



Universidad del
Rosario

Educación Continua
y Consultoría

Diplomado en Ciencia de Datos



#URSolucionesInnovadoras

#URConsultoría



Universidad del
Rosario

Educación Continua
y Consultoría

BASES DE DATOS 1

Introducción a las bases de datos relacionales



Universidad del
Rosario

Educación Continua
y Consultoría



Sobre el conferencista



Miguel Angel Orjuela Rocha

Ing. de Sistemas y Computación

Skills:

- Desarrollo Fullstack
- Analítica de datos
- Docencia universitaria

Contacto:



miguela.orjuela@urosario.edu.co



<https://www.linkedin.com/in/miguel-orjuela/>



<https://github.com/maorjuela73>



Universidad del
Rosario

Educación Continua
y Consultoría

Contenido

- Introducción
- Modelos de datos
- El modelo relacional
- Diagramas E-R
- ¿Qué es SQL?
- Entorno de trabajo
- Consultas básicas con SELECT





Universidad del
Rosario

Educación Continua
y Consultoría

Contenido

- **Introducción**
- Modelos de datos
- El modelo relacional
- Diagramas E-R
- ¿Qué es SQL?
- Entorno de trabajo
- Consultas básicas con SELECT





Introducción

Objetivo del módulo

Proporcionar una formación sólida en **bases de datos relacionales** y el **lenguaje SQL** que le permita al estudiante extraer y analizar información de forma eficiente.

Este módulo complementa la formación de científicos de datos para emplear fuentes de datos en bases de datos relacionales, de amplio uso en todos los sectores de la economía.





Introducción

Habilidades a desarrollar



- Automatización de tareas de extracción y transformación de información almacenada en bases de datos relacionadas.
- Montaje de sistemas manejadores de bases de datos relacionales de uso profesional.
- Desarrollo de consultas complejas sobre bases de datos relacionales usando el lenguaje SQL.
- Conexión a bases de datos relacionales con el lenguaje Python.



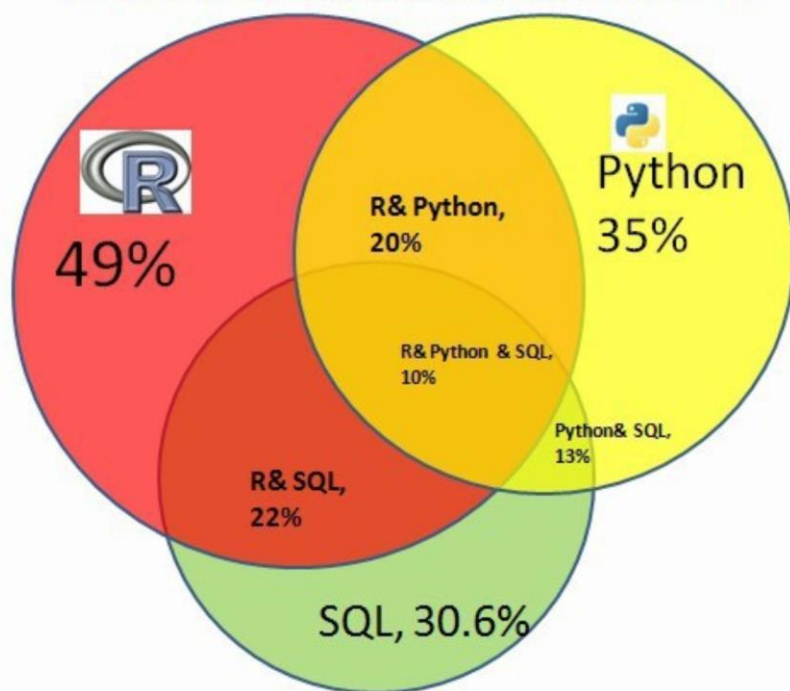
Universidad del
Rosario

Educación Continua
y Consultoría

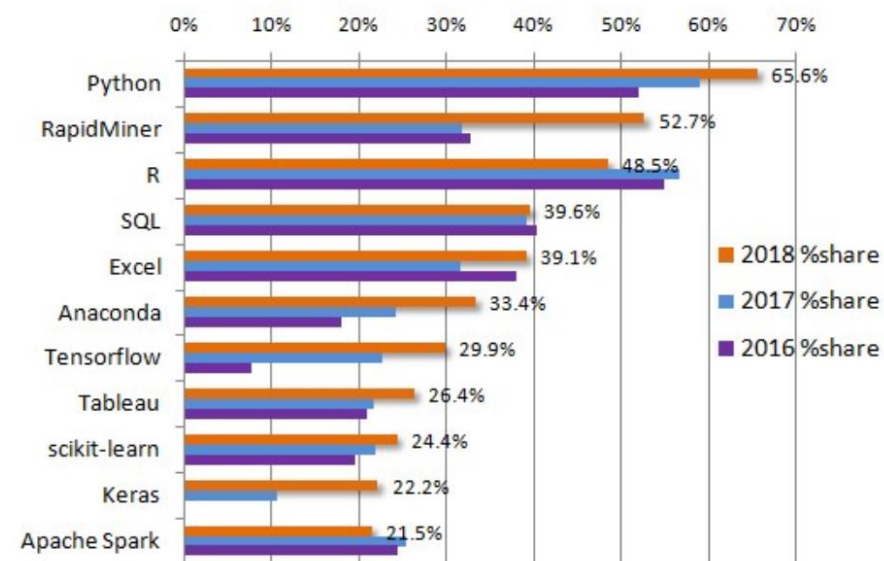
Introducción



KDnuggets 2014 Poll: Languages used for Analytics/Data Mining



KDnuggets Analytics, Data
Science, Machine Learning Software
Poll, 2016-2018





Introducción

SQL: Atemporal Y Eficiente



- SQL es más útil como lenguaje de procesamiento de datos que como herramienta analítica avanzada.
- Gran parte del proceso de la ciencia de datos depende de procesos de ETL (Extract – Transform –Load), y la longevidad y eficiencia de SQL son prueba de que es un lenguaje muy útil para el científico de datos moderno.



Contenido

- Introducción
- **Modelos de datos**
- El modelo relacional
- Diagramas E-R
- ¿Qué es SQL?
- Entorno de trabajo
- Consultas básicas con SELECT





Modelos de datos

Bases de datos

Son contenedores que guardan y organizan la información.

Ejemplos:

- Directorio telefónico
- Listado de pacientes
- Registros de ventas de una empresa





Modelos de datos

Aplicaciones de bases de datos



- Fueron de las primeras aplicaciones desarrolladas para computador.
- Mecanismos para almacenamiento y recuperación de información.
- Almacenan datos de forma electrónica.
- Indexan datos de múltiples maneras.
- Aseguran la información de fallos y accesos no autorizados.



Modelos de datos

Algunos tipos de aplicaciones

- Sistemas de ventas
- Sistemas de contabilidad
- Sistemas de recursos humanos
- Sector de manufactura
- Banca y finanzas
- Universidades
- Aerolíneas
- Telecomunicaciones
- Servicios basados en Web
- Bases de datos de documentos
- Sistemas de navegación





Universidad del
Rosario

Educación Continua
y Consultoría

Modelos de datos

Sobre los datos

- Altamente valiosos.
- Relativamente grandes.
- Son accedidos por múltiples usuarios y aplicaciones.





Modelos de datos

Rol del modelo de datos



- Determinar la estructura lógica de la base de datos.
- Determinar el modo de almacenar, ordenar y manipular los datos.

Algunos modelos de datos:

- Relacional
- Entidad-Relación
- De datos semiestructurados
- Basados en objetos



Modelos de datos

Modelo relacional

Postulado en 1970 por Edgar Frank Codd de IBM Research Labs.

La información se representa en tablas

- Cada tabla tiene filas y columnas.
- Cada columna tiene un nombre único.
- Cada tabla tiene información que identifica únicamente a cada fila.



