar un problema complejo en partes** más simples, abordables y analizables.



Descomponer = separar

# Descomposición de problemas

¿Por qué es importante?

- Permite entender mejor el problema
- Ayuda a planificar el análisis
- Facilita la colaboración en equipos
- Permite aplicar herramientas digitales por etapas

# Descomposición de problemas

## Pasos

1. Definir el objetivo con un verbo
2. Listar subtareas y dependencias
3. Definir entradas y salidas de cada subtask
4. Estimar reglas y excepciones
5. Priorizar lo crítico



# Lupa logística

¿Cómo resolverías este problema?



A partir de la tabla, determina cuántos pedidos fueron enviados.

## Tareas

- Indicar la cantidad de pedidos enviados
- Detallar la secuencia, condición y bucle existente

ID	Pago	Stock
O-502	Aprobado	Sí
O-506	Pendiente	No
O-501	Aprobado	No
O-505	Anulado	Sí
O-503	Pendiente	Sí
O-504	Aprobado	No
O-508	Aprobado	No
O-507	Aprobado	Sí

## Pseudocodigo

contador = 0

lista\_enviados = []

para cada fila en tabla:

    if fila.Pago == "Aprobado" and fila.Stock == "Sí":

        contador = contador + 1

        lista\_enviados.append(fila.ID)

# Pseudocódigo

**Definición estructurada de algoritmos** en un formato independiente de cualquier lenguaje de programación.

Debe ser fácil de leer para las personas y sencillo de convertir en código ejecutable.

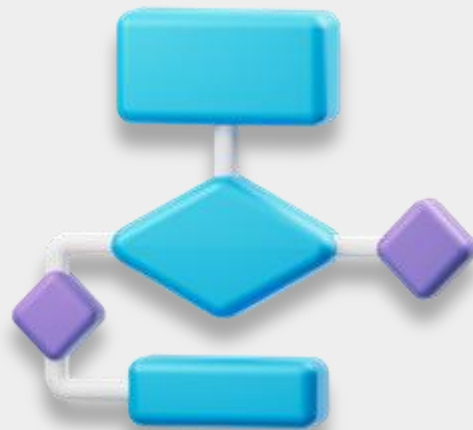
```
1  INICIO
2      LEER precio
3      SI precio > 1000 ENTONCES
4          MOSTRAR "producto caro"
5      SINO
6          MOSTRAR "Producto accesible"
7  FIN
```

## Sintaxis

- Bloques con sangría
- Palabras clave (INICIO/FIN, SI/ENTONCES/SINO)
- Entrada, Proceso, Salida
- Variables con nombres descriptivos
- Comentarios breves

# Diagrama de flujo

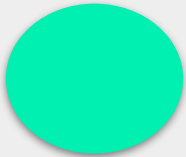
- **Permiten visualizar procesos y detectar huecos**
- Facilitan la comunicación de reglas con equipos
- Transforman pseudocódigos en símbolos conectados



