



Visualización Interactiva de la Información

Dra. Mariana-Edith Miranda-Varela

28-nov-2024

Agenda

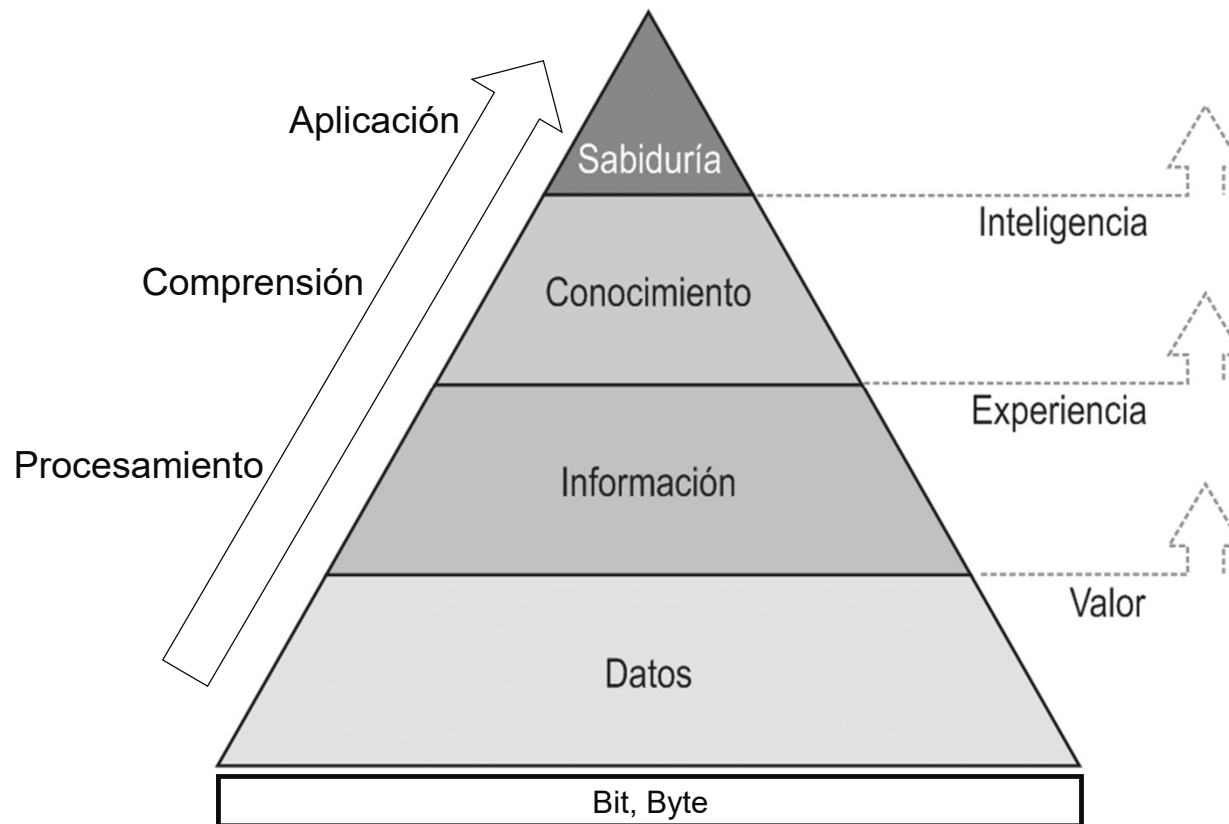
- Línea de tiempo
- Trabajar con datos
- Actividad



Línea de tiempo REVOLUCIÓN MEXICANA



Pirámide del conocimiento



FUENTE:
https://www.researchgate.net/publication/312201717_Knowledge_management_an_efficient_alternative_to_strengthening_micro_enterprises

