

## Tema 2. Introducción al Análisis de Datos y Microsoft Excel

### 2.1 Dato, información y conocimiento.

Generalmente, los términos dato, información y conocimiento se utilizan de forma indistinta en ambientes no formales. Sin embargo, los ámbitos más formales (el profesional y el académico) requieren de una distinción precisa de estos conceptos. De esta forma, se evitan malas interpretaciones durante las distintas fases del análisis de datos.

Tal y como puede esperarse, son múltiples las aproximaciones que se utilizan para la distinción de estos términos. Seguiremos la proporcionada por Davenport y Prusak (1998).

Un **dato** puede definirse como un hecho concreto y discreto acerca de un evento. La característica de ser discreto implica que, semánticamente, es la unidad mínima que puede comunicarse o almacenarse. Por sí solos, los datos no brindan detalles significativos del entorno del que fueron obtenidos.

Ejemplos de datos pueden ser:

- 2019.
- DPS.

La **información** puede definirse como un mensaje formado por la composición de varios datos. A diferencia del dato, la información sí posee un significado para un receptor u observador.

Por ejemplo, utilizando los datos anteriores se podría obtener la siguiente información:

- A finales del año 2019, surgió la pandemia de COVID-19 en China, que se extendería por todo el mundo en el 2020, teniendo un impacto global sin precedentes en la salud pública, la economía y la sociedad.
- El código del aeropuerto de Bali es DPS. Este código es utilizado por la Asociación de Transporte Aéreo Internacional (IATA) para identificar de manera única al Aeropuerto.

Los datos deben ser transformados para añadirles valor y convertirlos en información. Estas transformaciones incluyen métodos como:

- Contextualización: Conocer el propósito del dato obtenido.
- Categorización: Conocer la unidad de medida y los componentes del dato.
- Cálculo: Realizar una operación matemática sobre el dato.
- Corrección: Eliminar errores del dato.

- Agregación: Resumir o minimizar un dato de forma más concisa.

Por último, el **conocimiento** implica una combinación de experiencias, información contextual y relevancia sobre cierta información.

Así como la información se genera a partir de datos, el conocimiento surge de la agregación de información.

Ejemplos de métodos que generan esta transformación son:

- Comparación: Relación entre información obtenida en distintas experiencias.
- Repercusión: Implicación de la información en decisiones y acciones.
- Conexión: Relación entre distintos tipos de información.
- Conversación: Opinión de otras personas sobre la información.

La jerarquía del conocimiento suele representarse gráficamente por una pirámide en la que los datos suelen ser la base, mientras que el conocimiento se identifica con la cima.



## 2.2 Generalidades.

### *¿Qué es el análisis de datos?*

El análisis de datos se presenta como un enfoque integral que va más allá de la simple manipulación de números y cifras. Se trata de una travesía hacia el descubrimiento de información valiosa.

Este proceso multifacético se compone de diversas técnicas y metodologías, cada una con su propia lupa para observar los datos desde diferentes ángulos. Desde el análisis estadístico tradicional hasta las sofisticadas herramientas de inteligencia artificial, el analista

de datos tiene a su disposición un vasto repertorio para interpretar datos provenientes de diversas fuentes y formatos.

El análisis de datos no solo transforma la forma en que interpretamos el mundo, sino que también revoluciona la forma en que tomamos decisiones. Al basarnos en información sólida y evidencia concreta, podemos navegar con mayor seguridad por la incertidumbre, construyendo un futuro más informado, próspero y sostenible.

### ***La importancia del análisis de datos en el mundo digital actual***

En la era de la transformación digital, el análisis de datos ha trascendido su rol tradicional para convertirse en un motor fundamental del cambio. La explosión de datos generada por las tecnologías digitales ha dado lugar a un nuevo paradigma. Esta ingente cantidad de información, si se procesa y analiza adecuadamente, posee un potencial transformador sin precedentes para las empresas.

### ***El análisis de datos en acción***

Empresas de diversos sectores están aprovechando el poder del análisis de datos para transformar sus negocios y obtener ventajas competitivas.

Algunos ejemplos:

- e-commerce: Utilizan el análisis de datos para personalizar las recomendaciones de productos, optimizar la logística y mejorar la experiencia del cliente.
- Plataformas de streaming: Analizan los datos de visualización para recomendar películas y/o series que se ajusten a los gustos de cada usuario.
- Bancos: Utilizan el análisis de datos para detectar fraudes, evaluar la solvencia de los clientes y ofrecer productos financieros personalizados.
- Hospitales: Analizan los datos de pacientes para mejorar los diagnósticos, desarrollar tratamientos personalizados y optimizar la gestión de recursos.
- Redes sociales: Analizan datos de usuarios para personalizar el contenido que se muestra, y con base en ello, recomendar amigos, publicaciones y anuncios.

### ***El futuro del análisis de datos: Un horizonte de posibilidades***

El análisis de datos seguirá evolucionando a medida que se desarrollen nuevas tecnologías, tales como la inteligencia artificial y el aprendizaje automático. Las empresas que integren estas tecnologías en sus estrategias de análisis de datos estarán mejor posicionadas para liderar la transformación digital y alcanzar el éxito en el futuro.

## **2.3 Proceso de análisis de datos.**

### ***1.- Definición de objetivos y preguntas.***

En esta etapa inicial, se establecen los objetivos que se pretenden alcanzar con el análisis de datos. Esto implica definir claramente las preguntas que se buscan responder con los datos. Las preguntas deben ser específicas, medibles, alcanzables, relevantes y con un plazo determinado. Comprender el problema, identificar datos necesarios y definir parámetros son elementos clave.

## **2.- Recolección de datos.**

La recolección de datos implica identificar y reunir los datos necesarios para responder a las preguntas planteadas. Los datos pueden ser cuantitativos o cualitativos, dependiendo de la naturaleza del problema.

Esto puede involucrar diferentes fuentes de datos que pueden ser primarias, como experimentos o sensores propios, o secundarias, como datasets existentes o encuestas realizadas por terceros. Para optimizar la recolección, se emplean mecanismos de muestreo, como el muestreo aleatorio simple, estratificado o conglomerados. Estos métodos garantizan la representatividad de los datos y evitan sesgos.

Es crucial asegurar la calidad de los datos recolectados, verificando su precisión, completitud y consistencia. Por ello, es importante llevar a cabo una correcta interpretación de los datos, tomando en cuenta los siguientes aspectos:

- Contextualización: Entender el contexto en el que se recolectaron los datos es fundamental para interpretarlos correctamente, considerando factores económicos, sociales, políticos y tecnológicos.
- Crítica de Métodos: Evaluar críticamente los métodos de análisis utilizados, incluyendo la adecuación del modelo estadístico, la precisión de los algoritmos y la integridad de los datos.
- Consideración de Limitaciones: Reconocer las limitaciones del proceso de análisis, ya sean relacionadas con la calidad del dato, la metodología o las inferencias estadísticas.

