



## Análisis de datos y elaboración de indicadores para la toma de decisiones

---

Profesora: Katherine Barquero Mejías



# Clase 1: Datos y sus principales herramientas de análisis

# Los datos

- ❖ En la práctica, es probable que alguna vez hayamos utilizado la palabra *datos*, pero realmente sabemos ¿qué son? o ¿cuáles tipos existen?

Según UM (2020), los datos se definen como: *“aquella información extraída de la realidad qque tiene que ser registrada en algún soporte físico o simbólico, que implica una elaboración conceptual y además que se pueda expresar en alguna forma de lenguaje”*

- ❖ Los datos están muy relacionados con la labor de los (as) investigadores: en muchos casos el dato hay que construirlo. Los datos tienen una estructura compuesta por tres elementos:

Una **unidad de análisis** que en una **variable** asume un **valor** determinado.

# Definiciones básicas

## ❖ Unidad de análisis:

Elementos menores que constituyen el universo de estudio de una investigación a partir de los cuales se estudia el comportamiento de variables. Se refiere a una unidad de la población de estudio.

Población de estudio	Unidad de análisis
Estudiantes costarricenses de 15 años que participaron en PISA 2018	Individuo: Estudiante costarricense que 15 años que participó en PISA 2018
Humedales en el territorio de Costa Rica durante el 2020	Superficies: Estudiante costarricense que 15 años que participó en PISA 2018
Exportaciones de Costa Rica registradas en 2020	Bien intangible: Tipo de exportación registrada en 2020
Conflictos socioambientales en Centroamérica entre el 2000 y 2020	Tema social: Conflicto socioambiental en Centroamérica entre 2000 y 2020

# Definiciones básicas

- ❖ **Variables:** Característica, Aspecto, magnitud de un elemento o unidad de análisis con la propiedad de asumir diferentes valores
- ❖ **Valores:** se refieren a las categorías que presenta la **variable** y adopta alguna **unidad de análisis** y que se puede expresar de forma cualitativa o cuantitativa.

Variables	Valores
Edad de un estudiante	Cuantitativo: 15, 20, 10
Tipo de humedal	Cualitativa: grande, mediano o pequeño
Cantidad de Exportaciones	Cuantitativa: 3000, 3010, 3020
Duración de conflictos socioambientales	Cualitativa: corta duración, mediana duración o larga duración

# Tablas de datos

- ❖ Las tablas de datos, usualmente conocidas como bases datos, son una representación de cualquier tipo de información en forma tabular, es decir, ordenada en filas y columnas (por eso se conocen también como matrices).
- ❖ En general, pueden ser de muchos tipos, pero en nuestro caso vamos a analizar tablas tradicionales para el análisis estadístico: aquellas que tienen información representada por **números o cadenas de caracteres (string o letras)**.
- ❖ Las **matrices de datos** resumen unidades de análisis a las que asignamos valores a variables específicas. La tabla 1 muestra el modelo típico de una matriz o tabla de datos, notar que en este ejemplo cada columna es una variable y cada fila es la unidad de análisis.

Tabla 1: Modelo de una tabla de datos

MODELO DE MATRIZ DE DATOS

UNIDAD DE ANÁLISIS	VARIABLES						
	V 1	V 2	V 3	V 4	V 5	V 6	V N
UA1	1	20	1	2	4	3	
UA2	1	28	1	1	5	1	
UA3	2	46	2	1	6	6	
UA4	1	34	3	4	7	1	
UA5	2	29	2	2	9	1	
UA6	2	19	1	3	4	3	
UA7	2	54	2	2	4	5	
UA8	1	67	4	5	2	4	
UA9	1	32	2	1	9	1	
UA10	2	23	3	4	5	6	
UA11	1	19	1	3	2	2	
UA12	2	18	1	3	5	3	
UA13	1	36	3	1	8	1	
UA14	2	63	4	1	3	5	
UAN							

# Tablas de datos

- ❖ En los casos de los cuestionarios o encuestas, como los archivos de datos que vamos a trabajar en el curso, usualmente la información de las bases de datos tienen códigos u opciones de respuestas indicadas en el cuestionario. Esto facilita el procesamiento de los datos, como veremos más adelante

