

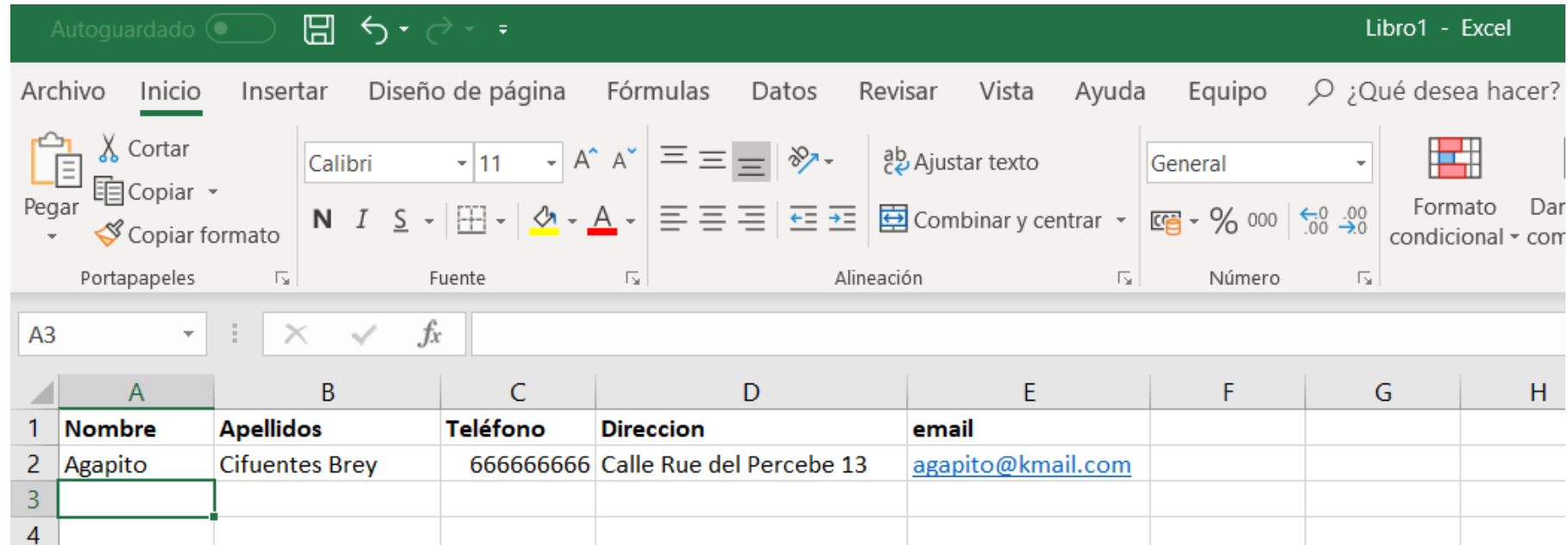
Introducción a las Bases de Datos

Máster en Data Science

M1967 - Modelos de Datos y Sistemas de Información 2019-2020

¿Qué es una base de datos?

- Podemos definir una base de datos como un conjunto de información relacionada que se almacena de forma agrupada para su posterior gestión, consulta o análisis.
 - ¿Es una hoja Excel un ejemplo de base de datos?