## **3.- Depuración de datos.**

La depuración de datos, también conocida como limpieza de datos, es el proceso de identificar y corregir errores en los datos recolectados. Esto puede incluir eliminar valores duplicados, corregir errores de formato, completar datos faltantes y tratar valores atípicos. Es crucial para garantizar la confiabilidad de los análisis posteriores.

## **4. Análisis de datos.**

En esta etapa, se aplican técnicas estadísticas y computacionales para analizar los datos depurados. Esto puede incluir análisis descriptivo, análisis exploratorio de datos, modelado estadístico, aprendizaje automático y minería de datos. La elección de las técnicas dependerá del tipo de datos, las preguntas planteadas y los objetivos del análisis.

## **5.- Interpretación y visualización de los datos.**

Una vez finalizado el análisis, se procede a interpretar los resultados obtenidos. Esto implica comprender el significado de los patrones, tendencias y relaciones identificadas en los datos. La visualización de datos juega un papel crucial en esta etapa, ya que permite presentar los resultados de manera clara, concisa y convincente, facilitando su comprensión e interpretación.

## **6. Narración de datos:**

La narración de datos es el arte de contar una historia con datos. Se trata de comunicar los resultados del análisis de una manera atractiva y persuasiva, utilizando técnicas de narración y storytelling. La narración de datos efectiva permite conectar con la audiencia, transmitir insights relevantes y generar acciones a partir de los datos.

El proceso del análisis de datos no es lineal, sino que implica una iteración constante entre las diferentes etapas. A medida que se avanza en el análisis, es posible que surjan nuevas preguntas o que sea necesario regresar a etapas anteriores para refinar los datos o el análisis. La clave para un análisis de datos exitoso es mantener una actitud curiosa y flexible, adaptándose a los hallazgos y aprendiendo de los datos en cada etapa del proceso.

## **2.4 Tipos de Análisis de Datos.**

Existen cuatro tipos fundamentales de análisis de datos, cada uno con un propósito único y perspectivas distintas.

### **1. Análisis descriptivo.**

El análisis descriptivo se centra en resumir y caracterizar los datos de manera sencilla y comprensible. Responde a preguntas como: ¿Qué ha pasado? ¿Cuál es la situación actual? ¿Cómo se distribuyen los datos?

*Técnicas:* Estadísticas básicas (media, mediana, moda, desviación estándar), tablas de frecuencias, gráficos (histogramas, diagramas de barras, diagramas de dispersión).

*Utilidad:* Obtener una visión general de los datos, identificar patrones básicos y tendencias generales.

*Ejemplo:* Analizar las ventas de una empresa en los últimos años para identificar los productos más vendidos y las tendencias de compra.

### **2. Análisis diagnóstico.**

El análisis diagnóstico profundiza en los datos para comprender las causas de los patrones y tendencias observados en el análisis descriptivo. Responde a preguntas como: ¿Por qué ha sucedido esto? ¿Cuáles son los factores que influyen en los resultados? ¿Qué relaciones existen entre las variables?

*Técnicas:* Análisis de varianza, pruebas de hipótesis, minería de datos.

*Utilidad:* Identificar las causas fundamentales de los problemas o desviaciones, comprender mejor el comportamiento de los datos y tomar decisiones más informadas.

*Ejemplo:* Analizar las causas de la baja tasa de retención de clientes para identificar los factores que contribuyen a la pérdida de clientes y desarrollar estrategias para mejorar la retención.

### **3. Análisis predictivo.**

El análisis predictivo utiliza modelos estadísticos y técnicas de aprendizaje automático para predecir eventos futuros o tendencias probables. Responde a preguntas como: ¿Qué es probable que suceda en el futuro? ¿Cuáles son los escenarios más probables? ¿Cómo podemos anticiparnos a los cambios?

*Técnicas:* Regresión lineal, árboles de decisión, redes neuronales artificiales.

*Utilidad:* Anticipar eventos futuros, predecir la demanda de productos o servicios, optimizar la gestión de riesgos y tomar decisiones estratégicas proactivas.

*Ejemplo:* Analizar los datos históricos de ventas para predecir la demanda futura de productos y optimizar la producción y el inventario.

### **4. Análisis prescriptivo.**

El análisis prescriptivo va más allá de la predicción y recomienda acciones para alcanzar objetivos específicos o mejorar resultados. Responde a preguntas como: ¿Qué podemos hacer para mejorar la situación? ¿Cuáles son las mejores alternativas? ¿Cómo podemos optimizar nuestros procesos?

*Técnicas:* Optimización matemática, simulación, análisis de costo-beneficio.

*Utilidad:* Tomar decisiones óptimas para alcanzar objetivos específicos, optimizar procesos, asignar recursos de manera eficiente y maximizar el valor.

*Ejemplo:* Analizar los datos de atención al cliente para identificar las áreas de mejora y desarrollar estrategias para optimizar la experiencia del cliente y reducir los costos.

Los cuatro tipos de análisis de datos se complementan entre sí, proporcionando una visión completa y profunda de la información. El análisis descriptivo ofrece una base fundamental, el análisis diagnóstico profundiza en las causas, el análisis predictivo anticipa el futuro y el análisis prescriptivo recomienda acciones para alcanzar objetivos.

La elección del tipo de análisis dependerá de las preguntas específicas que se busquen responder y los objetivos que se pretenden alcanzar.

## **2.5 Ética de datos.**

La ética de datos es un campo que estudia los aspectos morales del uso y manejo de la información personal en la era digital. Se centra en proteger los derechos individuales y garantizar un uso responsable de los datos.

Una de las consideraciones éticas más importantes cuando se trabaja con datos es la **privacidad**. La privacidad es un tema complicado, con muchas normas y reglamentos que dependen del contexto y del sector.

En términos generales, la privacidad es el derecho de una persona a controlar cómo se comparte la información sobre sí misma. Pero siempre hay que tener en cuenta que tus expectativas de privacidad pueden ser diferentes de las expectativas de privacidad de otras personas.

Otra consideración ética importante, relacionada con la privacidad, es el **consentimiento**. Una parte importante de la privacidad es poder controlar lo que ocurre con nuestra información. El consentimiento debe darse de forma clara e inequívoca por parte del usuario interesado, que debe haber recibido información detallada acerca de la gestión de sus datos por parte del servicio o página online.

Podemos ejercer control sobre nuestra información:

Dando nuestro consentimiento, o permiso, para que nuestra información se utilice de una determinada manera.

Al igual que la privacidad, el concepto de consentimiento tiene una historia diferente dependiendo el contexto.

#### *Importancia:*

La ética de datos es fundamental para proteger la privacidad individual y garantizar un uso responsable de la información en la era digital. Es crucial para:

- Construir confianza: Las personas deben confiar en que sus datos serán utilizados de manera responsable y ética.
- Fomentar la innovación: Un marco ético claro puede impulsar la innovación responsable en el uso de datos.
- Proteger los derechos: La ética de datos protege los derechos fundamentales de las personas, como la privacidad y la libertad de expresión.