Tabla 2: Modelo y referencia de la matriz de datos

## MODELO DE MATRIZ DE DATOS

UNIDAD DE ANÁLISIS	VARIABLES						
	V 1	V 2	V 3	V 4	V 5	V 6	VN
UA1	1	20	1	2	4	3	
UA2	1	28	1	1	5	1	
UA3	2	48	2	1	6	6	
UA4	1	34	3	4	7	1	
UA5	2	29	2	2	9	1	
UA6	2	19	1	3	4	3	
UA7	2	54	2	2	4	5	
UA8	1	67	4	5	2	4	
UA9	1	32	2	1	9	1	
UA10	2	23	3	4	5	6	
UA11	1	19	1	3	2	2	
UA12	2	18	1	3	5	3	
UA13	1	36	3	1	8	1	
UA14	2	63	4	1	3	5	
UAN							

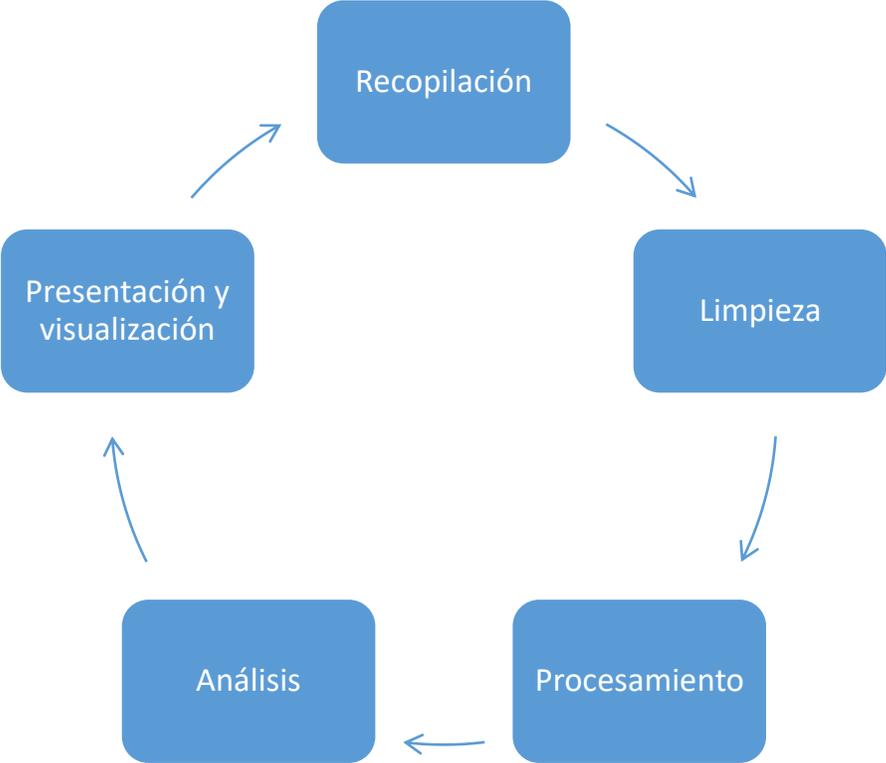
REFERENCIA DE LA MATRIZ DE DATOS	
UA= UNIDAD DE ANÁLISIS V= VARIABLES	
<b>V1</b> SEXO: 1. Masculino 2. Femenino	<b>V2</b> EDAD (Pregunta Abierta)
<b>V3</b> ESTADO CIVIL: 1. Soltero 2. Casado/Unido 3. Separado/Divorciado 4. Viudo 99. Ns./Nc.	<b>V4</b> ¿PODRÍA DECIRME QUE LUGAR OCUPA EN SU FAMILIA? 1. Cabeza de familia 2. Esposa/o cabeza de familia 3. Hijo/a 4. Hermano/a de 5. Padre/Madre 6. Otro especificar
<b>V5</b> NIVEL DE ESTUDIOS: 1. Sin Estudios 2. Primario Incompleto 3. Primario Completo 4. Secundario Incompleto 5. Secundario Completo 6. Terciario Incompleto 7. Universitario Incompleto 8. Terciario Completo 9. Universitario Completo o + 99. No sabe/No contesta	<b>V6</b> SITUACIÓN OCUPACIONAL: 1. Trabaja 2. Desocupado 3. Estudia 4. Jubilado o Pensionado 5. Rentista 6. Ama de Casa

