

Bases de Datos

Clase 1: Motivación

Información General

- Profesores:
 - Tamara Cucumides
 - Matías Toro
 - Domagoj Vrgoč
- Clases: W - 5, Talleres: W - 6
- Ayudantías: V - 5 (Se avisan con una semana de anticipación)
- Web: Syllabus en GitHub

Profesores



Tamara Cucumides

Profesor Adjunto IMC, PUC /

Magíster en Ciencia de la computación, PUC

- Tesis sobre cotas y algoritmos para evaluación de consultas de caminos en bases de datos de grafos.
- Data scientist 2 años de experiencia en la industria



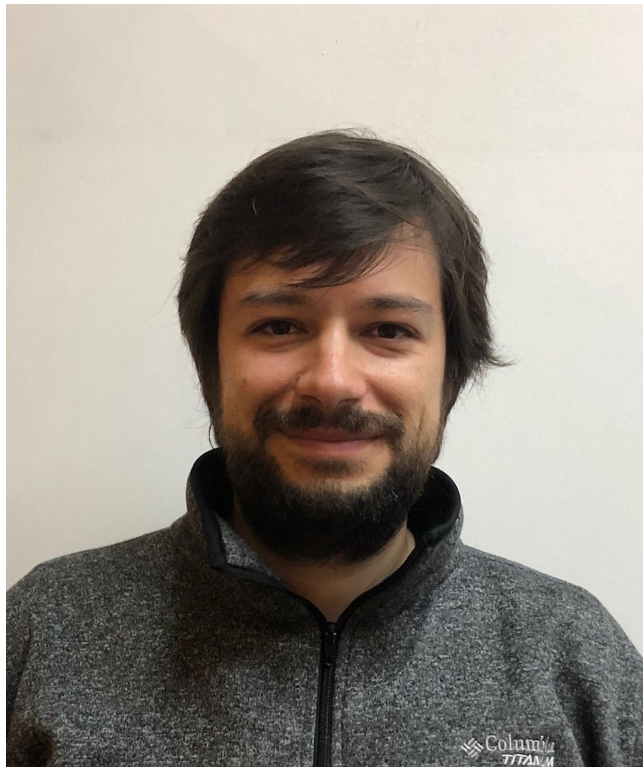
Domagoj Vrgoč

Profesor IMC, PUC

Investigador Instituto Milenio Fundamento de los Datos

- Experto en Bases de Datos de Grafos
- Implementación de sistemas de Bases de Datos
- Streaming y enumeración

Profesores



Matías Toro

Postdoc en el DCC, Uchile / Doctor en Ciencias de la Computación, Uchile

- Áreas de investigación:
- Lenguajes de programación
- Sistemas de tipos
- Privacidad diferencial
- Experiencia 13 años en la industria diseñando y desarrollando software aeronáutico.

Ayudantes jefes



Jerry Mendoza

Jefa de cátedra:

- Organizar ayudantías
- Evaluaciones y examen (incluye correcciones)
- Dudas de materia



Andrés Pincheira

Jefe de proyecto:

- Servidores
- Pautas y enunciados
- Correcciones y correcciones
- Dudas de proyecto

Evaluación

La evaluación se realizará en base a:

- 3 Controles evaluados y un control bonus, a realizarse principalmente en clases con plazo hasta el viernes de la misma semana.
- Examen, con el objetivo de evaluar conocimientos teóricos.
- Un proyecto semestral con 3 evaluaciones parciales. El objetivo del proyecto es que el alumno desarrolle una base de datos para una aplicación basada en un caso real. Las etapas avanzadas del proyecto introducirán al alumno otras problemáticas reales de las bases de datos.

La **nota de controles y exámenes** (NCE) corresponde al promedio de los controles y el examen. En otras palabras, $NCE = (C1 + C2 + C3 + Ex) / 4$. El control bonus puede reemplazar su peor control, pero no el examen.