#### *Desafíos:*

La aplicación de la ética de datos presenta desafíos, como:

- **Falta de conciencia:** Muchas personas no son conscientes de los riesgos y las implicaciones del uso de datos.
- **Tecnologías en constante evolución:** Es difícil establecer regulaciones que se adapten al ritmo del cambio tecnológico.
- **Aplicación desigual:** La aplicación de la ética de datos puede variar entre países y regiones.

Checklist para cumplir con la ética de datos:

a) Revisión de Datos Personales:

- Examina el conjunto de datos en busca de información personal identificable (IPI), como nombre, dirección, número de Seguridad Social, fecha de nacimiento, entre otros.
- Si se encuentra IPI, elimínalo o anonimízalo, y gestiona el acceso de manera controlada.

b) Investigación sobre Recopilación de Datos:

- Indaga cómo se obtuvo el conjunto de datos.
- Pregunta si las personas representadas dieron su consentimiento para ser incluidas y para el uso de su información de esa manera.
- En ausencia de consentimiento confirmado, busca otra fuente de datos o contacta a las personas para obtener su consentimiento informado.

En conclusión, la ética de datos es un campo complejo y en constante evolución que requiere un esfuerzo continuo para garantizar un equilibrio entre el conocimiento y la privacidad en la era digital.

## 2.6 ¿Cómo iniciarse en el mundo del análisis de datos?

El análisis de datos se ha convertido en una habilidad altamente demandada en la actualidad, abriendo puertas a un sinfín de oportunidades profesionales.

Para prosperar en el análisis de datos, se debe construir una base de conocimientos, perfeccionar las habilidades prácticas y acumular experiencia.

### *1. Comprende los conceptos básicos:*

Antes de sumergirte en herramientas y técnicas, es fundamental establecer una base sólida en los conceptos fundamentales del análisis de datos. Esto incluye:

- Tipos de datos: Familiarízate con los diferentes tipos de datos, como cuantitativos, cualitativos, estructurados y no estructurados.
- Medidas estadísticas: Aprende sobre medidas de tendencia central (media, mediana, moda), dispersión (rango, desviación estándar) y asociación (correlación, regresión).
- Proceso del análisis de datos: Comprende las etapas clave del análisis de datos, desde la definición de objetivos hasta la comunicación de resultados.

### *2. Aprende un lenguaje de programación:*



Los lenguajes de programación son herramientas esenciales para manipular, analizar y visualizar datos. Algunos de los lenguajes más populares para el análisis de datos incluyen:

- SQL: Un lenguaje fundamental para interactuar con bases de datos relacionales y extraer información relevante.
- Python: Un lenguaje versátil y fácil de aprender, con una amplia comunidad y bibliotecas especializadas en análisis de datos.
- R: Un lenguaje orientado al análisis estadístico y la creación de gráficos, ampliamente utilizado en investigación científica.

### *3. Domina la manipulación y visualización de datos:*

La capacidad de manipular y visualizar datos de manera efectiva es crucial para comunicar insights de manera clara y convincente. Aprende a:

- Limpiar y preparar datos: Elimina errores, inconsistencias y valores atípicos para garantizar la calidad de los datos.
- Transformar datos: Cambia la estructura o formato de los datos para adaptarlos a las necesidades del análisis.
- Crear visualizaciones: Utiliza gráficos, tablas y mapas para comunicar patrones, tendencias y relaciones en los datos.

### *4. Práctica:*

La mejor manera de dominar el análisis de datos es poner en práctica tus conocimientos. Cuanto más practiques, más habilidades y confianza desarrollarás.

## **MICROSOFT EXCEL**

### **2.7 ¿Qué es Excel?**

Excel es un software de hojas de cálculo desarrollado por Microsoft y utilizado por empresas de todo el mundo. Es una herramienta poderosa y versátil diseñada para usuarios técnicos y no técnicos, ya que ofrece una interfaz fácil de usar y facilita el almacenamiento, procesamiento, análisis y visualización de datos.

Excel permite a los usuarios:

- Crear y administrar tablas de datos: Con filas y columnas para almacenar y organizar datos numéricos, textuales y de fecha.
- Realizar cálculos matemáticos y estadísticos: Fórmulas y funciones integradas para realizar cálculos complejos, análisis de datos y generar informes.
- Crear gráficos y visualizaciones: Amplia variedad de gráficos, tablas dinámicas y herramientas de visualización para presentar datos de manera clara y atractiva.
- Automatizar tareas: Macros y herramientas de programación para automatizar tareas repetitivas y ahorrar tiempo.

- Colaborar con otros: Compartir y editar hojas de cálculo con otros usuarios en tiempo real.

## Usos Principales de Excel

### ***En contextos Profesionales***

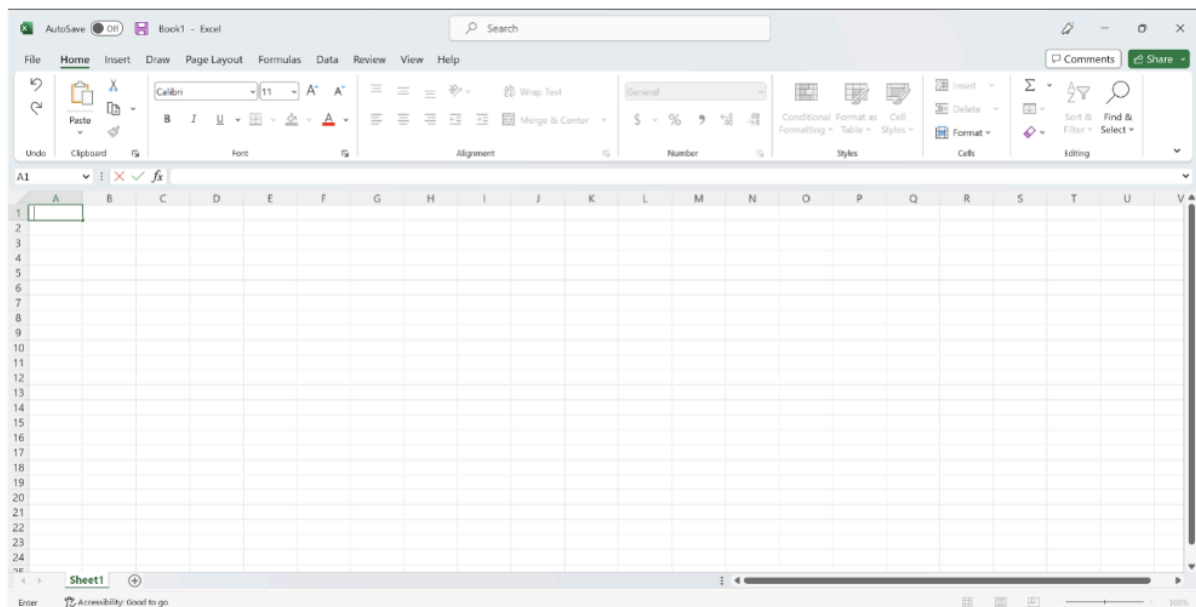
- Finanzas: Análisis de presupuestos, estados financieros y previsiones.
- Recursos Humanos: Seguimiento de empleados, gestión de nóminas y análisis de beneficios.
- Marketing: Análisis de campañas, seguimiento de ventas y estudios de mercado.
- Operaciones: Gestión de inventarios, programación de producción y análisis de logística.

