

Curso Data Engineer: Creando un pipeline de datos

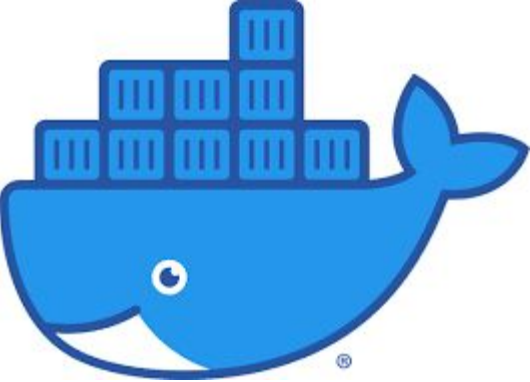
MÓDULO DOCKER SQL - Clase 1

Agenda

- Presentación del ejercicio por equipos (Max. 15 min.)
- ¿Qué es Docker?
- ¿Para que se utiliza?
- Comandos básicos
- DBeaver
- Realizar Queries



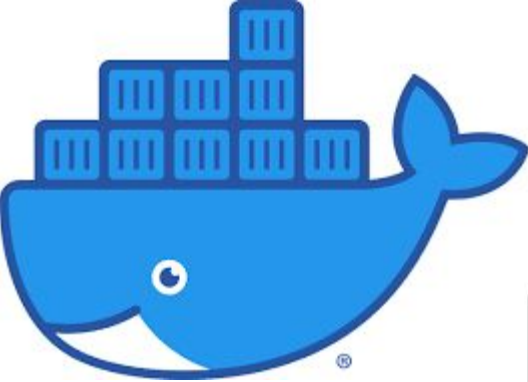
¿Qué es
Docker?



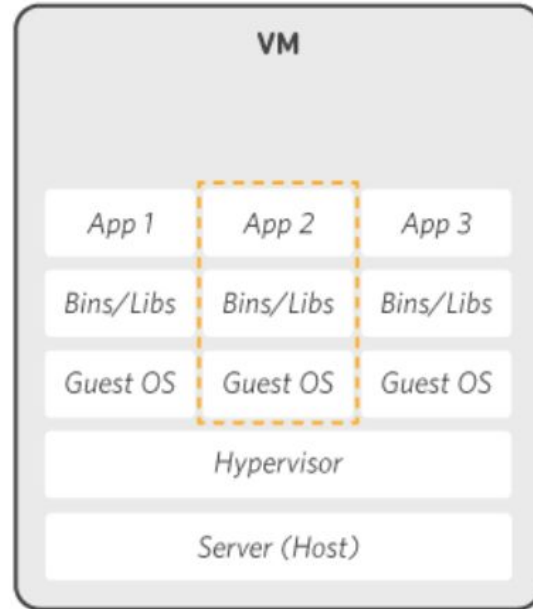
Docker

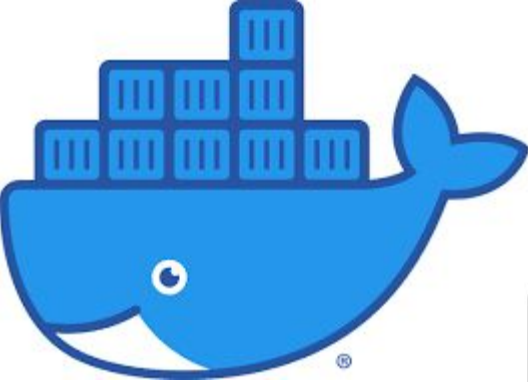


Docker, utiliza los contenedores como máquinas virtuales muy livianas y modulares, y obtiene la flexibilidad necesaria para crear Apps, implementarlas, copiarlas y trasladarlas de un entorno a otro