<i>ID</i>	<i>name</i>	<i>dept_name</i>	<i>salary</i>
22222	Einstein	Physics	95000
12121	Wu	Finance	90000
32343	El Said	History	60000
45565	Katz	Comp. Sci.	75000
98345	Kim	Elec. Eng.	80000

<i>dept_name</i>	<i>building</i>	<i>budget</i>
Comp. Sci.	Taylor	100000
Biology	Watson	90000
Elec. Eng.	Taylor	85000
Music	Packard	80000
Finance	Painter	120000
History	Painter	50000
Physics	Watson	70000



Modelos de datos

Modelo Entidad-Relación

- Usa objetos básicos llamados entidades, y relaciones entre ellos.
- Una entidad es una cosa u objeto que en el modelamiento se diferencia de otros objetos
- Se usa ampliamente en diseño de bases de datos





Modelos de datos

Modelo semiestructurado

Cada individuo puede tener el mismo o diferentes conjuntos de atributos

Ejemplos

- JSON
- XML



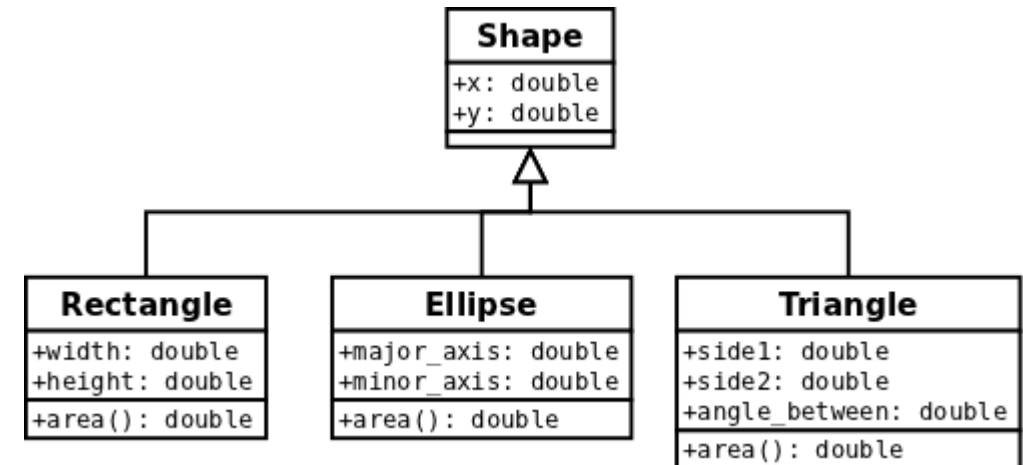
```
{  
  "type": "Feature",  
  "geometry": {  
    "type": "Point",  
    "coordinates": [125.6, 10.1]  
  },  
  "properties": {  
    "name": "Dinagat Islands"  
  }  
}
```



Modelos de datos

Modelo basado en objetos

- Parte del paradigma de programación orientada a objetos.
- Es un modelo relacional extendido a nociones de encapsulamiento.





Universidad del
Rosario

Educación Continua
y Consultoría

Contenido

- Introducción
- Modelos de datos
- **El modelo relacional**
- Diagramas E-R
- ¿Qué es SQL?
- Entorno de trabajo
- Consultas básicas con SELECT





El modelo relacional

Definición

Colección finita de relaciones de dos dimensiones: tuplas-atributos

Características principales:

- Simplicidad en su representación
- Precisión en su implementación
 - Matemáticamente riguroso
 - No ambiguo
- Flexibilidad en cuanto a estructura y contenido





El modelo relacional

Principios básicos del modelo



Una base de datos consiste en una o más relaciones.

El conjunto de esquemas de las relaciones de una base de datos se llama esquema de base de datos relacional, o solo esquema de base de datos.



El modelo relacional

Principios básicos del modelo

El modelo relacional requiere que cada componente de cada tupla sea atómico.

Cada atributo tiene un tipo de datos, llamado *dominio*

Movies(title:string, year:integer, length:integer, genre:string)



Atributo: Columna de la relación

<i>title</i>	<i>year</i>	<i>length</i>	<i>genre</i>
Gone With the Wind	1939	231	drama
Star Wars	1977	124	sciFi
Wayne's World	1992	95	comedy

Tupla: Fila de la relación



El modelo relacional

Principios básicos del modelo

El modelo relacional tiene una restricción importante para su correcto funcionamiento: Cada tabla DEBE TENER información que identifique **únicamente** a cada fila.

Una *llave* es el conjunto de atributos que identifica de forma única a cada fila de la relación

Ejemplo: Llave de Movies

Movies(title, year, length, genre)



Llave

<i>title</i>	<i>year</i>	<i>length</i>	<i>genre</i>
Gone With the Wind	1939	231	drama
Star Wars	1977	124	sciFi
Wayne's World	1992	95	comedy



El modelo relacional

Términos clave

- **Relación:** Una tabla
- **Tabla:** Conjunto de filas
- **Fila:** Conjunto de columnas que describen completamente una entidad. También se le llama *registro* o *tupla*
- **Columna:** Conjunto de información individual que describe a los elementos de interés. También se les llama *atributos*
- **Dominio:** Tipo de dato de cada columna
- **Llave primaria:** Identificador único de fila en una tabla
- **Llave foránea:** Identificador de fila en otra tabla
- **Esquema** de base de datos: Diseño lógico de la base de datos



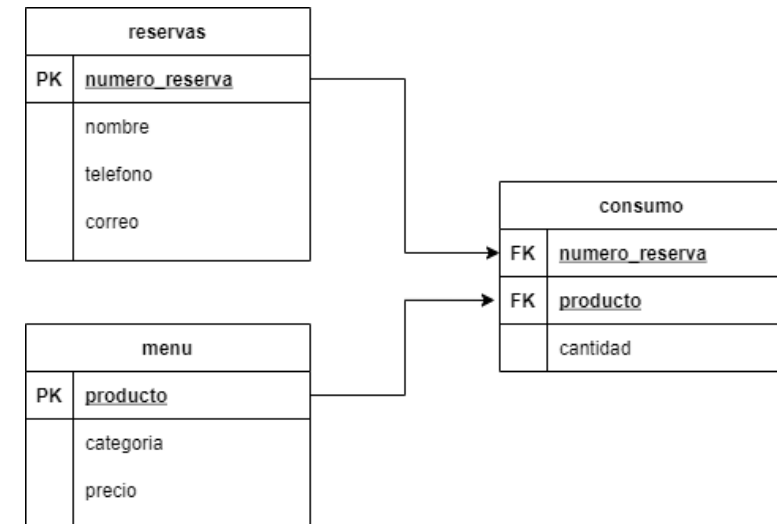


El modelo relacional

Llave primaria (PK)

Atributo o grupo de atributos que identifican de forma única cada tupla de una relación

- Toda relación debe tener una única PK
- Los atributos que forman la PK no pueden ser nulos





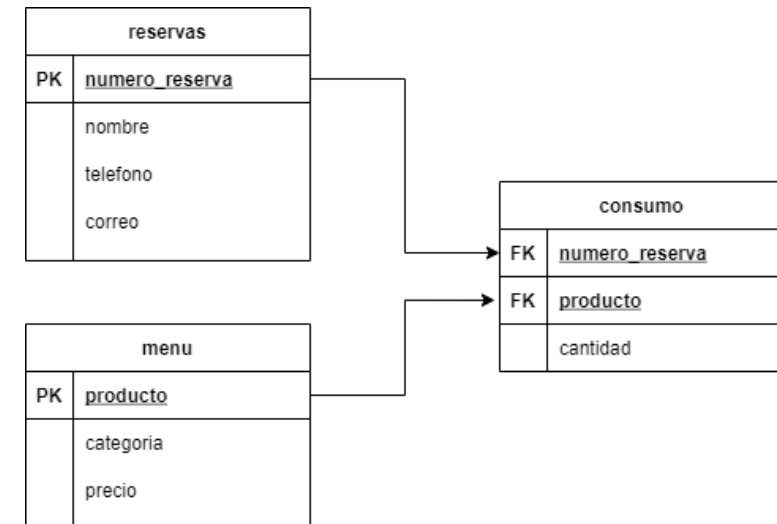
El modelo relacional

Llave foránea (FK)

Si una relación r_1 tiene dentro de sus atributos la llave primaria de una relación r_2 , ese atributo es una llave foránea de r_1 , referenciando a r_2

- r_1 es la relación que referencia.
- r_2 es la relación referenciada.