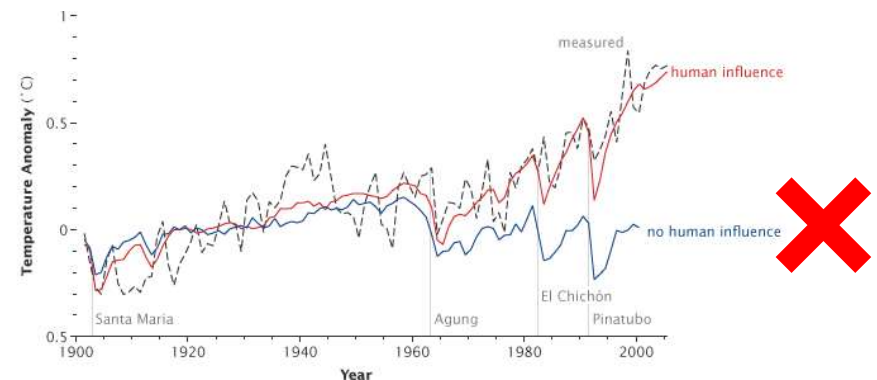
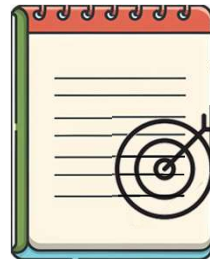
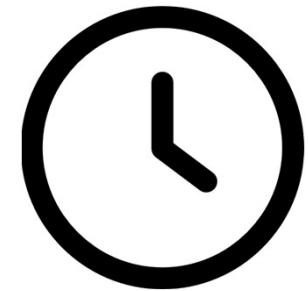


Trabajar con datos



Recolección de datos

- Correcto diseño
- Herramienta útil para:
 - Describir algo
 - Describir cómo están relacionadas las cosas
 - Explicar una relación
 - Influir sobre algo
- Identificar su propósito



Recolección de datos

Motivos erróneos

- Demostrar datos y números, no importan los resultados
- Se emplean mal los datos recolectados
- Preocupación por percepción de hallazgos
- Predisposición en “cómo” deben ser los resultados

Recolección de datos

Métodos complementarios

Método	Adecuado si
Medición directa	Los valores deben ser exactos
Grupo focal	<ul style="list-style-type: none">• No se sabe aún que preguntar• Si hay interés en dinámica de grupos o en procesos de decisión
Observación	Lo que se está midiendo es fácil y públicamente observable y se desea tomar notas del comportamiento de los participantes
Examen de documentos existentes	Los datos que nos interesan ya están registrados en algún sitio (recibos, facturas, logs de tráfico web, etc)
Diarios	Se necesita hacer un seguimiento de variables en el tiempo

Recolección de datos

Fuentes de datos externas

- Públicas
 - Contenidos abiertos
 - Análisis libre por terceros
 - Instituciones, gobiernos y organizaciones
- Privadas
 - Red en particular
 - Datos a la venta por terceros
 - Firma

Recolección de datos

Tipos de datos

Estructurados

ID	Name	Age	Degree
1	John	18	B.Sc.
2	David	31	Ph.D.
3	Robert	51	Ph.D.
4	Rick	26	M.Sc.
5	Michael	19	B.Sc.

Semiestructurados

```
<University>
  <Student ID="1">
    <Name>John</Name>
    <Age>18</Age>
    <Degree>B.Sc.</Degree>
  </Student>
  <Student ID="2">
    <Name>David</Name>
    <Age>31</Age>
    <Degree>Ph.D. </Degree>
  </Student>
  ....
</University>
```

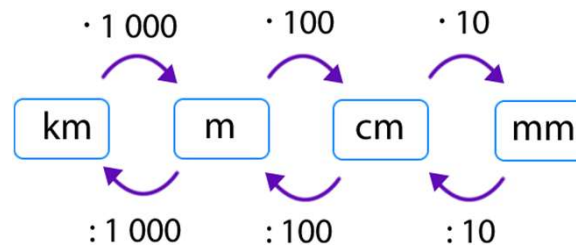
No estructurados

The university has 5600 students.
John's ID is number 1, he is 18 years old and already holds a B.Sc. degree.
David's ID is number 2, he is 31 years old and holds a Ph.D. degree. Robert's ID is number 3, he is 51 years old and also holds the same degree as David, a Ph.D. degree.