### ***En contextos Académicos***

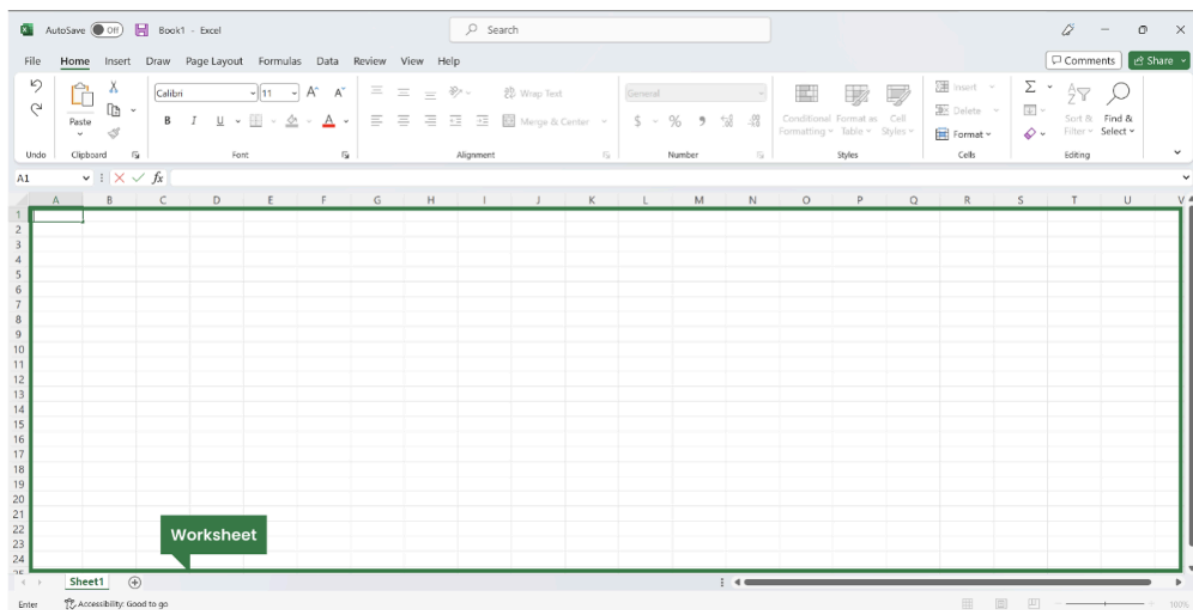
- Investigación: Análisis de datos experimentales, estadísticas y gráficos.
- Administración: Planificación de horarios y gestión de proyectos.
- Enseñanza: Creación de materiales didácticos, seguimiento del rendimiento de los estudiantes y organización de información.

## 2.8 Terminología básica.

**Libro:** un archivo Excel que contiene una o varias hojas de cálculo que ayuda a organizar los datos.



**Hoja de trabajo/cálculo:** área principal en la que trabajará, consta de celdas organizadas en columnas y filas.



**Filas:** Son las líneas horizontales de la hoja de cálculo. y se numeran consecutivamente desde 1 hasta 1,048,576. En Excel 2016 y versiones posteriores, cada fila se identifica por un número único ubicado en el extremo izquierdo de la hoja.

Las filas se utilizan para agrupar los datos de manera horizontal. Cada fila puede contener datos relacionados que se ingresan en las celdas de esa fila (en el análisis de datos esto se conoce como “registros”). Por ejemplo, en una lista de empleados, cada fila podría representar un empleado diferente.

**Columnas:** Son las líneas verticales y se etiquetan con letras comenzando desde A hasta Z. Después de la Z, las columnas continúan como AA, AB, y así sucesivamente, llegando hasta XFD en Excel 2016 y versiones posteriores.

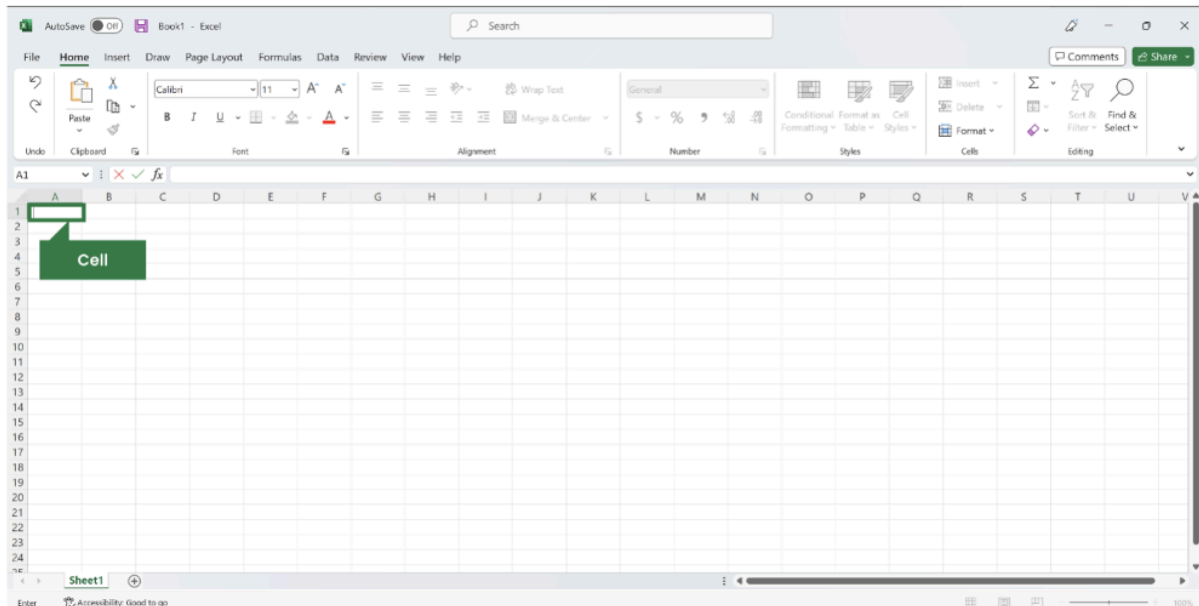
Las columnas se utilizan para organizar los datos de manera vertical. Cada columna puede contener un tipo específico de datos a lo largo de las diferentes filas (en el análisis de dato esto se conoce como “campos”). Por ejemplo, en la misma lista de empleados, cada columna podría representar una diferente característica del empleado, como nombre, apellido, departamento, etc.

**Celda:** intersección de una fila y una columna, toda la información de una hoja de cálculo se introduce en las celdas.