Se pueden configurar restricciones de integridad referencial

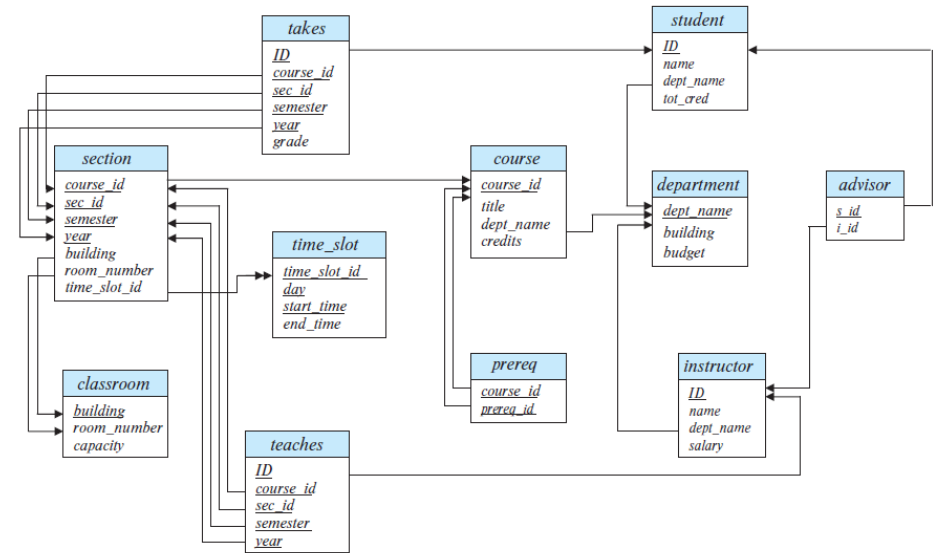




El modelo relacional

Esquema de base de datos

- Es el diseño lógico de la base de datos
- Cuando se le saca un pantallazo al esquema con datos, se le conoce como una **instancia** de la base de datos





Universidad del
Rosario

Educación Continua
y Consultoría

Contenido

- Introducción
- Modelos de datos
- El modelo relacional
- **Diagramas E-R**
- ¿Qué es SQL?
- Entorno de trabajo
- Consultas básicas con SELECT





Diagramas E-R

Modelo Entidad Relación (E-R)

Facilita el diseño permitiendo especificar la estructura lógica de las bases de datos

Tres conceptos básicos del modelo

- Conjuntos de **entidades**
- Conjuntos de **relaciones**
- **Atributos**





Diagramas E-R

Diferencias entre modelo relacional y E-R

	MODELO ER	MODELO RELACIONAL
Labor	Representa la colección de objetos llamada entidades y la relación entre esas entidades.	Representa la colección de tablas y la relación entre esas tablas.
Describir	El Modelo de relación de entidad describe los datos como conjunto de entidades, conjunto de relaciones y atributo.	El Modelo Relacional describe los datos en una tabla como Dominio, Atributos, Tuplas.
Relación	ER Model es más fácil de entender la relación entre las entidades.	Comparativamente, es menos fácil derivar una relación entre tablas en el Modelo Relacional.
Cartografía	El modelo de ER describe la asignación de cardinalidades.	El Modelo Relacional no describe las cardinalidades de mapeo.





Diagramas E-R

Conjunto de entidades

Entidad: Cosas u objetos del mundo real que se distinguen de otros objetos.

Conjunto de entidades: Entidades del mismo tipo.
Cada miembro del conjunto tiene las mismas propiedades o atributos.



Estudiante

Tipo de entidad

E1
E2
E3

Conjunto de entidades



Diagramas E-R

Atributos



- Cada entidad es representada por un conjunto de atributos
- Un atributo es una propiedad descriptiva que tiene cada miembro del conjunto de entidades
- Cada entidad tiene un **valor** para cada uno de sus atributos

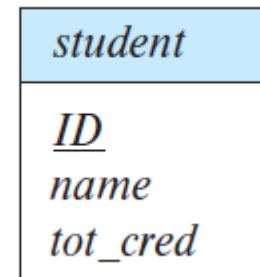
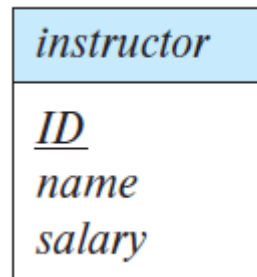


Diagramas E-R

Entidades y atributos

Ejemplo

Un diagrama en una notación cualquiera que muestra los conjuntos de entidades *instructor* y *student*.





Diagramas E-R

Conjuntos de relaciones

- Una relación es una asociación entre varias entidades
- Un conjunto de relaciones está compuesto por relaciones del mismo tipo



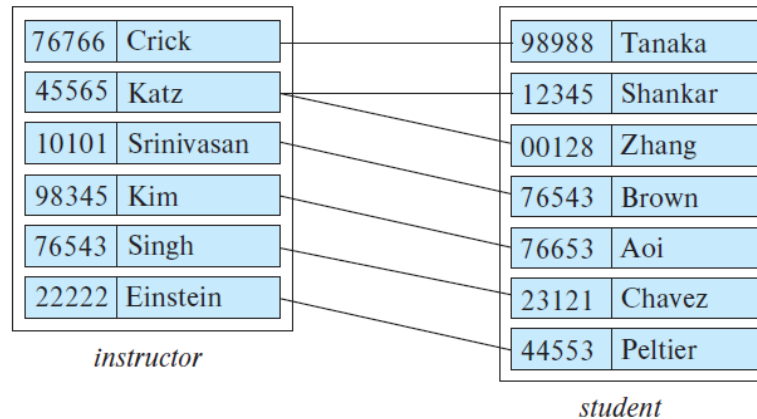


Diagramas E-R

Conjuntos de relaciones

Ejemplo

Relación *advisor*.



Instancia de relación

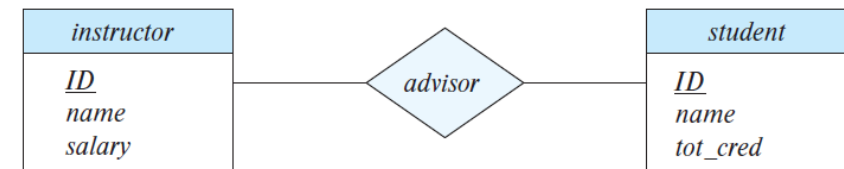


Diagrama E-R con la relación





Diagramas E-R

Tipos de relaciones según su cardinalidad

- **Uno a uno:** Gerente de punto y punto de venta
- **Uno a muchos:** Cliente y pedidos
- **Muchos a muchos:** Estudiantes y clases





Universidad del
Rosario

Educación Continua
y Consultoría

Diagramas E-R

Notaciones

- Notación Chen
- Notación UML
- Notación Crow's foot

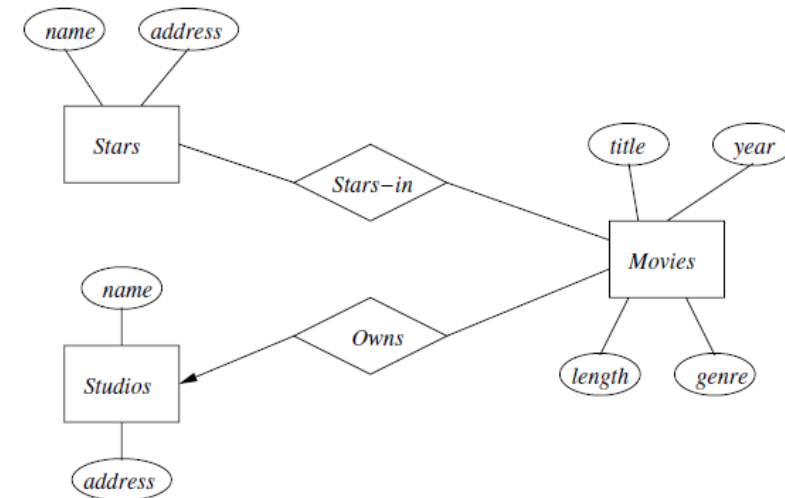




Diagramas E-R

Notación Chen

- **Cuadro:** Entidad
- **Rombo:** Asociación entre entidades
- **Elipse:** Atributo en una relación





Diagramas E-R

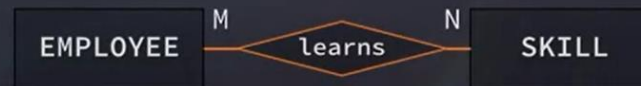
Notación Chen



A One-to-Many (1:M) Relationship: a PAINTER can paint many PAINTINGs; each PAINTING is painted by one PAINTER.



A Many-to-Many (M:N) Relationship: an EMPLOYEE can learn many SKILLs; each SKILL can be learned by many EMPLOYEEs.



A One-to-One (1:1) Relationship: an EMPLOYEE manages one STORE; each STORE is managed by one EMPLOYEE.