# Ciclo de análisis de datos

1 Partimos de un **problema de investigación** para el cual buscamos evidencia de soluciones

2 Recurrimos al **análisis de datos**

3 Brindamos **evidencia** y tomamos decisiones informadas



# Fuentes de información

## Registros administrativos

Información recopilada de datos y registros de las operaciones de una organización

Pueden ser producidos a lo interno de la organización o provenir de fuentes externas

## Cuestionarios o Encuestas

Se emplean para recolectar información de personas, grupos o temas específicos de los cuales no se tiene información

Usualmente se trabaja con muestras porque no se dispone de toda la información del universo de interés

## Información en la web

Se refiere a información disponible en la web

Pueden venir de diferentes fuentes: instituciones académicas, organismos internacionales, portales estadísticos

# Tipos de análisis de datos



# Análisis descriptivo

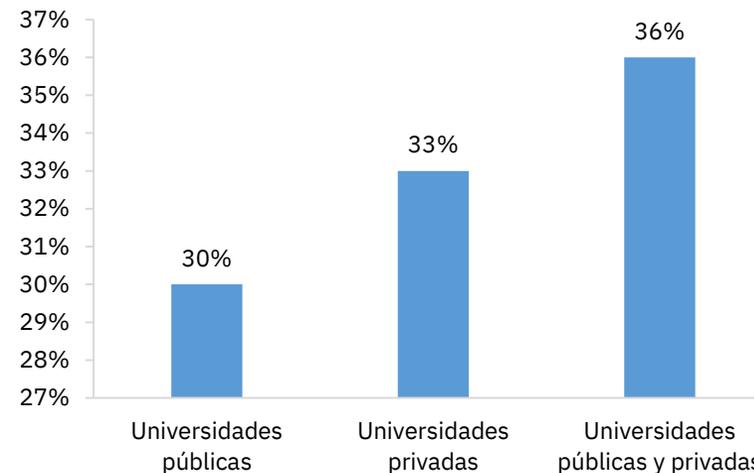
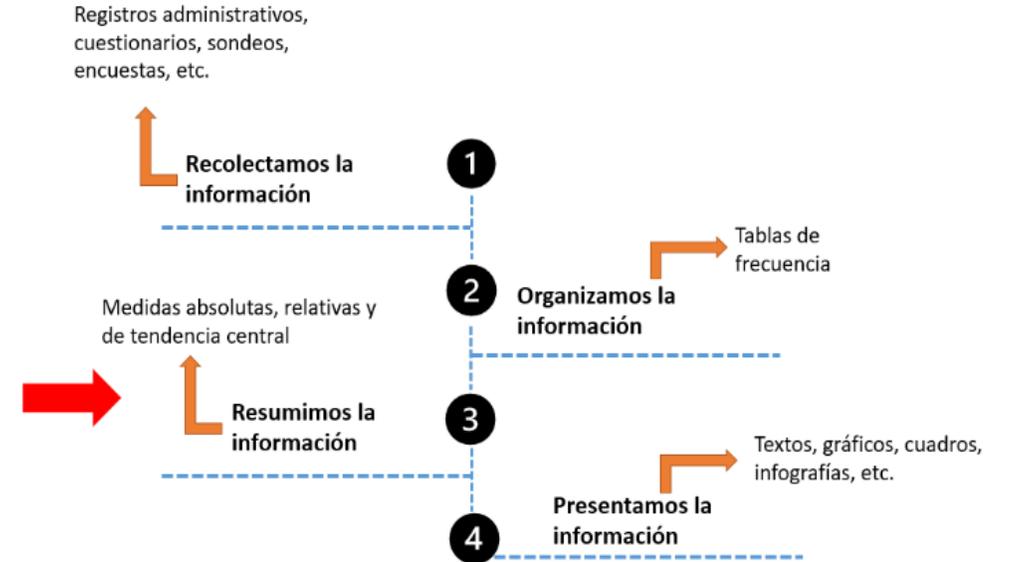
- ❖ Objetivo: resumir y describir los datos mediante estadísticas descriptivas (tablas de frecuencias, promedios, desviaciones estándar, gráficos)