Se podrán eximir del examen los alumnos que tengan un promedio entre los 3 controles $> 5,5$.

La **nota del proyecto** (NP) corresponde al promedio ponderado de todas las entregas del proyecto. La ponderación es: Entrega 1: 20 %, Entrega 2: 40 %, Entrega 3: 40 %.

Importante: La nota de las entregas del proyecto estarán sujetas a una evaluación de pares que puede alterar la nota del alumno. Si el profesor lo considera necesario, citará a los alumnos cuyas evaluaciones de pares indican un comportamiento irregular en las entregas.

Para aprobar el ramo, el alumno debe cumplir que NCE y NP sean $\geq 3,95$. En ese caso, la **nota final** se calcula como $NF = (0,5NCE + 0,5NP)$. En caso contrario, $NF = \min\{ NCE , NP \}$.

La fecha y forma de entrega de las entregas del proyecto será informada a medida que avanza el curso. El plazo final siempre es a las 20:00 del día de entrega.

¿Por qué este curso?

Un día cualquiera (9 de Marzo)

8:30 am	Despierto, reviso Telegram
---------	----------------------------

11:30 am	Veo los casos nuevos
----------	----------------------

12:00 pm	Se me acabo el pan, me pido un Cornershop
----------	---

12:30 pm	Reviso latercera.com
----------	----------------------

13:00 pm	Saco mi tercer 10% y lo deposito en mi banco
----------	--

13:30 pm	Actualizo unos documentos en Google Drive
----------	---

14:15 pm	Aburrido, reviso Twitter
----------	--------------------------

15:00 pm	Reviso el precio de mis Bitcoins
----------	----------------------------------

...	...
-----	-----

Un día cualquiera (9 de Marzo)

8:30 am Despierto, reviso **Telegram**

11:30 am Veo los **casos nuevos**

12:00 pm Se me acabo el pan, **me pido un Cornershop**

12:30 pm Reviso **latercera.com**

13:00 pm Saco mi tercer 10% y lo **deposito** en mi banco

13:30 pm Actualizo unos documentos en **Google Drive**

14:15 pm Aburrido, reviso **Twitter**

15:00 pm Reviso el precio de mis **Bitcoins**

...

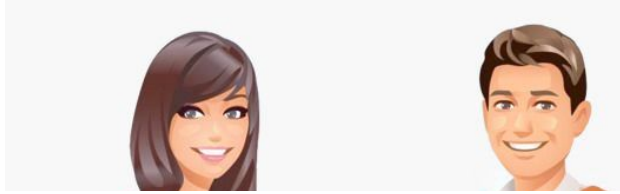
...

Un día cualquiera (9 de Marzo)

Todas estas actividades involucran alguna base de datos:

- Búsquedas en la web
- Datos públicos
- Redes sociales
- Métodos de pago
- Criptomonedas

Donde sea que trabajen,
tendrán que interactuar con
bases de datos



Oiga pero eso es tan noventero... Hoy día las bases de datos ya no se usan, lo que la lleva es

Big Data

Cloud Computing

Blockchain

quantum computing

Microservices

Buzzword

Deep Learning

Internet of Things

Data Science



Donde sea que trabajen,
tendrán que interactuar con
bases de datos

La competencia más buscada en trabajos de ciencia de datos?

¡Saber bases de datos!
modelo relacional, SQL, noSQL

Outline

- Qué son las bases de datos
- Por qué usar sistemas de bases de datos
- Roadmap del curso

Bases de Datos

Necesitamos una forma de almacenar los datos



- Para poder procesarlos de forma eficiente
- Que no haya que programar desde 0
- Que sea portable y estándar

Sistemas de Bases de Datos

Sistema de gestión de bases de datos (Database Management System - **DBMS**)

- Programa que facilite la lectura y almacenamiento de grandes volúmenes de datos.