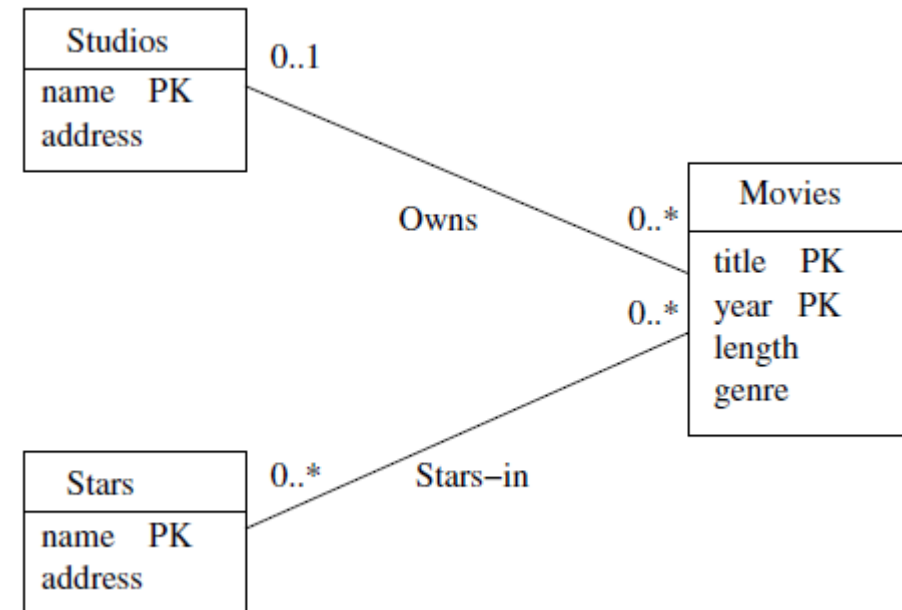




Diagramas E-R

Notación UML

- Las asociaciones se representan con una línea directa.
- Las restricciones de número de objetos en cada relación se representa con **m...n**



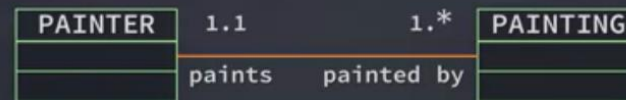


Diagramas E-R

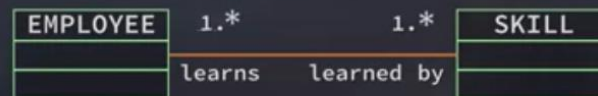
Notación UML



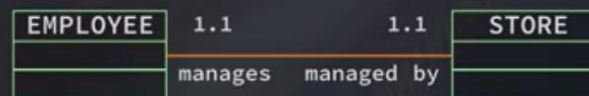
A One-to-Many (1:M) Relationship: a PAINTER can paint many PAINTINGs; each PAINTING is painted by one PAINTER.



A Many-to-Many (M:N) Relationship: an EMPLOYEE can learn many SKILLs; each SKILL can be learned by many EMPLOYEEs.



A One-to-One (1:1) Relationship: an EMPLOYEE manages one STORE; each STORE is managed by one EMPLOYEE.

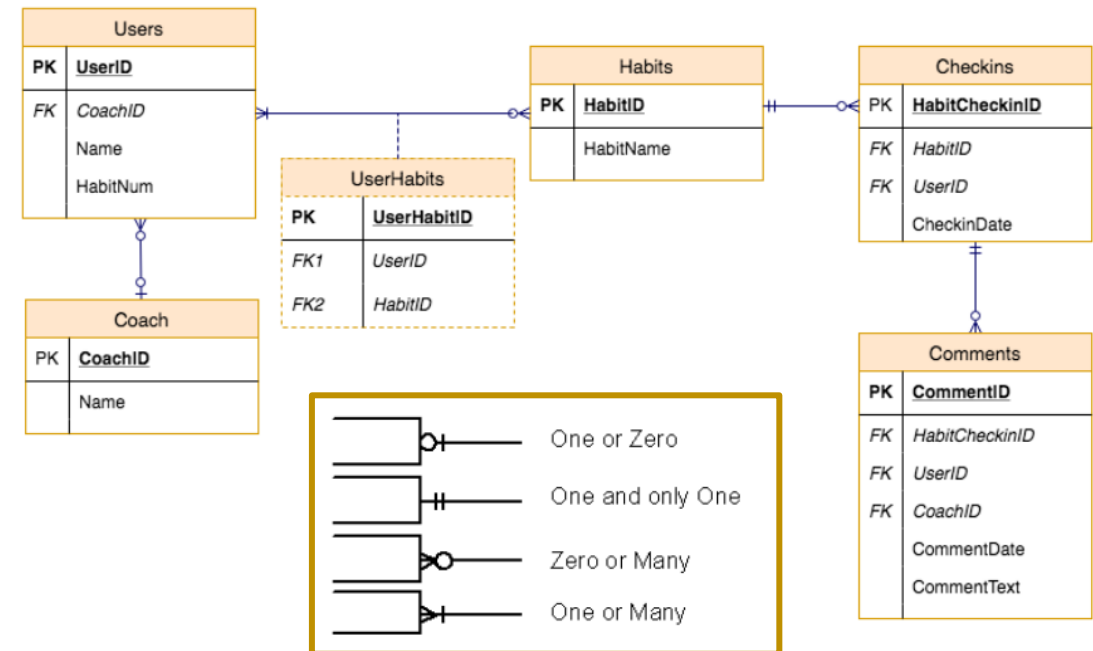




Diagramas E-R

Notación Crow's foot

- Las asociaciones se representan con una línea directa.
- Las restricciones de número de objetos en cada relación se representa con líneas y círculos al final de cada línea que conecta relaciones.



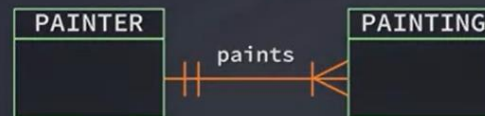


Diagramas E-R

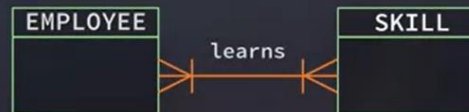
Notación Crow's Foot



A One-to-Many (1:M) Relationship: a PAINTER can paint many PAINTINGs; each PAINTING is painted by one PAINTER.



A Many-to-Many (M:N) Relationship: an EMPLOYEE can learn many SKILLs; each SKILL can be learned by many EMPLOYEEs.



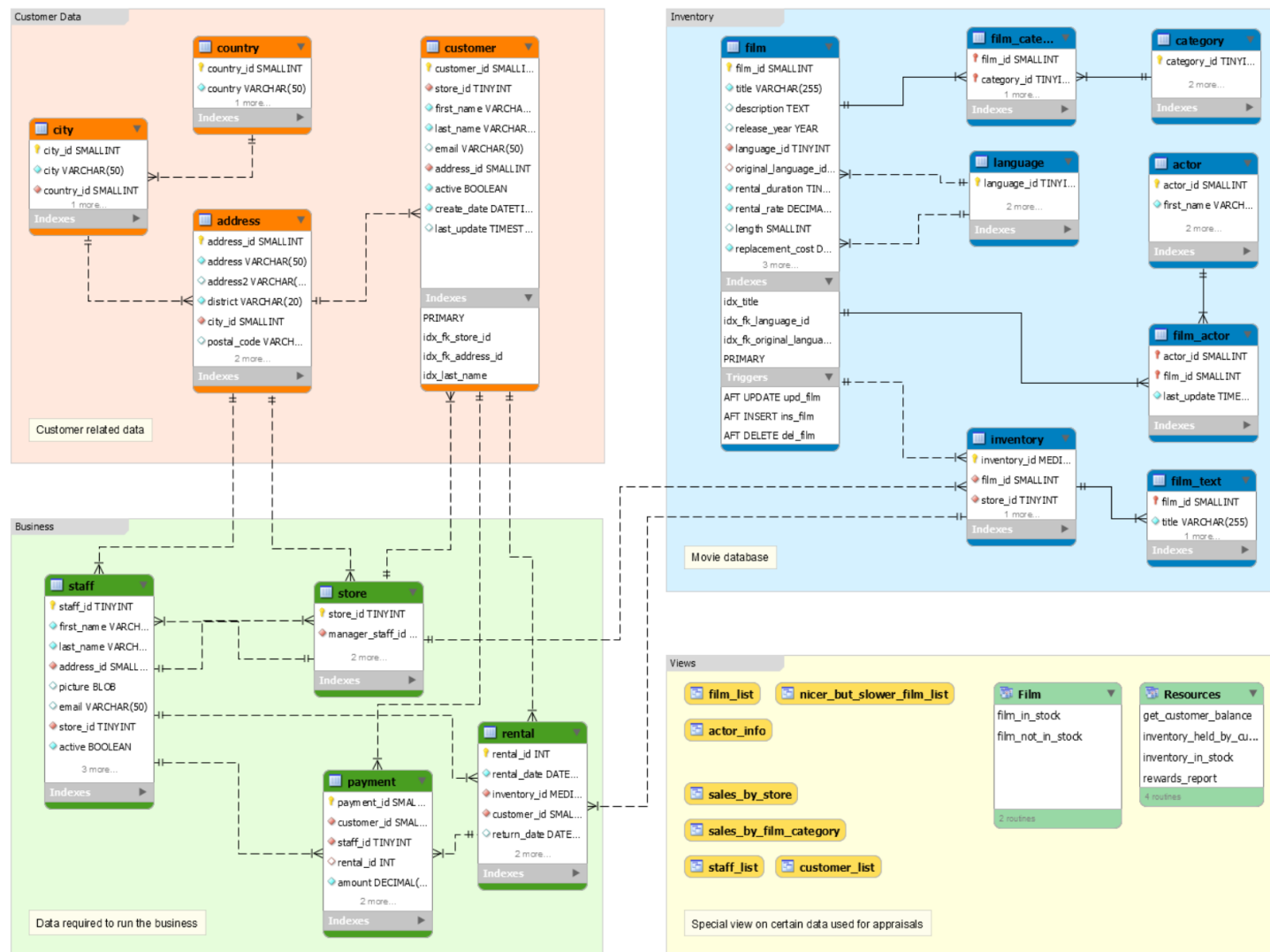
A One-to-One (1:1) Relationship: an EMPLOYEE manages one STORE; each STORE is managed by one EMPLOYEE.