## Ejemplo

- Se realizó una encuesta a 258 docentes
- Se indagó sobre su formación universitaria
- Se resume la información con una tabla de frecuencias

Distribución de docentes según el tipo de formación universitaria

Categoría	Frecuencia	Porcentaje
Universidades públicas	78	30.23
Universidades privadas	86	33.33
Universidades públicas y privadas	94	36.43



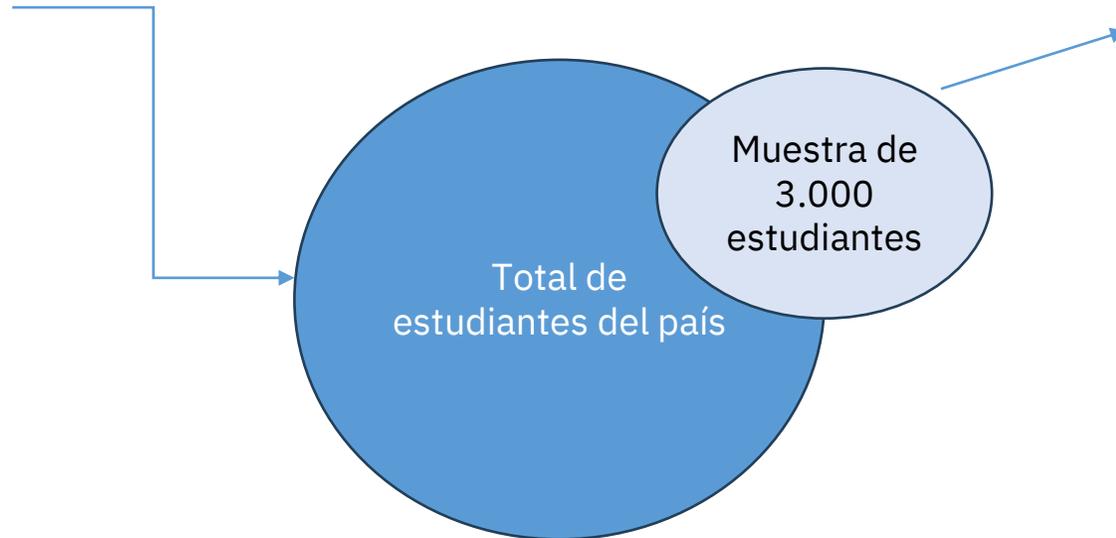


# Análisis inferencial

- ❖ Objetivo: realizar inferencias y generalizaciones a partir de una muestra de datos
- ❖ Técnicas propias del análisis inferencial (muestreo y análisis multivariado)
- ❖ Emplea el análisis descriptivo para presentar y resumir la información

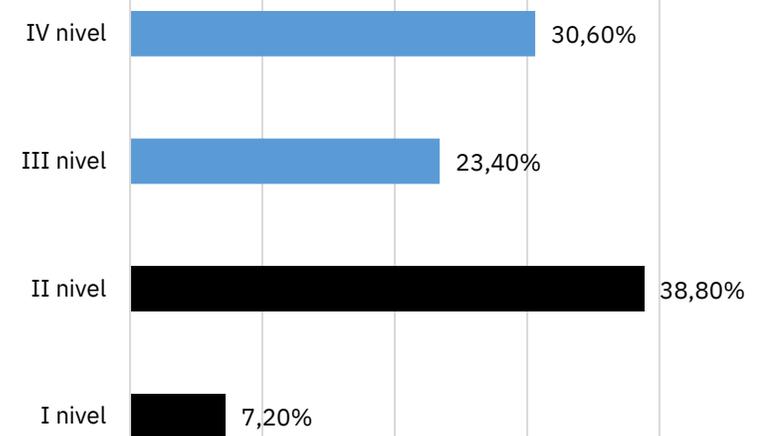
## Objetivo:

Conocer las habilidades en lectura que alcanzan los estudiantes al terminar la escuela