# Diagrama de flujo

## Símbolos

Inicio/Fin



Forma: Elipses

Proceso



Forma: Rectángulo

Decisión



Forma: Rombos

Entrada/Salida



Forma: Paralelogramo

Conector



Forma: Líneas

# Diagrama de flujo

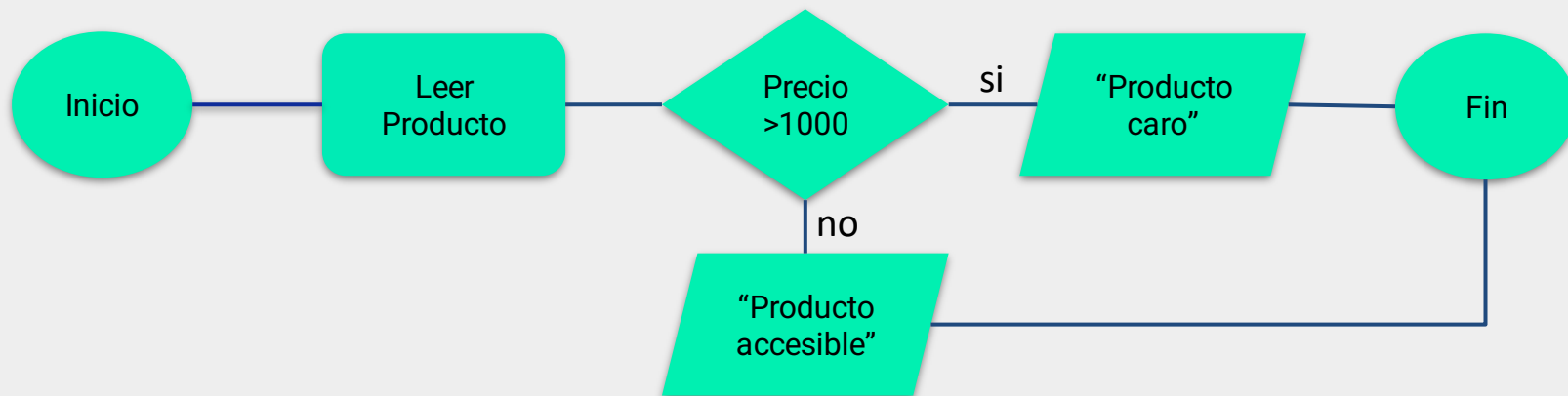
## Reglas

1. Una decisión = dos salidas (Sí/No)
2. Evitar cruces de flechas, usar conectores
3. Títulos breves por bloque
4. Mantener granularidad consistente

**Granularidad** es el nivel de detalle con que se describe una información

# Diagrama de flujo

## Ejemplo



# Diferencia entre Planillas de Calculo y Base de Datos

**Hojas de cálculo:** Lo mejor para conjuntos de datos pequeños y medianos; el rendimiento disminuye con grandes volúmenes de datos.

**Bases de datos:** Están diseñados para grandes conjuntos de datos y operaciones de alto rendimiento, incluso con millones de registros.



[illegible]

The image shows the top ribbon of the Microsoft Excel application. The ribbon is divided into several tabs: 'Inicio' (Home), 'Insertar' (Insert), 'Referencias' (References), 'Envío' (Mailings), 'Revisión' (Review), and 'Programas de complementos' (Add-ins). The 'Inicio' tab is currently selected and expanded, showing sub-sections for 'Fuente' (Font), 'Alineación' (Alignment), 'Número' (Numbers), 'Estilos' (Styles), 'Celdas' (Cells), 'Edición' (Editing), and 'Complementos' (Complements). Each sub-section contains various icons for formatting and editing text and cells. For example, the 'Fuente' section includes options for font face, size, bold, italic, underline, color, and background color. The 'Alineación' section includes options for text alignment, orientation, and wrapping. The 'Número' section includes options for number format, decimal places, and thousands separator. The 'Estilos' section includes options for conditional formatting, table styles, and cell styles. The 'Celdas' section includes options for inserting, deleting, and formatting cells. The 'Edición' section includes options for undo, redo, and find. The 'Complementos' section includes options for loading and managing add-ins.

[illegible]



# ¿Qué es la automatización?

Es el proceso de hacer que **tareas repetitivas o lógicas se realicen de forma automática, sin intervención manual.**

Esto se logra a partir de **reglas, condiciones o flujos** predefinidos que las herramientas tecnológicas pueden ejecutar por sí mismas.



# ¿Qué es la automatización?

## Ejemplos



### Python

Escribir un script que limpie un conjunto de datos y lo deje listo para el análisis.

### Power BI

Programar la actualización automática de un dashboard para mantener la información siempre al día.