Diagramas E-R

Ejemplo de un modelo E-R





Universidad del
Rosario

Educación Continua
y Consultoría

Contenido

- Introducción
- Modelos de datos
- El modelo relacional
- Diagramas E-R
- **¿Qué es SQL?**
- Entorno de trabajo
- Consultas básicas con SELECT





¿Qué es SQL?

Definición

Structured Query Language (SQL) es un lenguaje estandar de computador para administración de bases de datos relacionales y manipulación de datos.

- Se emplea para comunicarse con la base de datos
- Las sentencias de SQL tienen palabras clave en inglés altamente descriptivas.
- NO es un lenguaje procedimental.
- NO se pueden escribir aplicaciones completas con SQL.





¿Qué es SQL?

Componentes principales

Lenguaje de definición de datos (**DDL**)

- Permite definir la estructura de la base de datos

Lenguaje de manipulación de datos (**DML**)

- Son comandos que no afectan la estructura de la base de datos
- Pueden afectar el contenido de la base de datos o extraer información de la base de datos





Universidad del
Rosario

Educación Continua
y Consultoría

¿Qué es SQL?

Sistemas manejadores de bases de datos (DBMS)

Es una colección de **software muy específico**, cuya función es servir de **interfaz** entre la base de datos, el usuario y las distintas aplicaciones utilizadas.





Universidad del
Rosario

Educación Continua
y Consultoría

¿Qué es SQL?

Sistemas manejadores de bases de datos (DBMS)

El lenguaje SQL es un estándar (**ISO/IEC 9075**) pero cada DBMS tiene su propio dialecto.





Universidad del
Rosario

Educación Continua
y Consultoría

Contenido

- Introducción
- Modelos de datos
- El modelo relacional
- Diagramas E-R
- ¿Qué es SQL?
- **Entorno de trabajo**
- DML Parte 1: Consultas básicas con SELECT