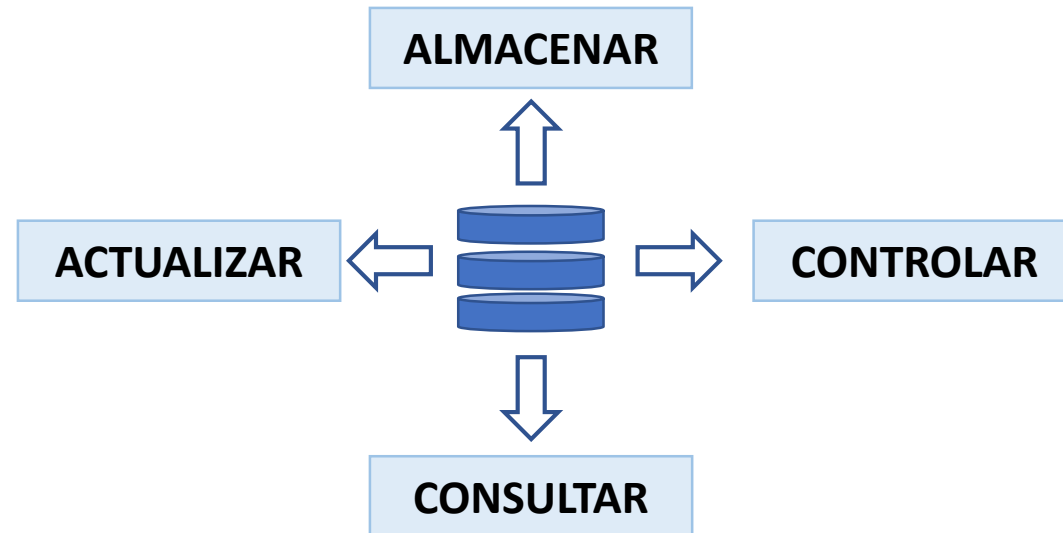


The screenshot shows the Microsoft Excel interface with a spreadsheet titled 'Libro1 - Excel'. The ribbon is set to 'Inicio' (Home). The spreadsheet contains a table with the following data:

	A	B	C	D	E	F	G	H
1	Nombre	Apellidos	Teléfono	Dirección	email			
2	Agapito	Cifuentes Brey	666666666	Calle Rue del Percebe 13	agapito@kmail.com			
3								
4								

¿Qué es una base de datos?

- Las bases de datos deben de permitir:
 - Almacenar los datos.
 - Controlar los datos mediante definición de restricciones e implementación de seguridad.
 - Consultar los datos.
 - Actualizar los datos.



¿Qué es una base de datos?

- Una base de datos refleja aspectos de la realidad, esto es, información que pueda ser procesada sobre un contexto.
 - Una base de datos de una Universidad reflejará información sobre:
 - Sus estudiantes (nombre, edad, dirección, contacto...)
 - Sus planes de estudios (nombre, grados, másters...)
 - Las asignaturas dentro de los planes (nombre, créditos, horarios...)
 - Las matrículas de los estudiantes en planes de estudios (fechas, asignaturas...)
 - Etc.
 - Una base de datos de un banco contendrá, entre otros, los siguientes datos:
 - Sus clientes (nombre, edad, dirección, contacto...)
 - Sus productos (cuentas corrientes, planes de pensiones...)
 - Los contratos de sus clientes (que productos contratan, fechas...)
 - Etc.

¿Qué es un Gestor de Bases de Datos?

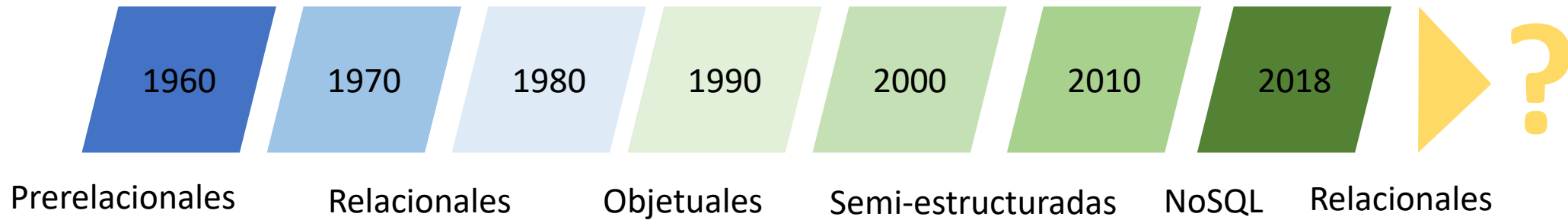
- Un Sistema Gestor de Bases de Datos (SGDB) es un programa (o un conjunto de programas) que permite almacenar, mantener, actualizar y acceder a los datos, así como definir restricciones de integridad o seguridad, entre otro.
 - En definitiva, permite gestionar los datos y garantizar su persistencia.
 - Hay diferentes tipos de SGDBs:
 - Jerárquicos
 - En Red
 - Relacionales
 - Semiestructurados
 - Geográficos
 - Temporales
 - NoSQL
 - NewSQL
 - Etc.