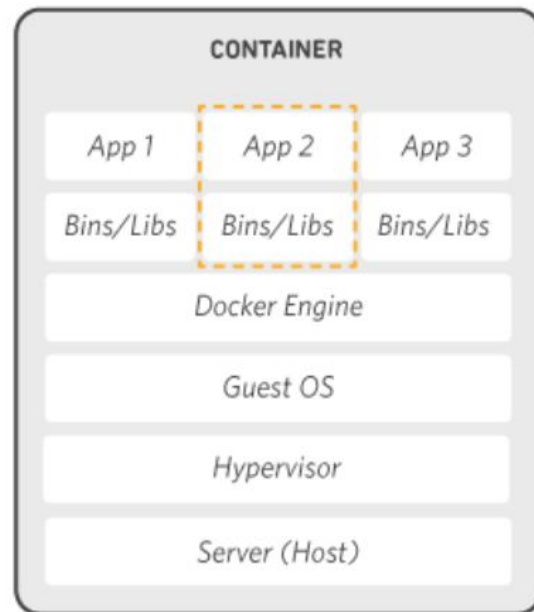
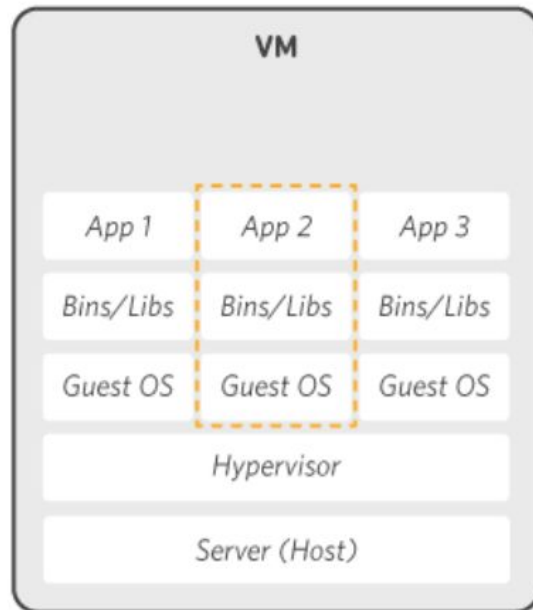


Docker



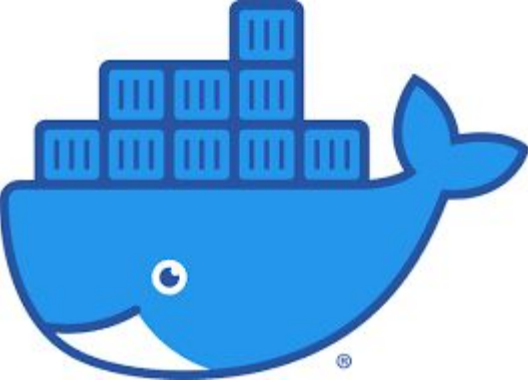


Docker





¿Para que se
utiliza?



Docker



MICROSERVICIOS

Cree y escale arquitecturas de aplicaciones distribuidas al utilizar las implementaciones de código estandarizadas que los contenedores de Docker proporcionan.

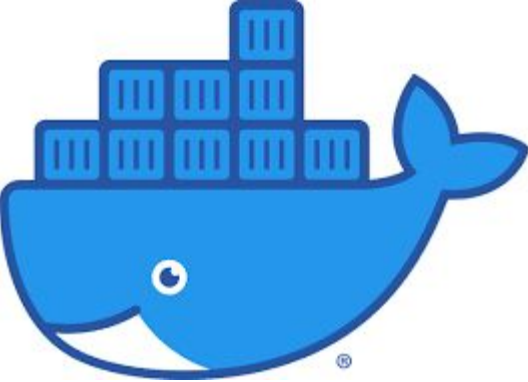


PROCESAMIENTO DE DATOS

Proporcione como servicio el procesamiento de big data. Datos de paquetes y paquetes de análisis en contenedores portátiles que pueden ejecutar usuarios sin conocimientos técnicos.