Sistemas de Bases de Datos



- Datos se almacenan en disco
- Pero los usuarios interactúan con una **capa lógica** (ej. tablas), fácil de acceder

ID Actor	Nombre Actor
1	Leonardo DiCaprio
2	Matthew McConaughey
3	Daniel Radcliffe
4	Jessica Chastain
...	...

ID Película	Nombre Película
1	Interstellar
2	The Revenant
3	Harry Potter
4	The Wolf of Wall Street
...	...

ID Actor	ID Película
1	2
1	4
2	1
3	3
...	...

Cómo funciona un DBMS



¿Cuál es la mejor película de Christopher Nolan?



Cómo funciona un DBMS



¿Cuál es la mejor película de Christopher Nolan?



```
SELECT movies,  
rating FROM Movies
```



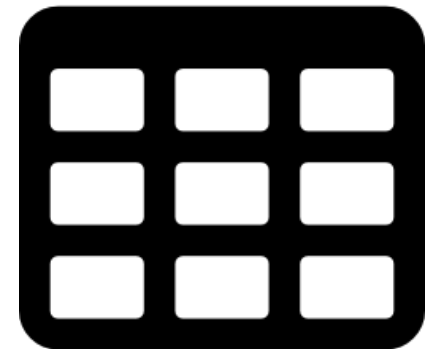
Cómo funciona un DBMS



¿Cuál es la mejor película de Christopher Nolan?



```
SELECT movies,  
rating FROM Movies
```



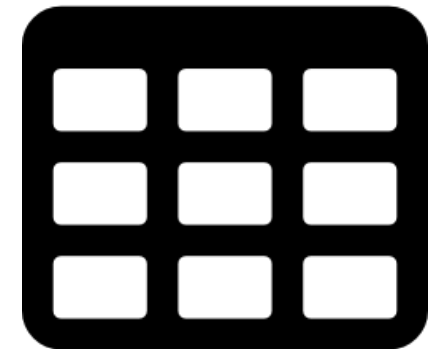
Cómo funciona un DBMS



¿Cuál es la mejor película de Christopher Nolan?



```
SELECT movies,  
rating FROM Movies
```



Interstellar

Por qué usar DBMS

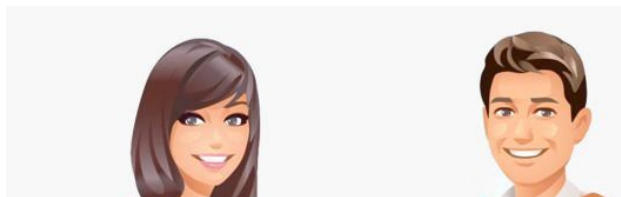
- Almacenar datos (insertar)
- Encontrar datos (búsquedas y consultas)
- Modificar datos (update)
- Asegurar la consistencia de los datos
- Seguridad y privacidad de los datos

En este curso usaremos:

- PostgreSQL

Por qué usar DBMS

- Almacenar datos (insertar)
- Encontrar datos (búsquedas y consultas)
- Modificar datos (update)
- Asegurar la consistencia de los datos
- Seguridad y privacidad de los datos



Pero podríamos programar todo esto en python y nos ahorramos aprender?

Ah claro, pero tendríamos que trabajar mucho resolviendo cosas que ya han sido resuelto por otros más talentosos que nosotros.

En este curso trabajaremos
principalmente con bases de
datos relacionales

Lo primero, es aprender a
cómo funciona el modelo
relacional

Modelo Relacional

El modelo de las bases de datos relacionales se basa en:

- Tablas (relaciones)
- Columnas de las tablas (atributos con sus tipos)
- Filas de las tablas (tuplas) que contienen los datos