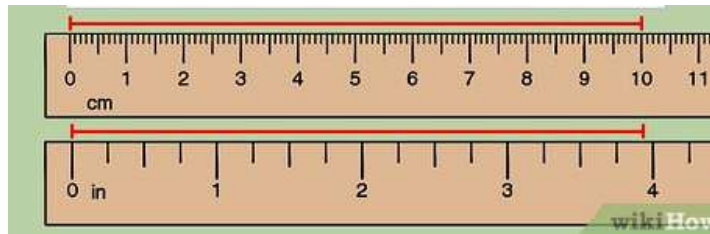
Preparación y limpieza de datos

Preparación de datos

- Separar campos que serán útiles
- Identificar datos incompletos y vacíos
- Uniformizar unidades



- Conversión de unidades de medida



Preparación y limpieza de datos

Limpieza

Nombre	Edad	Estatura	Correo electrónico
Juan Popoca Reyes	35	1.8	juan.p@gmail.com
Marcela López Ixtlamati	29	1.63	ixtlamati_mgmail.com
Natalia Espino Tellez	45	1.71	nat.tellez@gmail.com
Juan Popoca Reyes	35	1.8	juan.p@gmail.com
Estela Flores Campos	50	1.45	estela.fc@yahoo.com
Alfredo González Rojas	65	3.1	fredy_rojas@gmail.com
Pedro 1345	29	1.67	pedro_ac@gmail.com
Irlanda Tepox Luna	39	1.75	irlanda_xluna@gmail.com
Luz María Chairez Juárez	48	1.78	luh_cjz@gmail.com
Amlet Vázquez Lara	27	1.85	hvl@yahoo.com

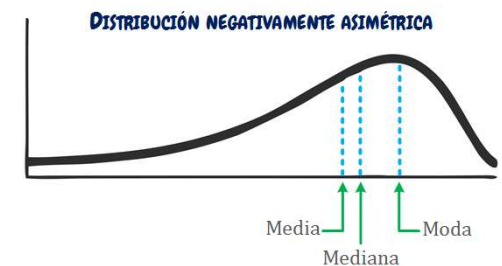
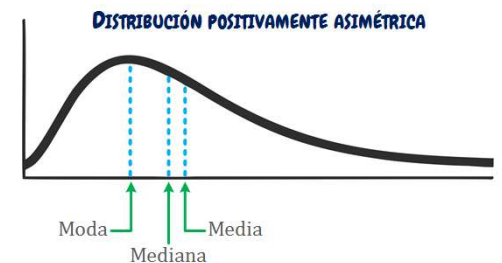
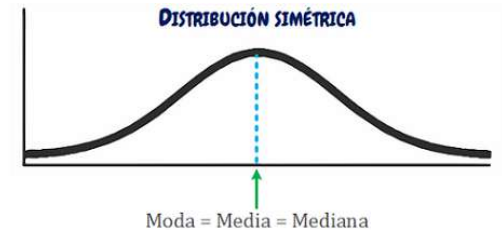
Preparación y limpieza de datos

Limpieza

Nombre	Edad	Estatura	Correo electrónico
Juan Popoca Reyes	35	1.8	juan.p@gmail.com
Marcela López Ixtlamati	29	1.63	ixtlamati_mgmail.com
Natalia Espino Tellez	45	1.71	nat.tellez@gmail.com
Juan Popoca Reyes	35	1.8	juan.p@gmail.com
Estela Flores Campos	50	1.45	estela.fc@yahoo.com
Alfredo González Rojas	65	3.1	fredy_rojas@gmail.com
Pedro 1345	29	1.67	pedro_ac@gmail.com
Irlanda Tepox Luna	39	1.75	irlanda_xluna@gmail.com
Luz María Chairez Juárez	48	1.78	luh_cjz@gmail.com
Amlet Vázquez Lara	27	1.85	hvl@yahoo.com

Transformación de datos

- Procedimiento de manipulación más común
- Revela características ocultas
- Transformar la distribución
 - Más fáciles de observar
 - Cumplir los requisitos de muchas pruebas estadísticas
 - Distribución normal (suposición)
 - Valores sesgados
 - Transformación logarítmica
 - Transformación mediante raíz cuadrada





Actividad



Dataset Titanic

- Analiza los datos del dataset Titanic e identifica sus principales características.
- ¿Qué gráficos se pueden realizar?
- Agrega tu gráfico en la liga que se compartirá durante la sesión (vigente únicamente durante la clase).

| Próxima sesión

Semana 4
27-nov-25

「 Muchas gracias. 」