SGDB: Un poco de historia

- Previamente al desarrollo de los SGDB, los datos se almacenaban directamente en Sistemas de Ficheros.

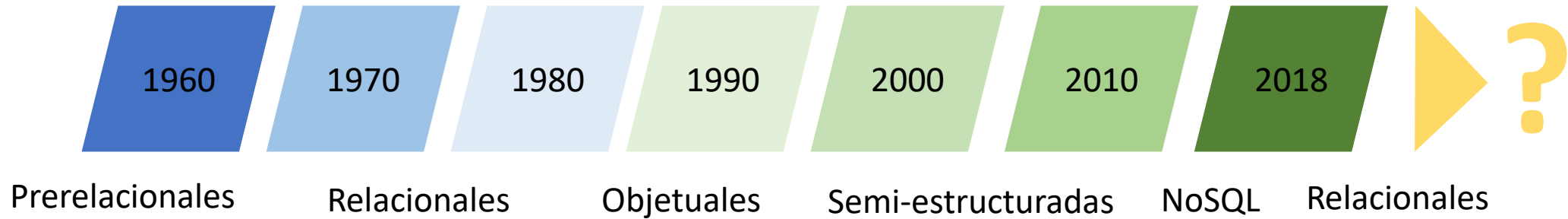
Desventajas	Motivo
Redundancia e inconsistencia de datos	La falta de control provoca que se pueda duplicar la información
Problemas de integridad	No se puede indicar que un campo como, por ejemplo, la edad, ha de ser numérico
Imposibilidad de acceso concurrente	Varios usuarios no pueden acceder a un fichero de forma simultánea
Sin control de seguridad	Difícil implementación de una seguridad a diferentes niveles para diferentes usuarios
Diferentes formatos de datos	En cada fichero los datos pueden estar en un formato diferente
...	...

SGDB: Un poco de historia

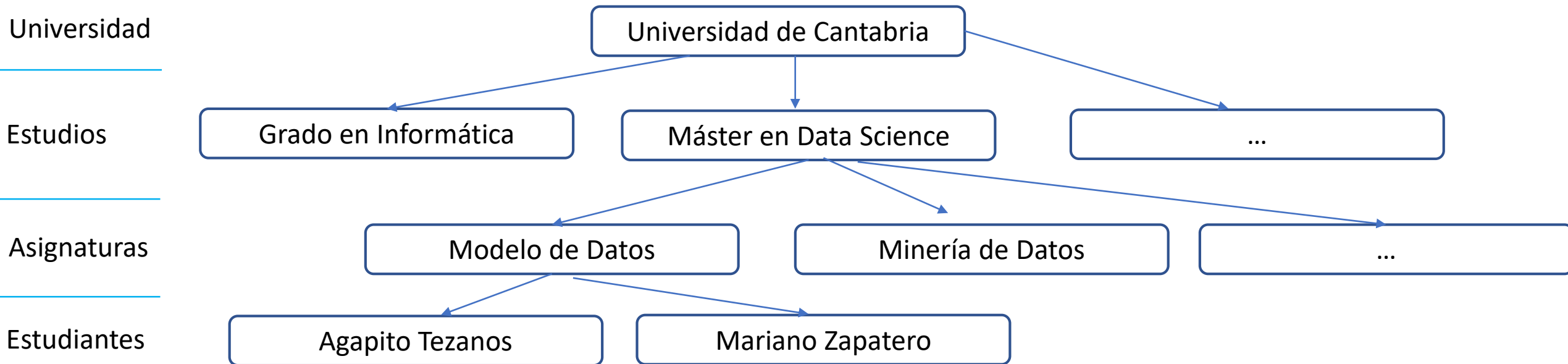


- Hasta 1960 (aprox.) se usaban cintas magnéticas y el lenguaje COBOL, que fue el primero en describir los datos y los ficheros con independencia de las acciones.
- Entre 1960 a 1970 se comenzaron a utilizar bases de datos jerárquicas y en red, que requería el uso de punteros para el acceso a los datos.