## Conclusión:

Distribución de estudiantes según el nivel de desempeño alcanzado en lectura, 2019



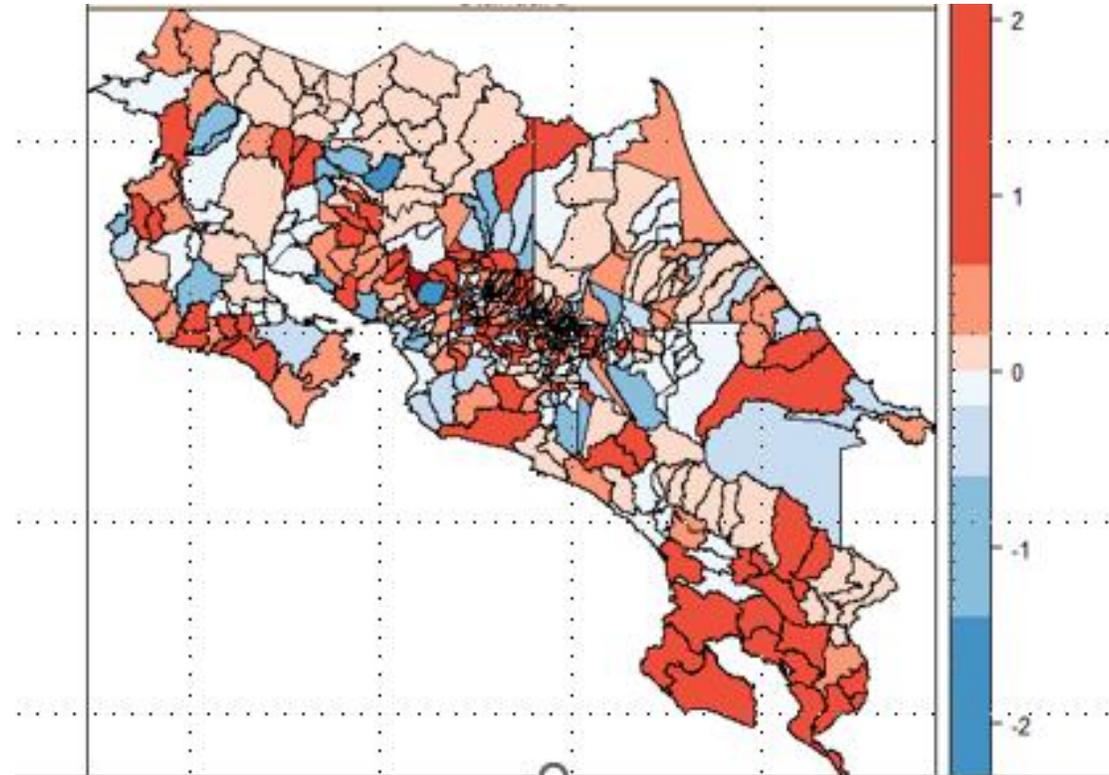
Fuente: Unesco, 2019

# Análisis espacial

- Objetivo: Comprender patrones geográficos y relaciones entre datos en un contexto geográfico
- Técnicas de Sistemas de información Geográfica (SIG)