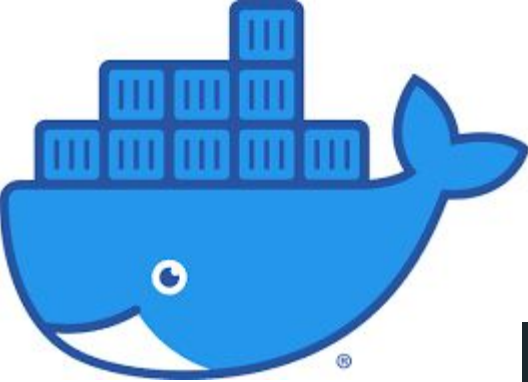
Comandos básicos



Docker



- Instalación
- Probar instalación
- Bajar imagen del repositorio (Docker Hub)
- Correr la imagen
- Listar imágenes/Dockers
- Borrar imágenes/Dockers
- Listar características del Docker
- Copiar files a Docker
- Guardar cambios a las imágenes



Instalación



<https://docs.docker.com/engine/install/>

Supported platforms

Docker Engine is available on a variety of Linux platforms, macOS and Windows 10 through Docker Desktop, and as a static binary installation. Find your preferred operating system below.

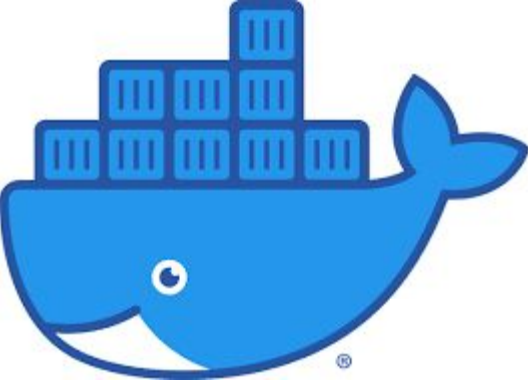
Desktop

Platform	x86_64 / amd64	arm64 (Apple Silicon)
Docker Desktop for Linux	✓	
Docker Desktop for Mac (macOS)	✓	✓
Docker Desktop for Windows	✓	

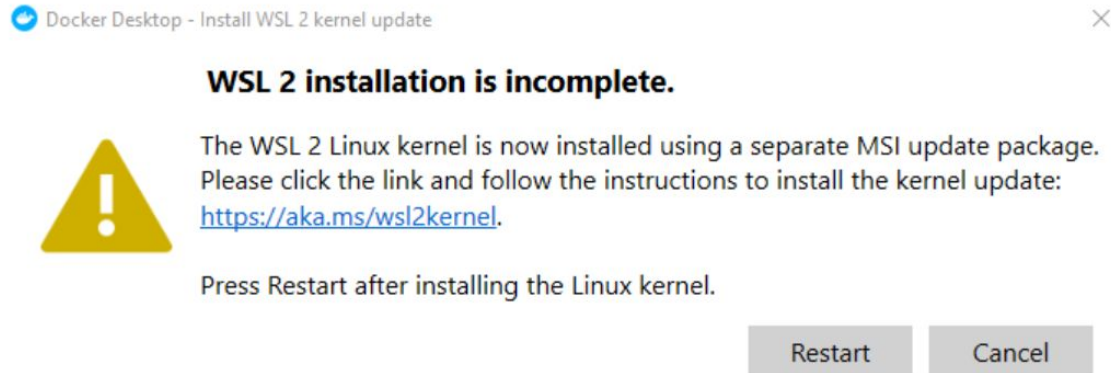
Server

Docker provides `.deb` and `.rpm` packages from the following Linux distributions and architectures:

Platform	x86_64 / amd64	arm64 / aarch64	arm (32-bit)	s390x
CentOS	✓	✓		
Debian	✓	✓	✓	
Fedora	✓	✓		
Raspbian			✓	
RHEL				✓
SLES				✓
Ubuntu	✓	✓	✓	✓
Binaries	✓	✓	✓	