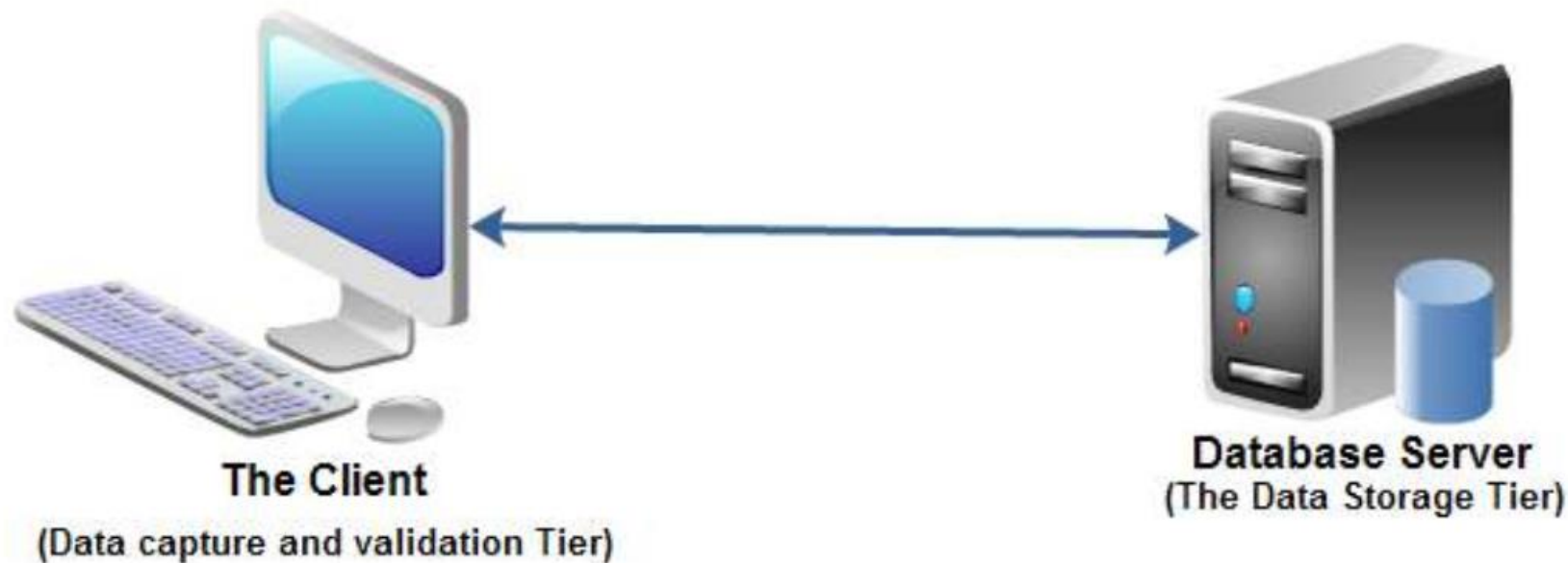


Universidad del
Rosario

Educación Continua
y Consultoría

El entorno de trabajo

Arquitectura Cliente - Servidor



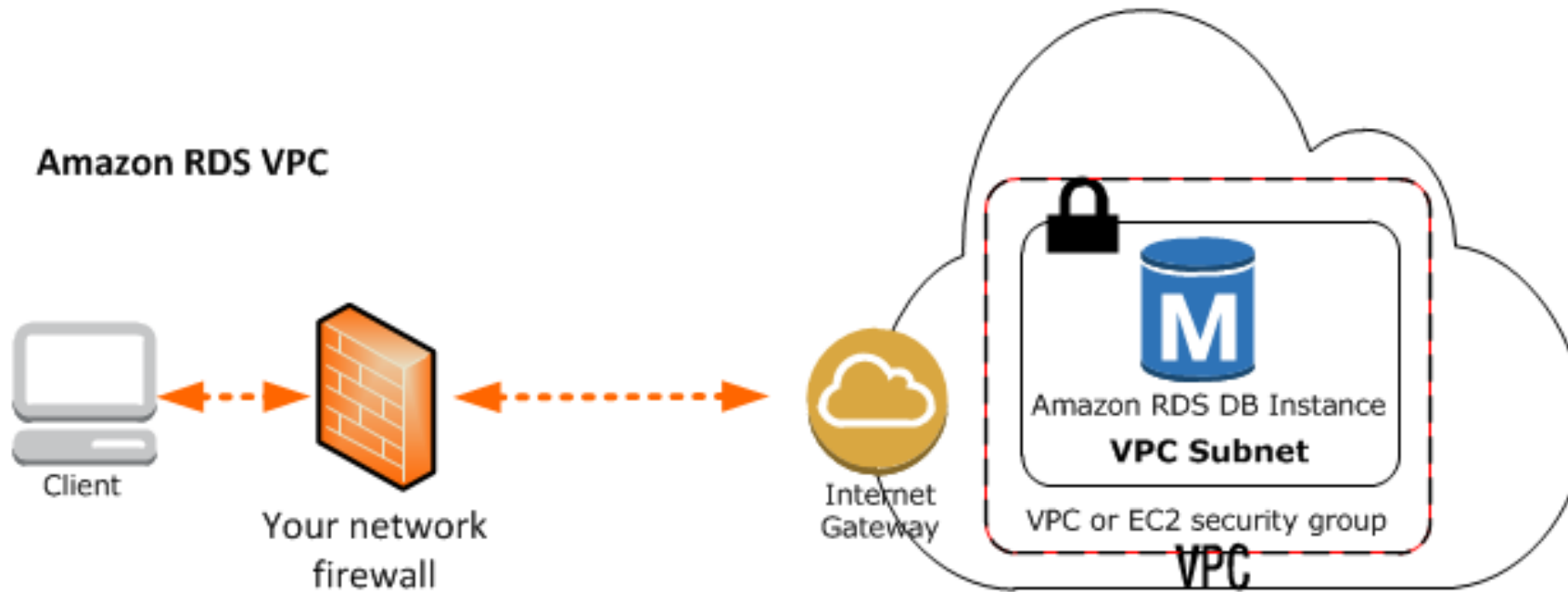


Universidad del
Rosario

Educación Continua
y Consultoría

El entorno de trabajo

Arquitectura Cliente - Servidor





El entorno de trabajo

Cliente



Es un programa que permite conectarse al servidor de base de datos
Puede ser de comandos (**CLI**) o tener interfaz gráfica (**GUI**)

Debe por lo menos solicitar la siguiente información:

- **Server:** Dirección IP del servidor
- **Port:** Puerto de conexión
- **User:** Nombre de usuario autorizado
- **Pass:** Contraseña de conexión