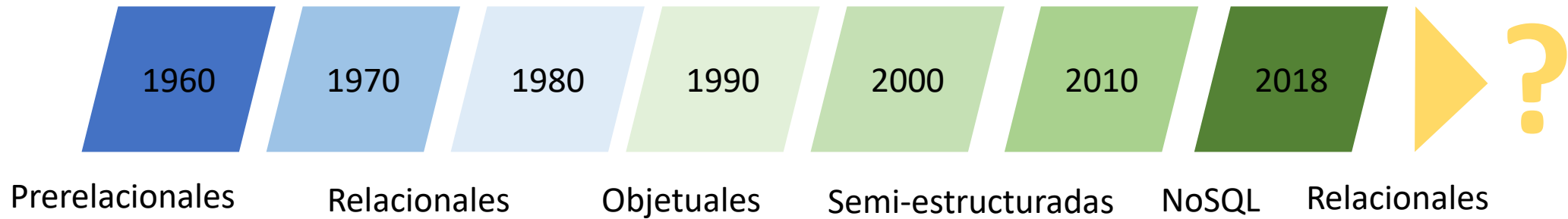
SGDB: Un poco de historia



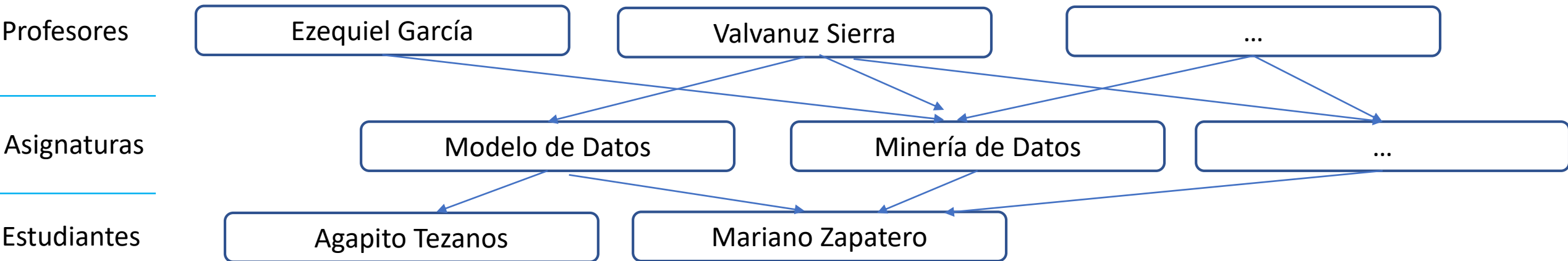
- Ejemplo de base de datos jerárquica:



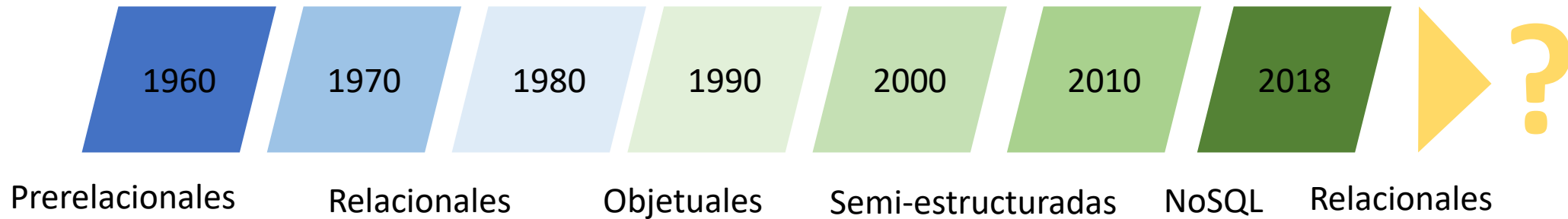
SGDB: Un poco de historia



- Ejemplo de base de datos en red:

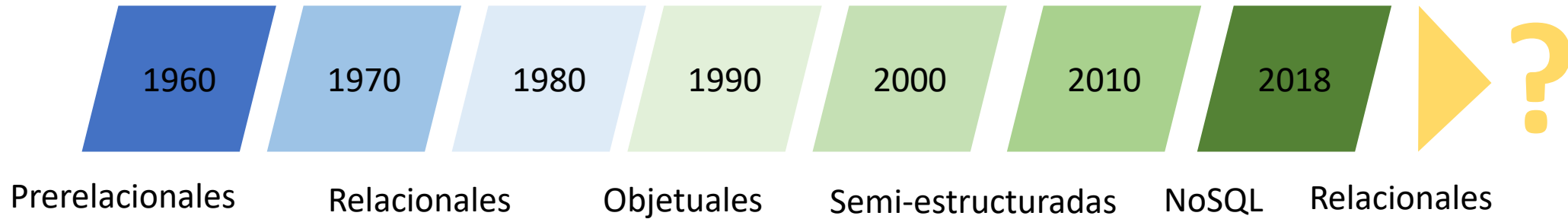


SGDB: Un poco de historia



- En la década de los 70, Codd definió el modelo relacional.
 - Modelo teórico bien fundamentado.
 - Garantiza la independencia lógica y física de los datos.
 - Provee de un lenguaje declarativo “sencillo”, llamado SQL (Structured Query Language).
 - Se basa en tablas y sus relaciones.

SGDB: Un poco de historia



- Ejemplo de base de datos relacional

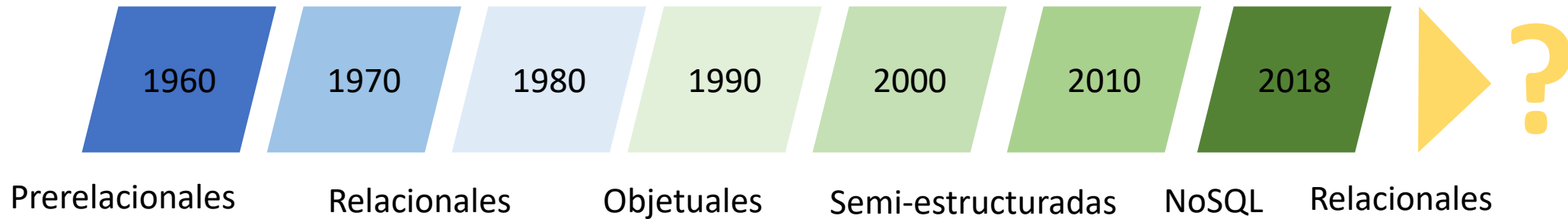
Identificador	Nombre	Apellidos	...
Uc0001	Diego	García Saiz	...
Uc0002	Valvanuz	Sierra	...
...

Tabla Profesor

Identificador	Nombre	Profesor Responsable
m301	Modelado de Datos	Uc0001
m389	Minería de datos	Uc002
...

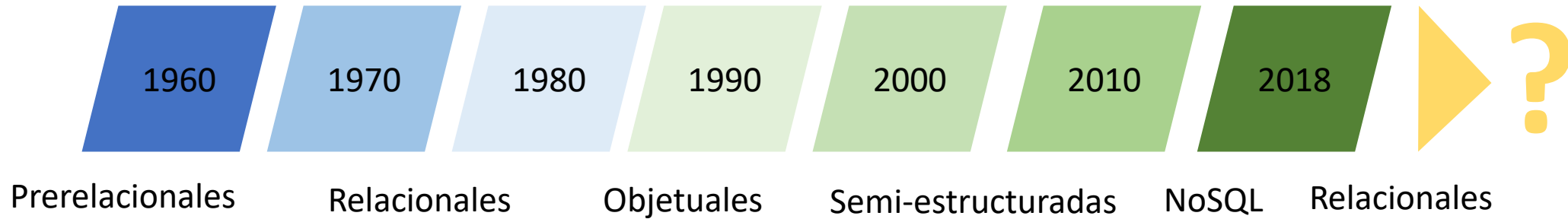
Tabla Asignatura

SGDB: Un poco de historia



- En la década de los 70, Codd definió el modelo relacional.
 - Modelo teórico bien fundamentado.
 - Garantiza la independencia lógica y física de los datos.
 - Provee de un lenguaje declarativo “sencillo”, llamado SQL (Structured Query Language).
 - Se basa en tablas y sus relaciones.
 - Se “exige” un control de la integridad y consistencia de los datos y su normalización:
 - No admite redundancia (el dato de profesor responsable de una asignatura sólo aparece en una tabla).
 - No hay mecanismos de tipo colección (no se usan listas, arrays...)
 - Hay que respetar el tipo de datos de cada campo (no se pueden insertar letras en campos numéricos como la edad)

SGDB: Un poco de historia



- A partir de la aparición de los modelos relacionales, surgen otros paradigmas:
 - Semi-estructuradas: XML, JSON...