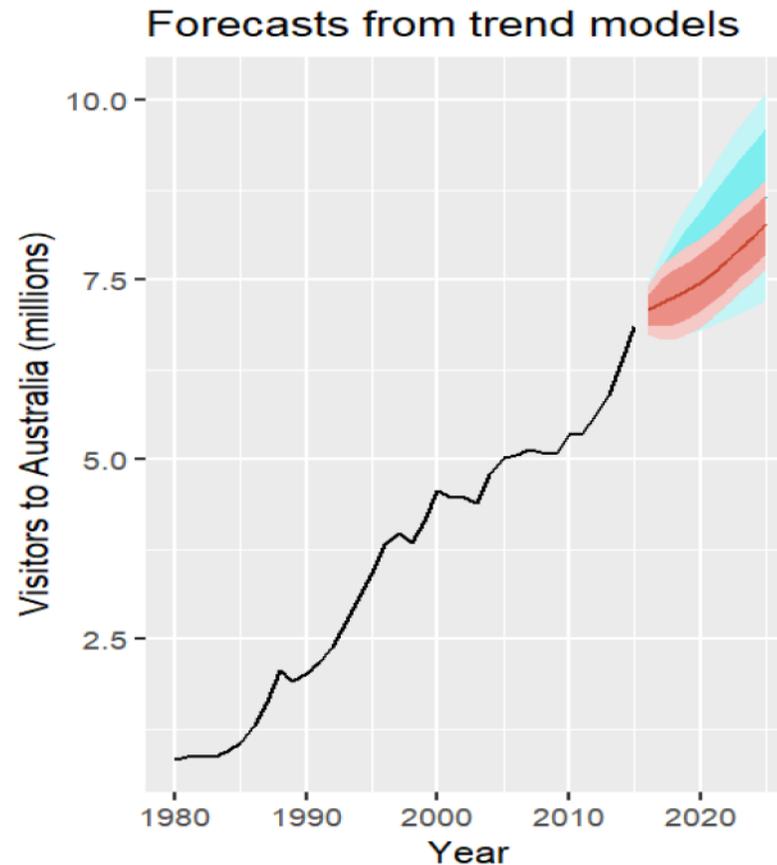
## Pregunta de investigación

¿Existen patrones espaciales en la disminución de la matrícula de las escuelas públicas de Costa Rica?



# Análisis de series de tiempo

- Objetivo: analizar el comportamiento de variables cuyos datos se registran en el tiempo. Se busca proyectar el comportamiento futuro
- Estadística multivariada, técnicas de minería de datos y machine learning



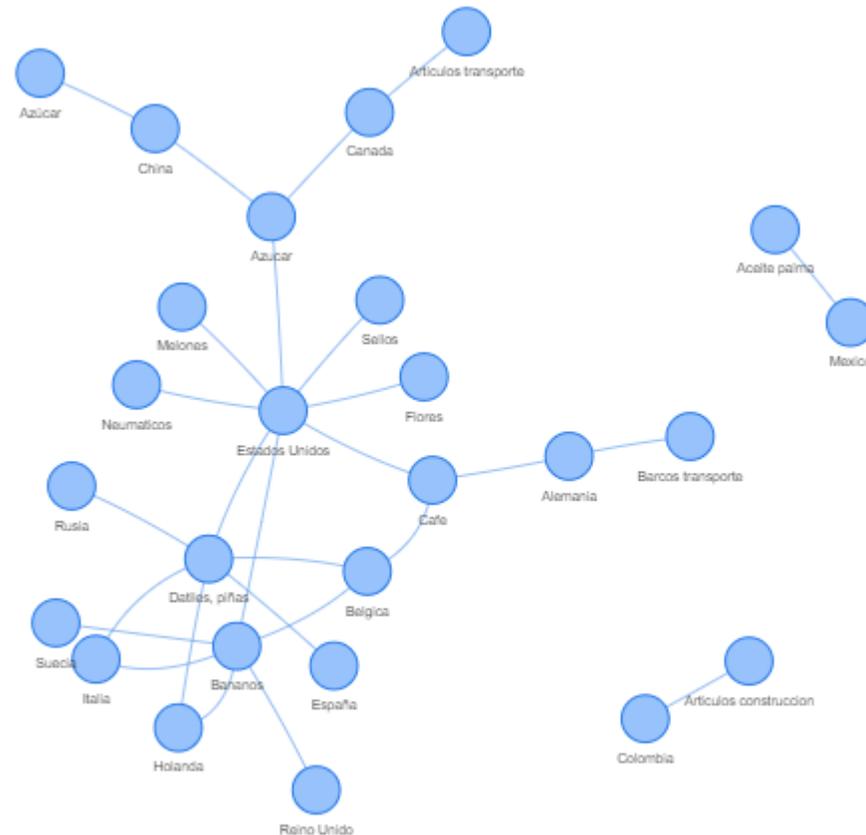


# Análisis de redes

- Identificar relaciones entre entidades, redes de colaboración, transporte, etc.
- Estadística multivariada

**Objetivo: Identificar los principales socios comerciales con respecto al peso de los principales productos de exportación.**

Redes entre productos de exportación y socios comerciales en Costa Rica. 2018



Muchas gracias