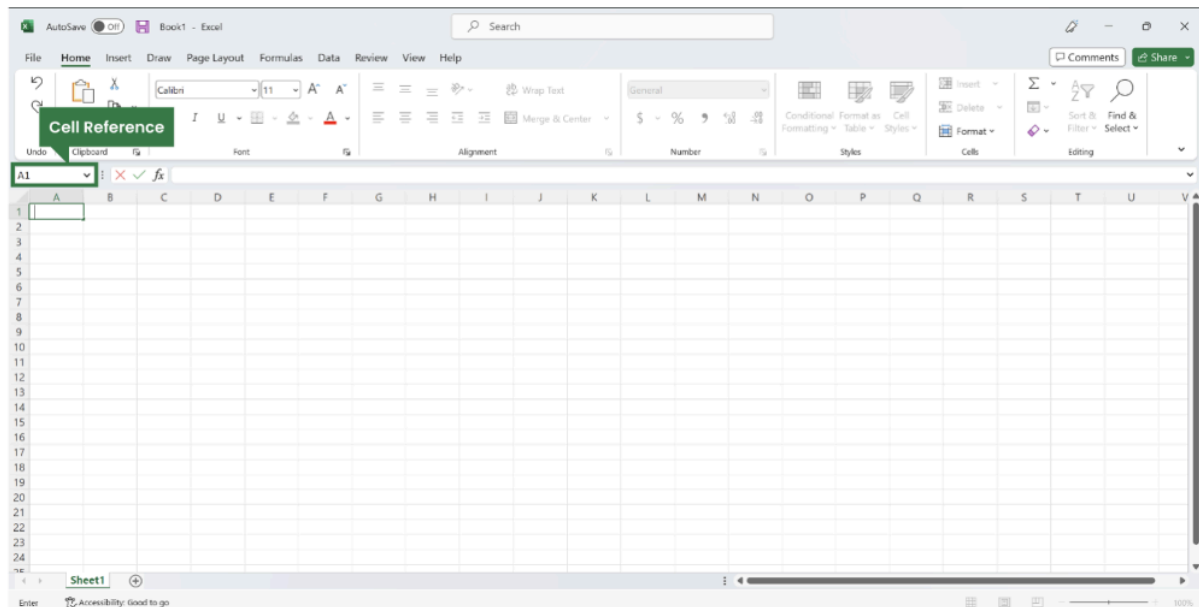
### **Utilización de Celdas**

- **Seleccionar una Celda:** Haciendo clic en una celda específica. La celda seleccionada se resalta y su referencia aparece en el cuadro de nombre, situado en la esquina superior izquierda.
- **Ingresar Datos:** Haciendo doble clic en una celda o seleccionándola y comenzando a escribir. Una vez ingresados los datos, se puede presionar Enter para mover a la siguiente fila o Tab para moverse a la siguiente columna.
- **Editar Datos:** Haciendo doble clic en una celda con datos.

- **Rango de Celdas:** Un grupo de celdas seleccionadas que se identifican por la referencia de la celda superior izquierda y la celda inferior derecha, separadas por dos puntos. Por ejemplo, A1:B10 incluye todas las celdas desde A1 hasta B10.
- **Copiar y Pegar:** Seleccionando una celda o rango de celdas, utilizando Ctrl+C para copiar y Ctrl+V para pegar en otra ubicación.
- **Arrastrar y Soltar:** Seleccionando una celda o rango de celdas y arrastrando desde el borde para mover los datos a una nueva ubicación.



**Referencia de celda:** combinación de una letra de columna y un número de fila que identifica una celda de una hoja de cálculo.



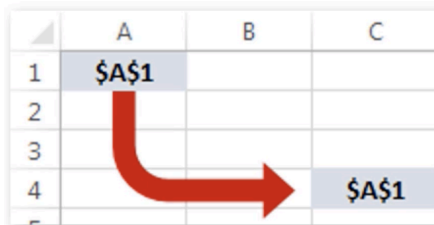
Quando se crea una fórmula, se usan referencias de celdas en vez de valores, así que, cuando se hacen cambios en las celdas, también cambia el resultado final.

Existen dos tipos de referencias: relativas y absolutas. Las referencias relativas cambian cuando se copian de una celda a otra. Por otra parte, las referencias absolutas se mantienen constantes sin importar la celda en que se copien.

Por default, las celdas son **relativas**, por lo tanto, cuando se copia una fórmula de una celda a otra, la fórmula cambia con base en la posición de las filas y las columnas. Si se copia una fórmula como =C1 + D1 de la fila 1 a la fila 2, la fórmula se transformará en =C2 + D2.

En ocasiones no se desea un cambio en la fórmula al copiarla de una a otra celda y en esos casos se usan las referencias **absolutas**. Para crear una referencia absoluta, basta con añadir el símbolo (\$). Dependiendo de cuál es la parte de la fórmula que no se desea que cambie, se procede de la siguiente manera:


- \$A\$1: ni la columna ni la fila cambia.



A diagram illustrating an absolute reference. It shows a grid with columns A, B, and C, and rows 1, 2, 3, and 4. A red arrow originates from cell A1, which contains the text '\$A\$1', and points to cell C4, which also contains the text '\$A\$1'. This demonstrates that the reference remains constant regardless of the destination cell.

|   | A      | B | C      |
|---|--------|---|--------|
| 1 | \$A\$1 |   |        |
| 2 |        |   |        |
| 3 |        |   |        |
| 4 |        |   | \$A\$1 |

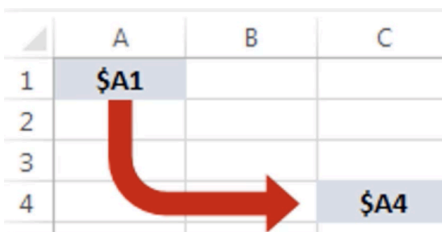
- A\$1: la fila no cambia.



A diagram illustrating a row absolute reference. It shows a grid with columns A, B, and C, and rows 1, 2, 3, and 4. A red arrow originates from cell A1, which contains the text 'A\$1', and points to cell C4, which contains the text 'C\$1'. This demonstrates that the row number remains constant while the column letter changes.

|   | A    | B | C    |
|---|------|---|------|
| 1 | A\$1 |   |      |
| 2 |      |   |      |
| 3 |      |   |      |
| 4 |      |   | C\$1 |