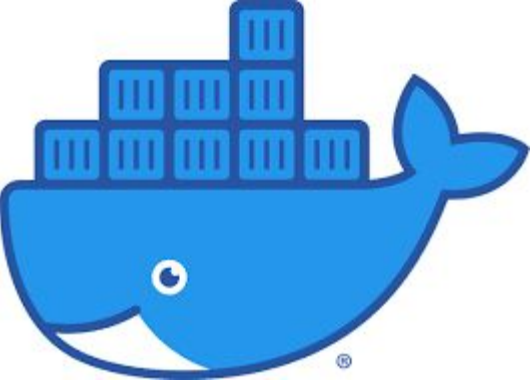


Instalación



WSL 2 = Windows Subsystem for Linux

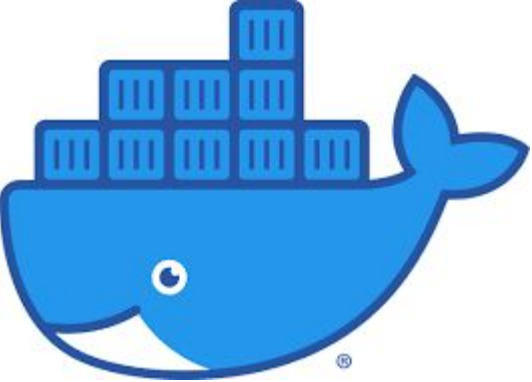
WSL2 nos permite ejecutar el kernel nativo de Linux al completo ejecutándose desde un equipo Windows.



info de Docker



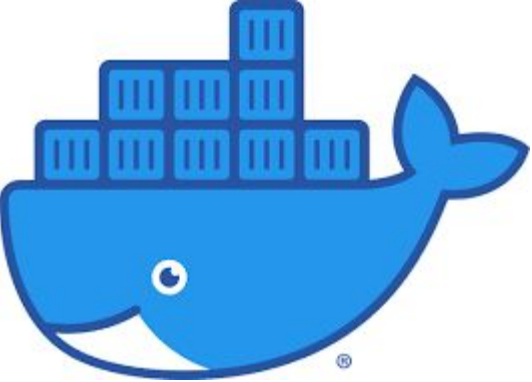
docker info



Probar la Instalación



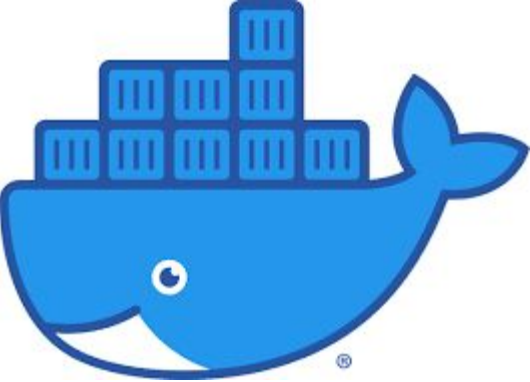
`docker pull hello-world`



Bajar imagen



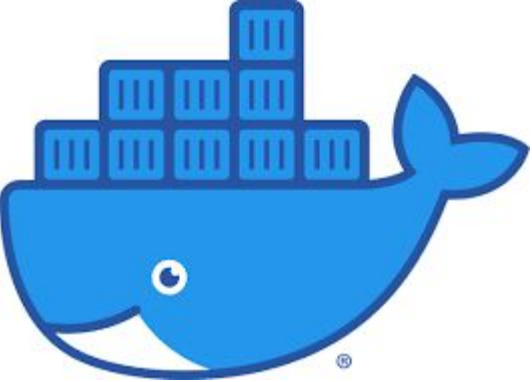
```
docker pull fedepineyro/edvai_postgres:v1
```



Correr imagen



```
docker run --name edvai_postgres -e  
POSTGRES_PASSWORD=edvai -d -p 5432:5432  
fedepineyro/edvai_postgres:v1
```