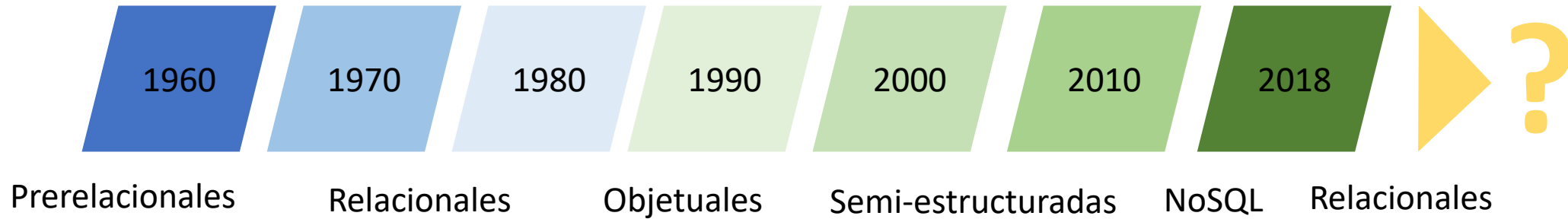
```
<profesor id = 1>  
  <nombre> Diego </nombre>  
  <apellidos> García Saiz </apellidos>  
  ...  
</profesor>
```

Ejemplo XML

```
profesor {  
  id: 1,  
  nombre: "Diego"  
  ...  
}
```

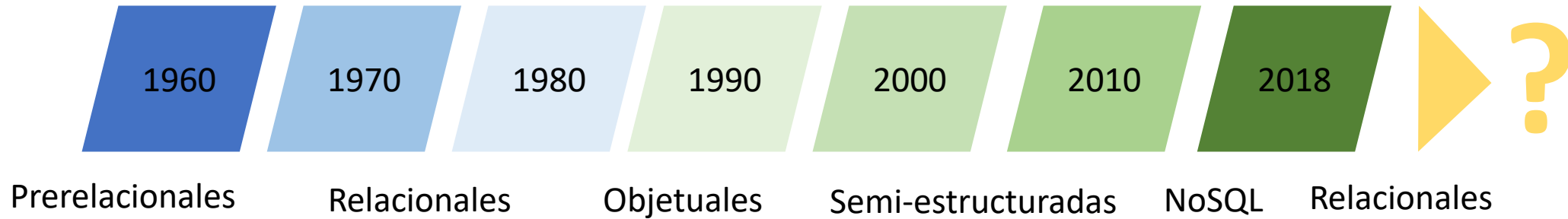
Ejemplo JSON

SGDB: Un poco de historia



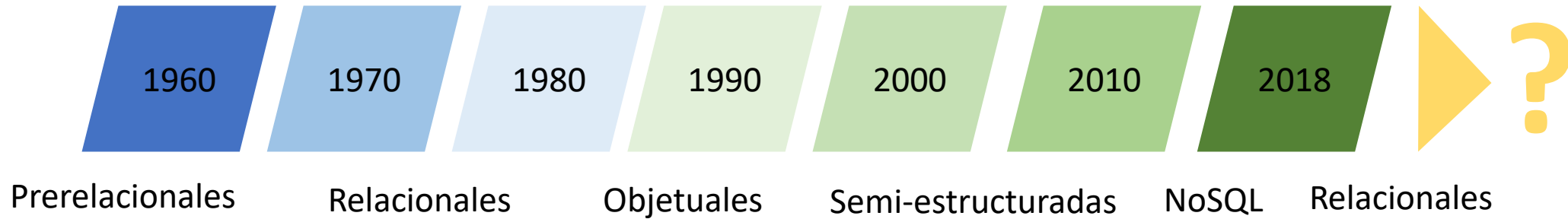
- A partir de la aparición de los modelos relacionales, surgen otros paradigmas:
 - Semi-estructuradas: XML, JSON...
 - Objetuales y la extensión objeto-relacional.