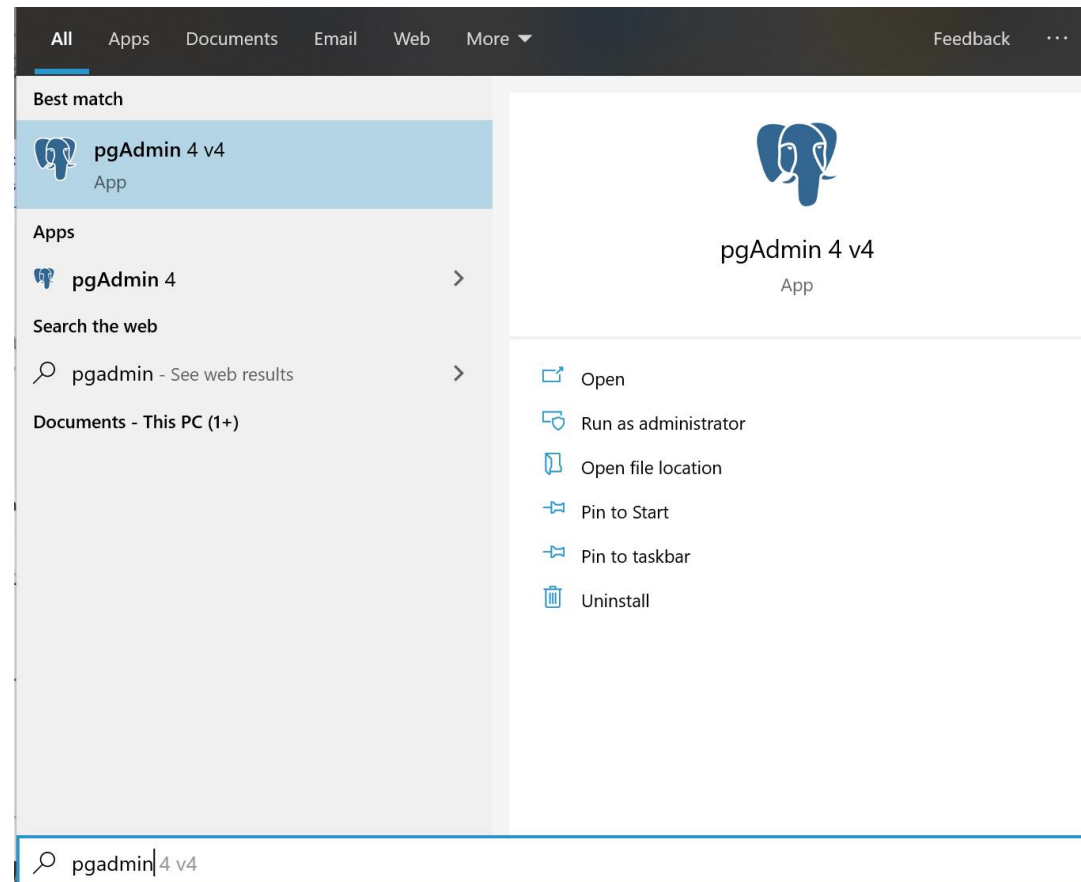


Universidad del
Rosario

Educación Continua
y Consultoría

El entorno de trabajo

Cliente



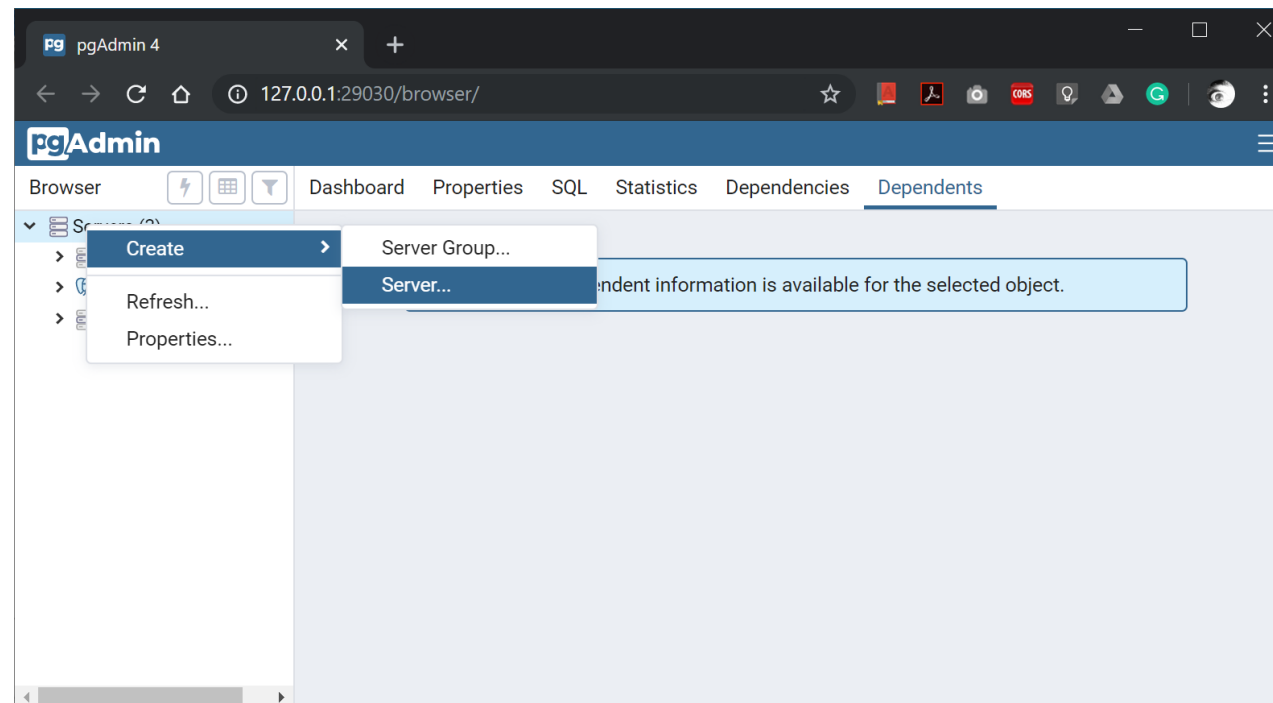


Universidad del
Rosario

Educación Continua
y Consultoría

El entorno de trabajo

Cliente



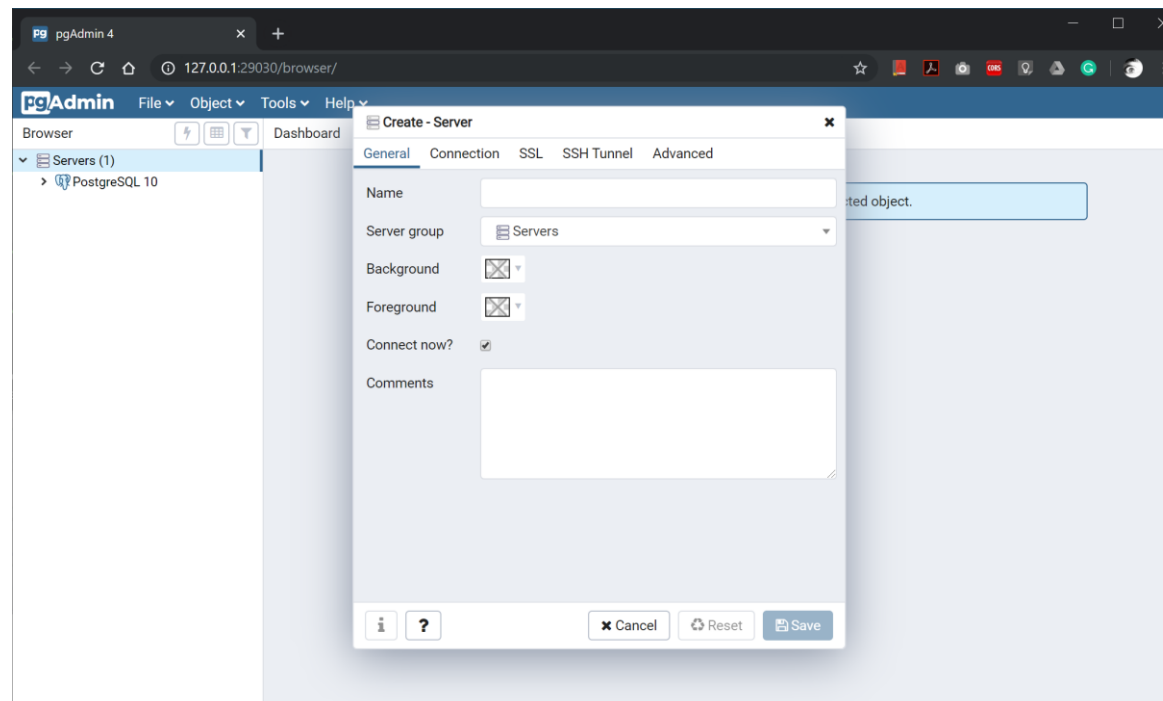


Universidad del
Rosario

Educación Continua
y Consultoría

El entorno de trabajo

Cliente



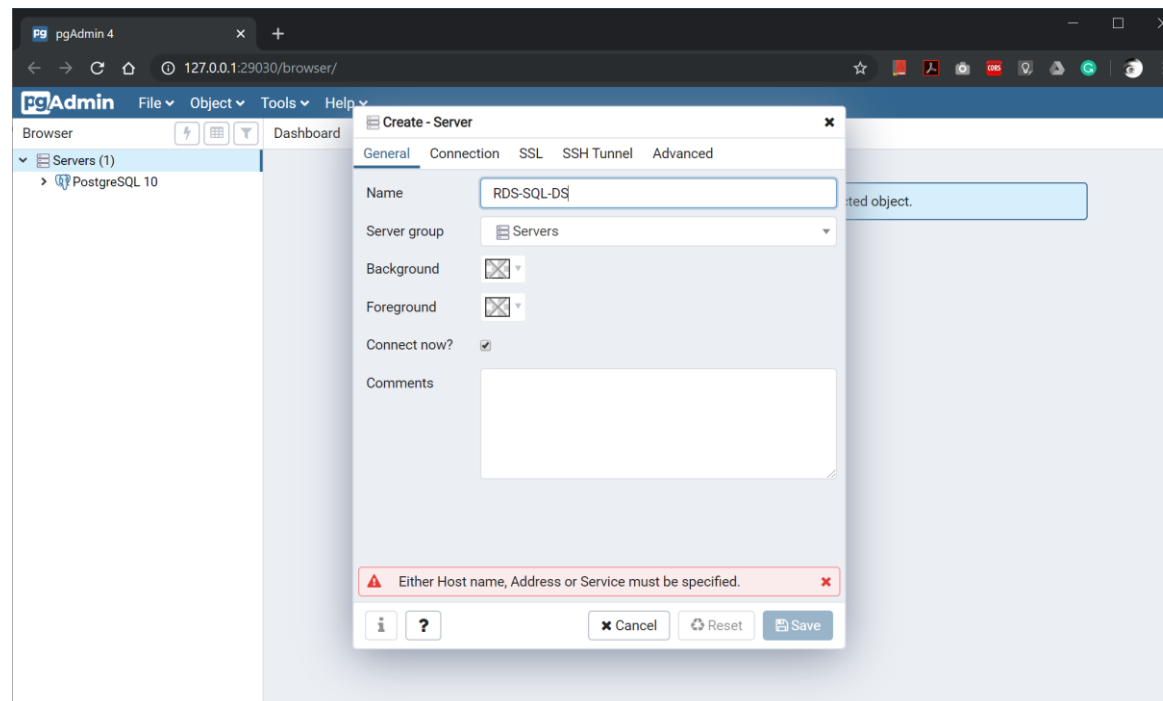


Universidad del
Rosario

Educación Continua
y Consultoría

El entorno de trabajo

Cliente



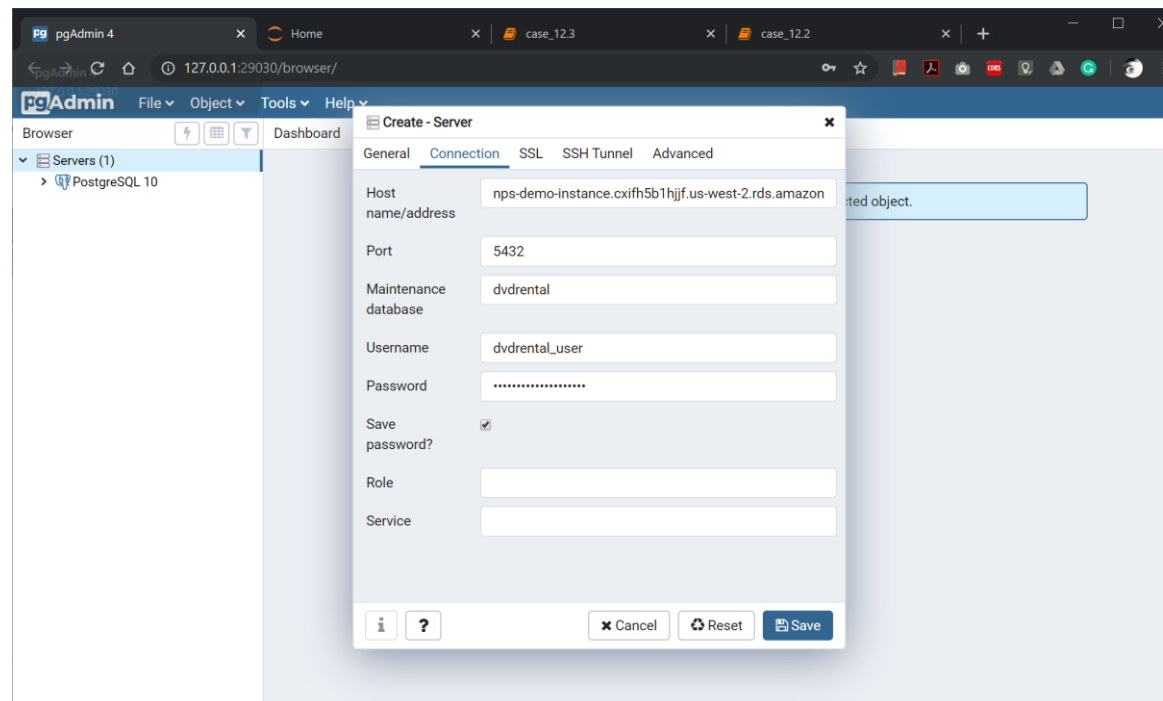


Universidad del
Rosario

Educación Continua
y Consultoría

El entorno de trabajo

Cliente





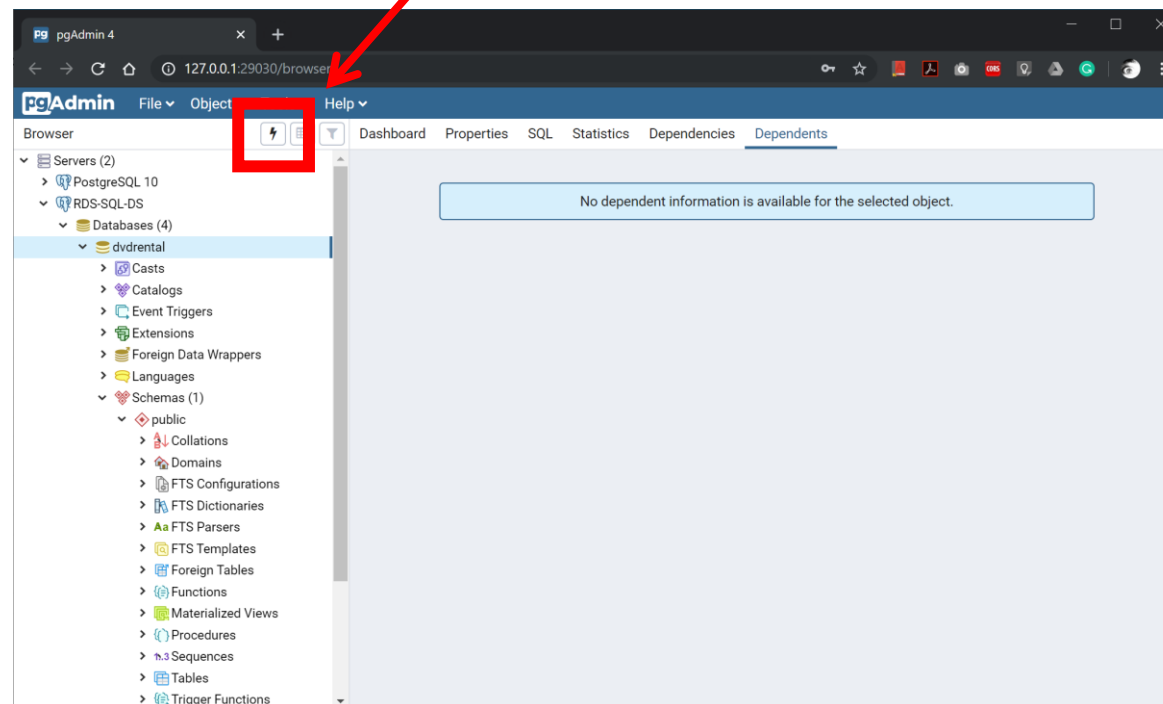
Universidad del
Rosario

Educación Continua
y Consultoría

El entorno de trabajo

Cliente

Editor de consultas





Universidad del
Rosario

Educación Continua
y Consultoría

El entorno de trabajo

El entorno de trabajo



Credenciales de conexión:

Server: database-1.cdqqgt4oejnt.us-west-2.rds.amazonaws.com

Port: 5432

User: postgres

Pass: MEeLaN2z



Universidad del
Rosario

Educación Continua
y Consultoría

Contenido

- Introducción
- Modelos de datos
- El modelo relacional
- Diagramas E-R
- ¿Qué es SQL?
- Entorno de trabajo
- **Consultas básicas con SELECT**



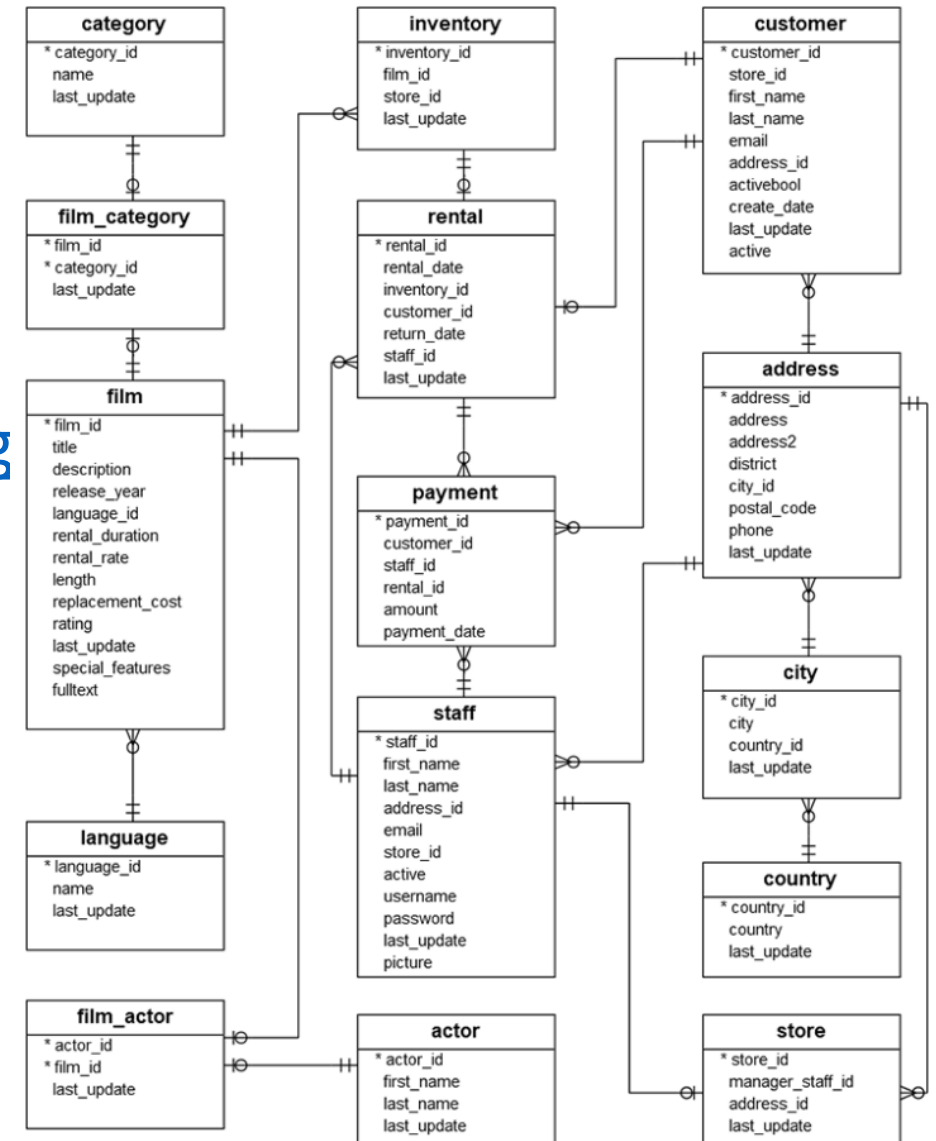


Consultas básicas con SELECT

Sobre el modelo a consultar