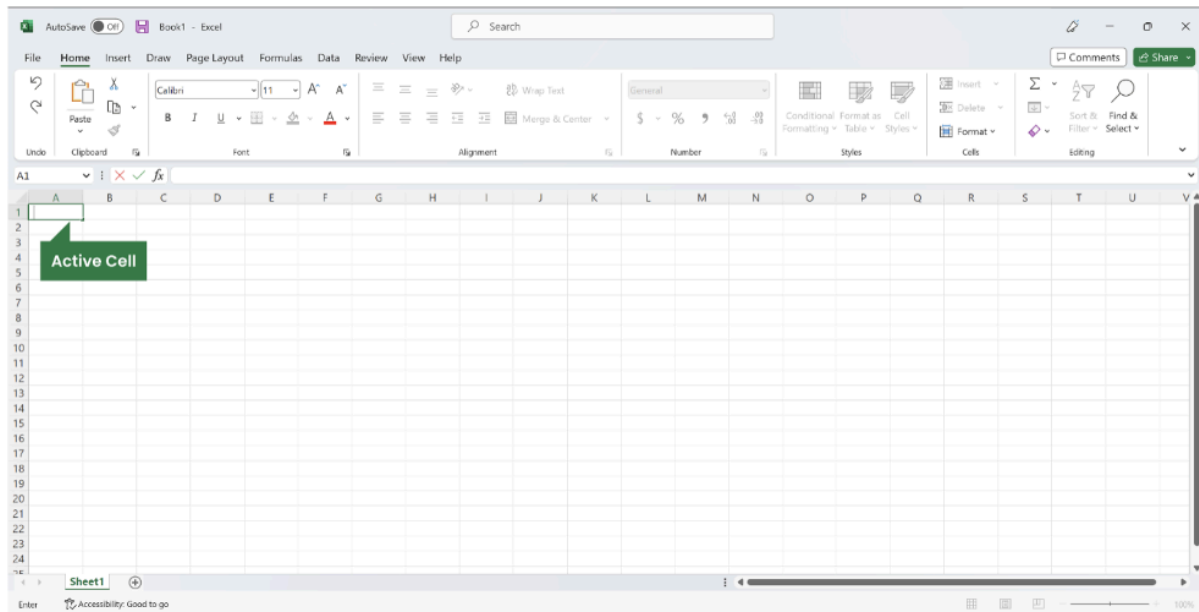
- \$A1: la columna no cambia.



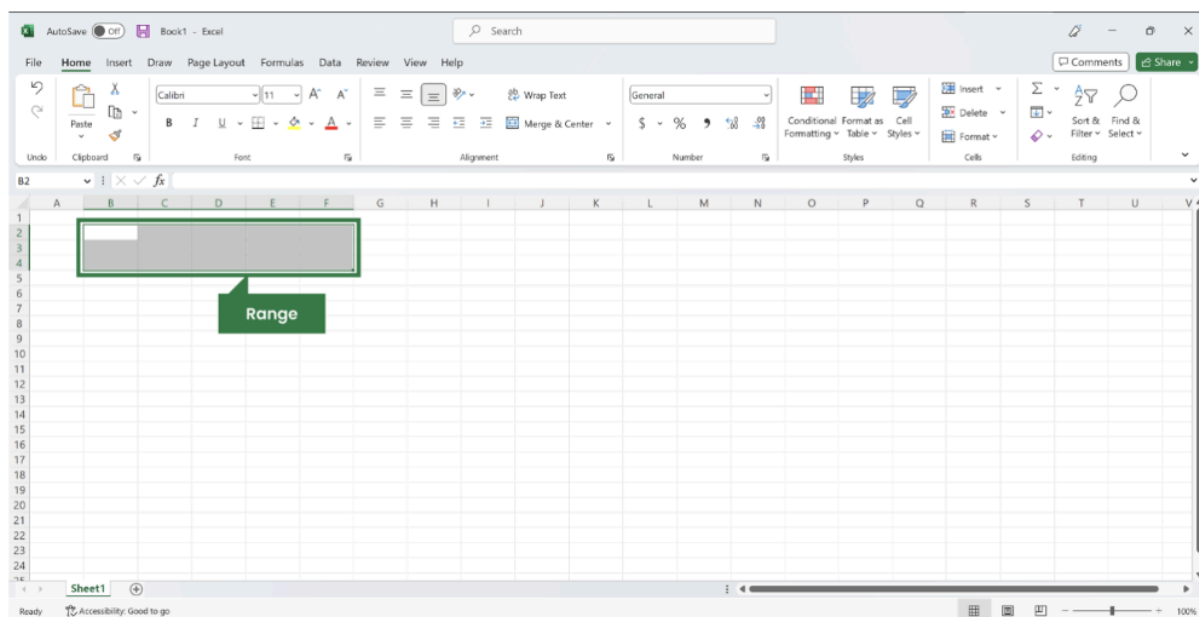
A diagram illustrating a column absolute reference. It shows a grid with columns A, B, and C, and rows 1, 2, 3, and 4. A red arrow originates from cell A1, which contains the text '\$A1', and points to cell C4, which contains the text '\$A4'. This demonstrates that the column letter remains constant while the row number changes.

|   | A    | B | C    |
|---|------|---|------|
| 1 | \$A1 |   |      |
| 2 |      |   |      |
| 3 |      |   |      |
| 4 |      |   | \$A4 |

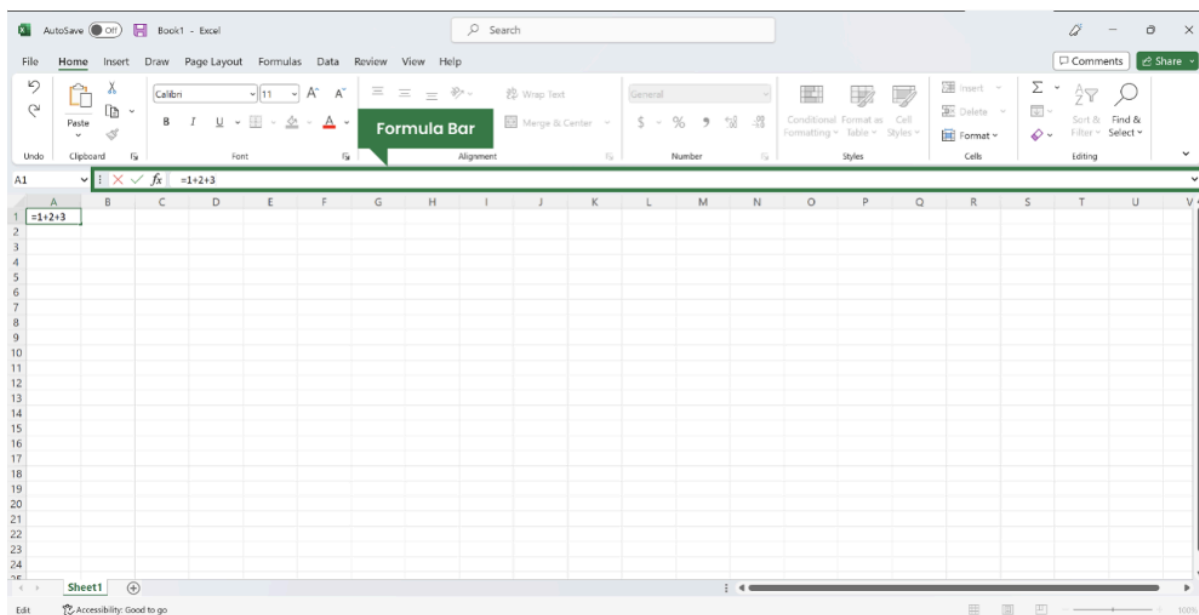
**Celda activa:** se refiere a la celda con la que se está trabajando.



**Rango:** se refiere a dos o más celdas contiguas de una hoja de cálculo.



**Barra de fórmulas:** espacio en el que se ven los datos o las fórmulas contenidas en la celda activa.



## 2.9 Funcionalidades básicas.

**Fórmulas:** Ecuaciones que realizan cálculos de los valores en la hoja, todas las fórmulas comienzan con un signo igual (=).