SGDB: Un poco de historia



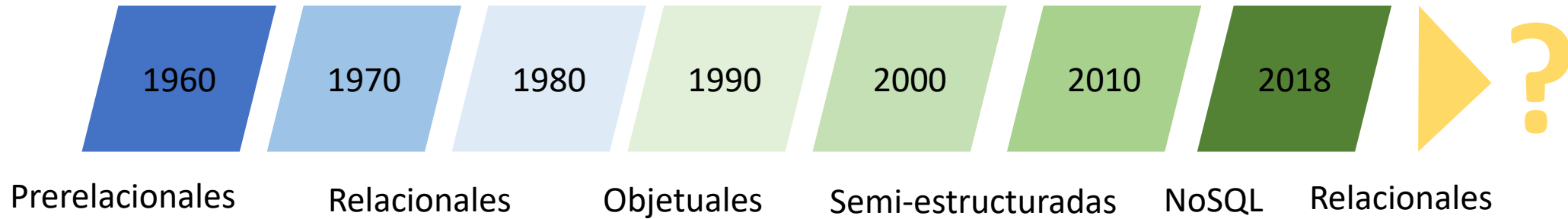
- A partir de la aparición de los modelos relacionales, surgen otros paradigmas:
 - Semi-estructuradas: XML, JSON...
 - Objetuales y su extensión objeto-relacional.
 - Distribuidas (pueden ser relacionales)

SGDB: Un poco de historia



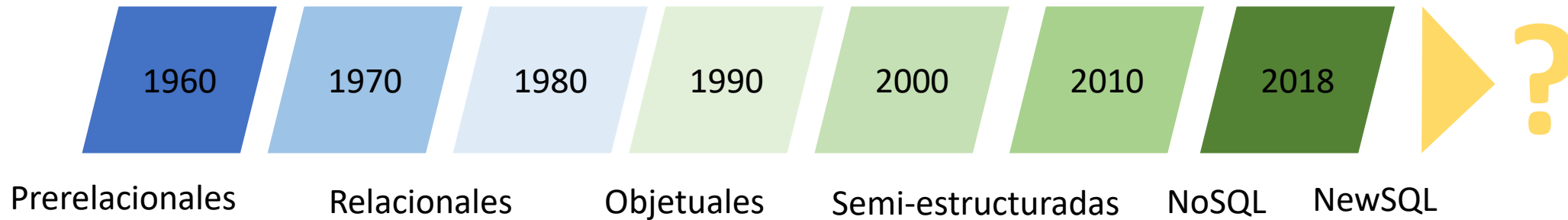
- A partir de la aparición de los modelos relacionales, surgen otros paradigmas:
 - Semi-estructuradas: XML, JSON...
 - Objetuales y su extensión objeto-relacional.
 - Distribuidas (pueden ser relacionales)
 - Geográficas

SGDB: Un poco de historia



- A partir de la aparición de los modelos relacionales, surgen otros paradigmas:
 - Semi-estructuradas: XML, JSON...
 - Objetuales y su extensión objeto-relacional.
 - Distribuidas (pueden ser relacionales)
 - Geográficas
 - Multimedia
 - ...

SGDB: Un poco de historia



- Uno de los paradigmas con más empuje actualmente son las bases de datos NoSQL.
 - Surgen, en contraposición a las bases de datos relacionales, para ofrecer un mayor rendimiento a costa de sacrificar la consistencia de los datos.
 - Muchas admiten redundancia.
 - En ocasiones el control de la integridad de los datos depende de quien los introduce (el cliente).
 - No es una tecnología concreta, sino una “filosofía”. Diferentes SGDB no relacionales tienen su propio modelo de datos.
 - Cada tipo de SGDB NoSQL responde a problemas diferentes.
 - Otro de los paradigmas que están surgiendo con fuerza son las bases de datos NewSQL, que tratan de obtener el rendimiento de las NoSQL sin dejar de lado las ventajas de los SGDB relacionales.