Lenguajes de Consultas

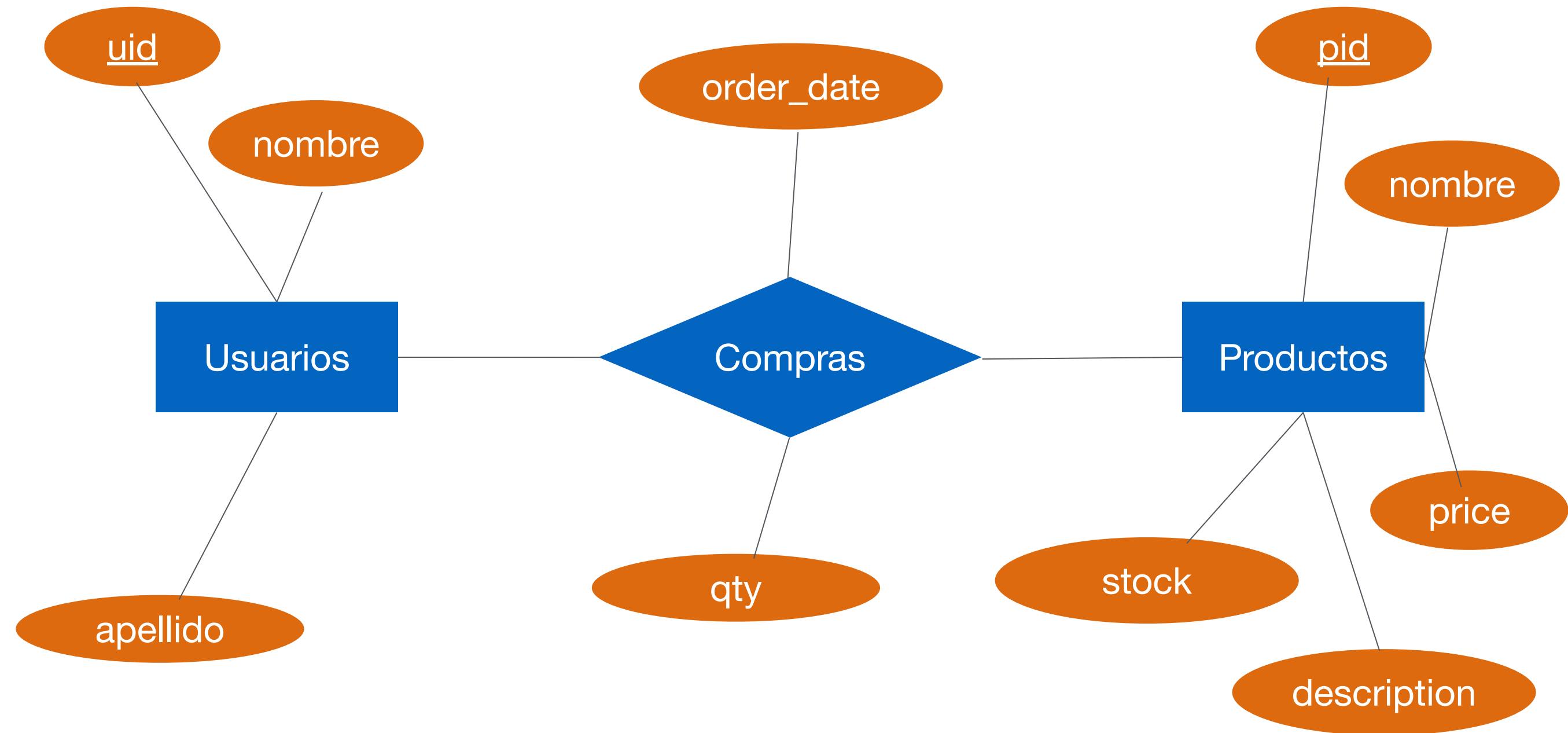
En este curso aprenderemos el lenguaje de consultas del modelo relacional: **SQL**

También aprenderemos el lenguaje teórico sobre el que se sustenta: **álgebra relacional**