**Funciones:** Fórmulas predefinidas que te permiten realizar cálculos complejos y automatizados. Son herramientas esenciales para ahorrar tiempo, reducir errores y aumentar la eficiencia al analizar datos.

Las funciones se componen de:

- **Nombre:** Identifica la función, por ejemplo, SUMA, PROMEDIO, SI.
- **Argumentos:** Valores o referencias de celda que se proporcionan a la función para realizar el cálculo. Se escriben entre paréntesis, separados por comas.
- **Sintaxis:** La estructura correcta en la que se escribe la función, incluyendo el nombre, los paréntesis y los argumentos.

*Sintaxis*

**=COINCIDIR** (valor\_buscado,matriz\_buscada, [tipo\_de\_coincidencia])

**Nombre de la función:** Le dice a Excel que tipo de función estás a punto de ejecutar. Excel ofrece aproximadamente 500 funciones diferentes.

- 1.- Todas las funciones empiezan con un signo de =.
- 2.- Las funciones **no** son *case sensitive* (No tiene importancia escribir los caracteres en mayúsculas o minúsculas).

**Argumentos:** Los argumentos varían dependiendo de la función a utilizar, y estos son los que le proveen a Excel la información necesaria para arrojar un resultado.

\*No todos los argumentos son obligatorios, aquellos que están entre corchetes [ ] son opcionales. Casi todas las funciones tienen al menos un argumento obligatorio.

1. Los argumentos van entre paréntesis y están separados por comas.

*Tipos de funciones:*

Excel ofrece una amplia variedad de funciones categorizadas por su propósito:

- Matemáticas: SUMA, PROMEDIO, PRODUCTO, POTENCIA, RAÍZ CUADRADA, etc.
- Estadísticas: CONTAR, MEDIANA, MODA, DESVIACIÓN ESTÁNDAR, etc.
- Texto: CONCATENAR, LARGO, MAYÚSCULAS, MINÚSCULAS, etc.
- Fecha y hora: HOY, AHORA, FECHA, HORA, etc.
- Lógicas: SI, Y, O, NO, etc.
- Búsqueda y referencia: BUSCAR, BUSCARV, INDICE, COINCIDIR, etc.
- Financieras: PAGO, TIR, NPV, etc.

*Ventajas del uso de funciones:*

- Ahorran tiempo: Automatizan cálculos repetitivos.
- Reducen errores: Minimizan la posibilidad de errores manuales.
- Aumentan la eficiencia: Simplifican el análisis de datos complejos.
- Agregan flexibilidad: Permiten realizar cálculos personalizados.

**Formato de celdas:** Excel permite formatear las celdas para diferentes tipos de datos, como texto, números, fechas y monedas.

**Estilos y Temas:** Se pueden aplicar diferentes estilos y temas para mejorar la presentación visual de los datos.

**Ordenar:** Permite organizar los datos en una columna de manera ascendente o descendente.

**Filtrar:** Facilita la visualización de datos específicos mediante la aplicación de criterios de filtrado.

**Gráficos:** Excel ofrece diversos tipos de gráficos (barras, líneas, circulares, de superficie, etc.) para visualizar datos de manera efectiva.

**Tablas:** Herramienta que facilita la organización y el análisis de datos mediante funcionalidades adicionales como filtrado y ordenación.

**Tablas Dinámicas:** Permiten resumir, analizar, explorar y presentar datos de manera interactiva. Son ideales para analizar grandes volúmenes de información.

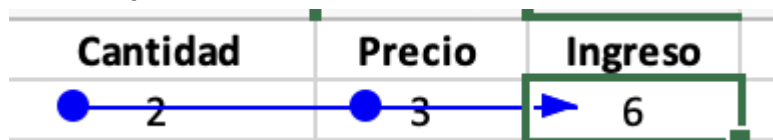


## 2.10 Tipos de errores comunes en Excel.

| Error    | ¿Qué significa?                                                                     | ¿Cómo arreglarlo?                                                                                                                                              |
|----------|-------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------|
| #####    | La columna no está lo suficientemente ancha para mostrar los valores.               | Ensanchando la columna:<br>1.- Dando doble click en el borde de ésta para que se ajuste automáticamente.<br>2.- Arrastrando el borde para hacerlo manualmente. |
| #NOMBRE? | Excel no reconoce el texto en una fórmula, ya sea en el nombre o en los argumentos. | Asegurarse que el nombre de la función sea correcto y los argumentos tengan la gramática adecuada.                                                             |
| #VALOR!  | La fórmula tiene argumentos erróneos.                                               | Asegurarse que los argumentos sean del tipo correcto.                                                                                                          |
| #DIV/0   | La fórmula está dividida entre 0 o entre una celda vacía.                           | Checar bien los valores, y en caso de querer dividir entre cero se tendrá que usar la función SI.                                                              |
| #REF!    | La fórmula se está refiriendo a una celda inválida.                                 | Asegurarse que no se haya movido, borrado o reemplazado la celda que está referenciada en la fórmula.                                                          |
| #N/A     | La fórmula no es capaz de encontrar el valor referenciado.                          | Asegurarse que todas las referencias se puedan evaluar correctamente. Este error es muy común en la función BUSCARV, BUSCARH.                                  |

## 2. 11 Herramientas de auditoría de fórmulas.

**Rastrear precedentes:** Identifica las celdas afectan al valor seleccionado.



**Rastrear dependientes:** Identifica las celdas que fueron afectadas por el valor seleccionado.

| Precio | Ingreso |
|--------|---------|
| 3      | 6       |

**Mostrar fórmulas:** Muestra las fórmulas como cadenas de texto.

| Código de área   |
|------------------|
| =IZQUIERDA(D2,3) |
| =IZQUIERDA(D3,3) |
| =IZQUIERDA(D4,3) |

**Evaluar fórmula:** Permite revisar cada cálculo de manera independiente, para identificar la causa del error entre sus componentes. También se puede usar para desglosar fórmulas desconocidas.

**Comprobación de error:** Escanea la hoja en busca de errores y proporciona un resumen con opciones para rastrear la fuente, ignorar el error o modificar tus opciones.

**Validación de datos:** Restringe los valores que un usuario puede ingresar en una celda determinada, basándose en

- tipo de número
- valor
- lista
- fecha y hora
- longitud del texto
- personalización

Microsoft Excel es una herramienta indispensable en el proceso de análisis de datos debido a su versatilidad, facilidad de uso, amplia gama de funciones y capacidad de integración con otras herramientas.

## 2.11 Estructuración de Datos: Importancia de Tener Datos Estandarizados en Excel

Tener datos estandarizados es fundamental para garantizar la precisión y eficacia en el análisis de información en Excel. La estandarización de datos facilita su captura, almacenamiento y presentación, lo que a su vez mejora la consistencia y claridad de las planillas de Excel. Este aspecto es especialmente crítico cuando se trabaja en equipos, donde la coherencia en los datos permite una colaboración más efectiva y reduce la probabilidad de errores.