<https://www.postgresqltutorial.com/postgresql-sample-database/>

DVD Rental ER Model





Consultas básicas con SELECT

Recuperar todas las filas y todas las columnas



```
SELECT  
  Code,  
  Name,  
  Continent,  
  Region,  
  SurfaceArea,  
  IndepYear,  
  Population,  
  LifeExpectancy,  
  GNP,  
  GNPOld,  
  LocalName,  
  GovernmentForm,  
  HeadOfState,  
  Capital,  
  Code2  
FROM  
  country;
```



```
SELECT *  
FROM country;
```




Consultas básicas con **SELECT**

Recuperar todas las filas y columnas específicas

```
SELECT  
Code,  
Name,  
Continent,  
Population,  
LifeExpectancy  
FROM  
country;
```





Consultas básicas con SELECT

Uso del **DISTINCT**



```
SELECT Continent  
FROM country;
```

Entrega todos los
resultados, así tengan
columnas repetidas

Elimina los repetidos

```
SELECT DISTINCT Continent  
FROM country;
```



Consultas básicas con SELECT

Recuperar datos con SELECT

Columnas calculadas/derivadas

- Se pueden realizar operaciones en una columna
- La columna toma el nombre de la operación (por defecto)

```
SELECT Name, SurfaceArea, Population, Population/SurfaceArea  
FROM country;
```

- *Alias AS* en el nombre de la columna

```
SELECT Name, SurfaceArea, Population, Population/SurfaceArea AS HabsPorKm2  
FROM country;
```

Unidad de Educación Continua y Consultoría
construimos país desde

#URSolucionesInnovadoras
#URConsultoría



@RosarioContinua



/EduContinuaURosario



@RosarioContinua