# ¿Qué se puede automatizar en la IA?

- Limpieza y procesamiento de datos
- Aplicación de reglas y condiciones
- Ejecución de cálculos
- Predicciones y proyecciones
- Generación de reportes o resultados



**IA y automatización se complementan** para mejorar  
procesos y decisiones

# Despacho inteligente



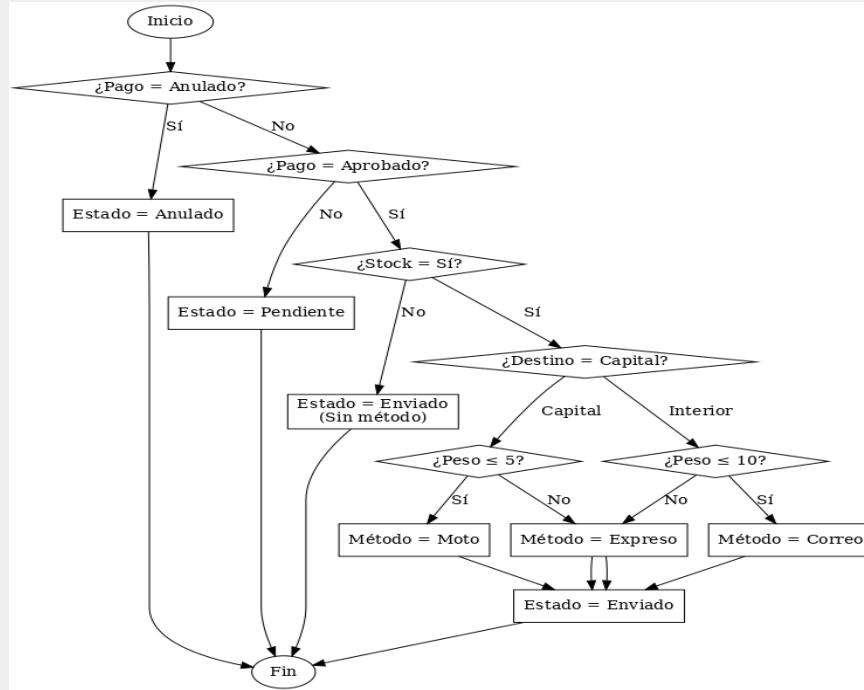
Se debe calcular el recuento de pedidos por estado y método de envío, realiza para ello un pseudocódigo y diagrama de flujo y ten en consideración las siguientes reglas:

- Si Pago = Anulado >> Estado = Anulado
- Si Pago  $\neq$  Aprobado >> Estado = Pendiente
- Si Pago = Aprobado y Stock = No >> Estado = Enviado
- Si Pago = Aprobado y Stock = Sí >> Método de envío:
  - Moto si Destino = Capital y Peso  $\leq 5$
  - Correo si Destino = Interior y Peso  $\leq 10$
  - Expreso en cualquier otro caso

ID	Pago	Stock	Destino	Peso
O-702	Pendiente	Sí	Interior	7
O-708	Aprobado	Sí	Interior	10
O-705	Aprobado	No	Capital	2
O-701	Aprobado	Sí	Capital	3
O-703	Aprobado	Sí	Interior	8
O-707	Aprobado	Sí	Capital	6
O-704	Aprobado	Sí	Interior	12
O-706	Anulado	Sí	Capital	1



```
para cada pedido en tabla:
  si pago == "Anulado":
    estado = "Anulado"
  sino si pago != "Aprobado":
    estado = "Pendiente"
  sino si pago == "Aprobado" y stock == "No":
    estado = "Enviado"
  sino si pago == "Aprobado" y stock == "Sí":
    estado = "Enviado"
    si destino == "Capital" y peso <= 5:
      metodo = "Moto"
    sino si destino == "Interior" y peso <= 10:
      metodo = "Correo"
    sino:
      metodo = "Expreso"
```





# Retro

¿Cómo nos vamos?

- ¿Qué fue lo más útil de la clase?
- ¿Qué parte te costó más?
- ¿Qué te gustaría repasar o reforzar?