## ***Captura de Datos***

Buenas Prácticas para la Captura de Datos:

- Definir Formatos Estándar: Asegúrate de que todos los datos se ingresen siguiendo formatos predefinidos (por ejemplo, fechas en formato "dd/mm/aaaa", números con dos decimales, etc.).
- Validación de Datos: Utiliza reglas de validación para restringir la entrada de datos incorrectos o fuera de rango. Esto incluye listas desplegables, validación de números, fechas y textos.
- Evitar Redundancias: Evita ingresar la misma información en múltiples lugares. Utiliza referencias a celdas o crea tablas de datos centralizadas para reducir redundancias.

## ***Almacenamiento de Datos***

Buenas Prácticas para el Almacenamiento de Datos:

- Estructura Clara y Organizada: Organiza los datos en tablas bien definidas con encabezados claros. Cada columna debe representar un tipo de dato específico.
- Uso de Tablas de Excel: Las tablas de Excel proporcionan ventajas como el autofiltro y la posibilidad de agregar fácilmente nuevas filas y columnas sin perder la coherencia de los datos.
- Nombres de Hojas y Archivos Descriptivos: Utiliza nombres de hojas y archivos que reflejen claramente el contenido y la fecha de actualización. Esto facilita la búsqueda y el mantenimiento de los archivos.
- Copias de Seguridad: Realiza copias de seguridad periódicas para evitar la pérdida de datos importantes. Almacena versiones anteriores del archivo para poder revertir cambios si es necesario.

## ***Presentación de Datos***

Buenas Prácticas para la Presentación de Datos:

- Formato Consistente: Asegúrate de que todos los datos sigan un formato consistente para facilitar la lectura y el análisis. Utiliza colores, fuentes y tamaños de manera uniforme.
- Uso de Gráficos y Tablas Dinámicas: Utiliza gráficos y tablas dinámicas para resumir y visualizar datos de manera clara y comprensible.
- Comentarios y Notas: Añade comentarios y notas explicativas donde sea necesario para proporcionar contexto adicional y aclarar cualquier duda que pueda surgir.

## ***Criterios para Mantener la Consistencia y Claridad***

- Documentación de Procedimientos: Mantén una documentación clara de los procedimientos y reglas de estandarización que todos los miembros del equipo deben seguir. Esto incluye formatos de fecha, numeración, nomenclatura de archivos, etc.
- Revisiones y Auditorías: Realiza revisiones periódicas y auditorías de las planillas para asegurar que se cumplan los estándares establecidos. Identifica y corrige cualquier inconsistencia encontrada.

- Capacitación Continua: Asegura que todos los miembros del equipo estén capacitados en las buenas prácticas de captura, almacenamiento y presentación de datos. Ofrece formación continua para mantener a todos actualizados con las mejores prácticas y nuevas herramientas.

### ***Beneficios de la Estandarización de Datos***

- Mejora la Precisión: Reduce los errores en la entrada y manipulación de datos, lo que mejora la precisión general del análisis.
- Facilita la Colaboración: Permite a los miembros del equipo trabajar juntos de manera más eficiente, ya que todos siguen las mismas reglas y formatos.
- Aumenta la Eficiencia: Ahorrar tiempo en la revisión y corrección de datos, lo que permite concentrarse más en el análisis y la toma de decisiones.
- Facilita el Análisis de Datos: Datos estandarizados son más fáciles de analizar y comparar, lo que facilita la obtención de insights valiosos.

### ***Normalización de Datos***

La normalización de bases de datos es un proceso fundamental en el diseño de bases de datos relacionales. Su objetivo es organizar los datos para minimizar la redundancia y mejorar la integridad de los datos, asegurando que la información esté estructurada de manera eficiente.

¿Por qué es importante la normalización?

Una base de datos no normalizada puede tener problemas como:

- Redundancia de datos: La misma información almacenada en múltiples lugares.
- Anomalías de actualización: Cambiar un dato en un solo lugar sin actualizarlo en otros puede generar inconsistencias.
- Anomalías de inserción y eliminación: Puede haber dificultades para agregar o eliminar datos sin afectar otros registros.

### **Buenas prácticas**

#### Evitar datos repetidos en una misma tabla

Si notas columnas con valores repetidos, considera dividir la tabla en dos para eliminar redundancias.

#### Documentar el diseño de la base de datos

Mantén un esquema bien documentado para que cualquier persona pueda entender la estructura de la base de datos.

#### Revisar y optimizar regularmente

Con el crecimiento de los datos, evalúa si es necesario realizar ajustes para mejorar la eficiencia.

## Ejemplo de Normalización: Datos de Ventas

| Pedido | Cliente     | Email          | Producto | Cantidad | Precio | Fecha Pedido |
|--------|-------------|----------------|----------|----------|--------|--------------|
| 101    | Ana Pérez   | ana@email.com  | Laptop   | 1        | 1200   | 24-02-15     |
| 102    | Juan García | juan@email.com | Mouse    | 2        | 250    | 24-02-16     |
| 103    | ANA PEREZ   | ana@email.com  | Teclado  | 1        | 500    | 24-02-16     |

*Tabla original (No Normalizada)*

### Pasos para Normalizar

1. Crear una tabla separada para los Clientes.

- Copiar y pegar las columnas Cliente y Email en una nueva tabla.
- Formatear los datos uniformemente y eliminar espacios de ser necesario.
- Quitar duplicados para eliminar registros repetidos.
- Agregar una columna ClientelD con un número único para cada cliente.

| ClientelD | Nombre      | Email          |
|-----------|-------------|----------------|
| 1         | Ana Pérez   | ana@email.com  |
| 2         | Juan García | juan@email.com |

*Tabla Clientes*




2. Modificar la tabla de Pedidos

- En la tabla original, reemplazamos Nombre y Email por ClientelD.

| Pedido | ClientelD | Producto | Cantidad | Precio | Fecha Pedido |
|--------|-----------|----------|----------|--------|--------------|
| 101    | 1         | Laptop   | 1        | 1200   | 24-02-15     |
| 102    | 2         | Mouse    | 2        | 250    | 24-02-16     |
| 103    | 1         | Teclado  | 1        | 500    | 24-02-16     |

*Tabla Pedidos (Normalizada)*

### **Beneficios de Normalizar**

-  Menos duplicados: Menos riesgo de errores al actualizar datos.
-  Más flexibilidad: Se pueden agregar más detalles en cada tabla sin afectar las demás.
-  Análisis más eficiente: Con tablas separadas, es más fácil hacer reportes y cálculos.

## **2.12 Enlaces para información adicional**

Funciones existentes en Excel por categoría:

<https://support.microsoft.com/es-es/office/funciones-de-excel-por-categor%C3%ADa-5f91f4e9-7b42-46d2-9bd1-63f26a86c0eb>

Tipos de datos en Excel:

<https://exceltotal.com/tipos-de-datos-en-excel-2013/>

Guía de atajos en Excel:

<https://exceltotal.com/atajos-de-teclado-en-excel/>