Pero también hablaremos de otros sistemas fuera del modelo relacional

Y lo más importante...
Aprenderemos a modelar los
problemas

Modelo Relacional



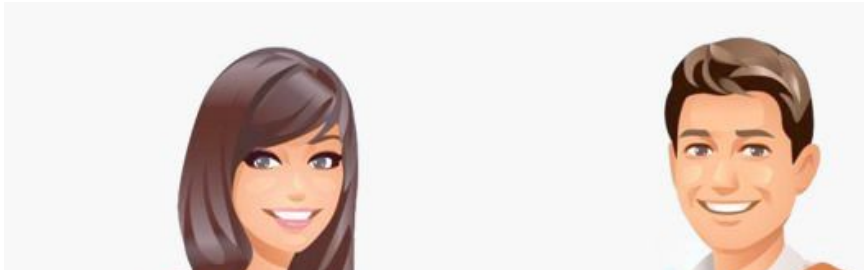
Y también aprenderemos
conceptos avanzados sobre el
lenguaje SQL

En este curso

Aprenderemos a realizar consultas complejas.

Aprenderemos de como el DBMS ejecuta una consulta.

Y varias cosas más fundamentales para el uso práctico de un DBMS



Y qué hay de mi aplicación Web?

Las aplicaciones web suelen conectarse a algún sistema de bases de datos.

SQL + Programación

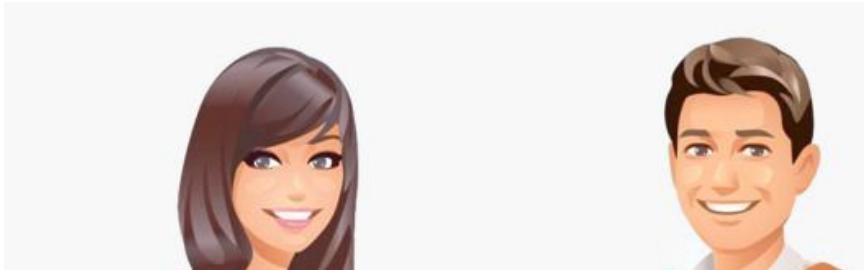
Es posible conectar un DBMS a los lenguajes de programación (Python, Java, C#, etc) y consultar la base de datos de forma programática.

Las aplicaciones que comúnmente usamos, se conectan a una base de datos (muchas de ellas, una base de datos SQL) .

En este curso

Aprenderemos cómo consumir una base de datos desde un lenguaje de programación

En el proyecto, construiremos una aplicación web que hace uso de sistemas de bases de datos



Pero nos contaron que si usábamos un
framework Web no hay que saber usar SQL

Quizás es posible, pero también es posible hacer una aplicación web en una interfaz gráfica sin saber cómo funcionan.

[Funciones](#)[Explora](#)[Suscripciones](#) [Plantillas](#)[Ayuda](#)[Entrar](#)

Todo empieza con tu increíble página web

Wix combina la belleza con la más avanzada tecnología para crear tu increíble página web. Es fácil y gratis.

[Empieza ya](#)

De la misma forma que es posible
armar un puente sin saber cómo
hacerlo





*“El que no domina el lenguaje de consulta nunca
dominará construir aplicaciones de calidad”*

El tío Iroh... creo

En este curso además...

Aprenderemos nociones de cómo funciona un sistema de bases de datos por dentro

- Índices
- Algoritmos internos
- Transacciones
- Recuperación de fallas

¿Es SQL suficiente?

Un sistema puede funcionar lento porque:

- La base de datos está mal indexada
- Los datos están mal modelados
- Hay una cantidad de datos demasiado grande
- ...

¿Es SQL suficiente?

Aplicaciones grandes, como una red social de millones de usuarios, o un laboratorio de millones de registros, necesita técnicas diferentes

Pero ojo, una aplicación típica puede funcionar perfecto con una base de datos SQL!

En este curso

Aprenderemos tópicos sobre bases de datos NoSQL

- JSON y MongoDB
- Búsqueda de texto
- Herramientas de análisis de datos

Y lo más importante, cuando es pertinente usar cada una de estas tecnologías

¿Preguntas?