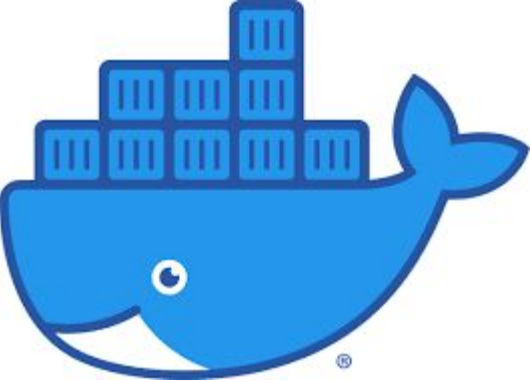
Listar imágenes/contenedores



docker images

docker ps

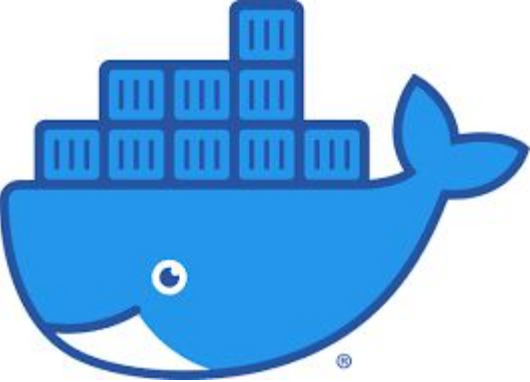
docker ps -a



Listar características



`docker inspect <contenedor>`



Copiar files a Docker

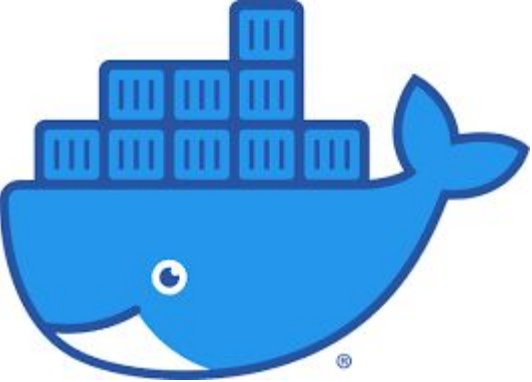


wget <https://data-engineer-edvai.s3.amazonaws.com/2021-informe-ministerio.csv>

curl -o 2021-informe-ministerio.csv

<https://data-engineer-edvai.s3.amazonaws.com/2021-informe-ministerio.csv>

docker cp 2021-informe-ministerio.csv edvai_postgres:/home

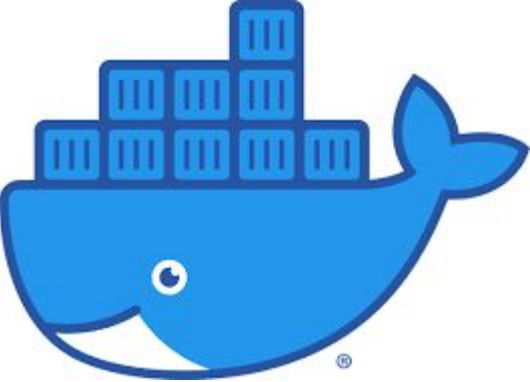


Borrar imágenes/contenedores



docker rmi <imagen>

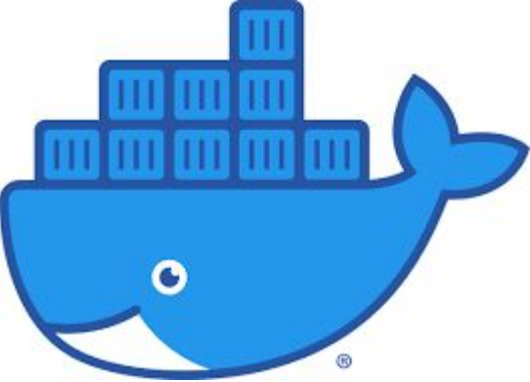
docker rm <contenedor>



Guardar cambios a las imágenes



```
docker commit edvai_postgres fedepineyro/edvai_postgres:v1
```



Otros comandos



parar / detener contenedor: `docker start/stop <contenedor>`

Ingresar al shell del contenedor: `docker exec -it <contenedor> bash`



DBeaver



DBeaver

Instalación: deberás instalar esta herramienta para gestionar tu BD.

(<https://dbeaver.io/download/>)

Windows

- [Windows \(installer\)](#)
- [Windows \(zip\)](#)
- [Install from Microsoft Store](#)
- [Chocolatey](#) (`choco install dbeaver`)

Mac OS X

- [MacOS for Intel \(dmg\)](#)
- [MacOS for Apple Silicon \(dmg\)](#)
- [Brew Cask](#) (`brew install --cask dbeaver-community`)
- [MacPorts](#) (`sudo port install dbeaver-community`)

Linux

- [Linux Debian package \(installer\)](#)
- [Linux RPM package \(installer\)](#)
- [Linux \(zip\)](#)
- [Linux \(zip without Java included\)](#)
- [Linux ARM \(zip without Java included\)](#)
- [Snap](#) (`sudo snap install dbeaver-ce`)
- [Flatpak](#) (`flatpak install flathub io.dbeaver.DBeaverCommunity`)

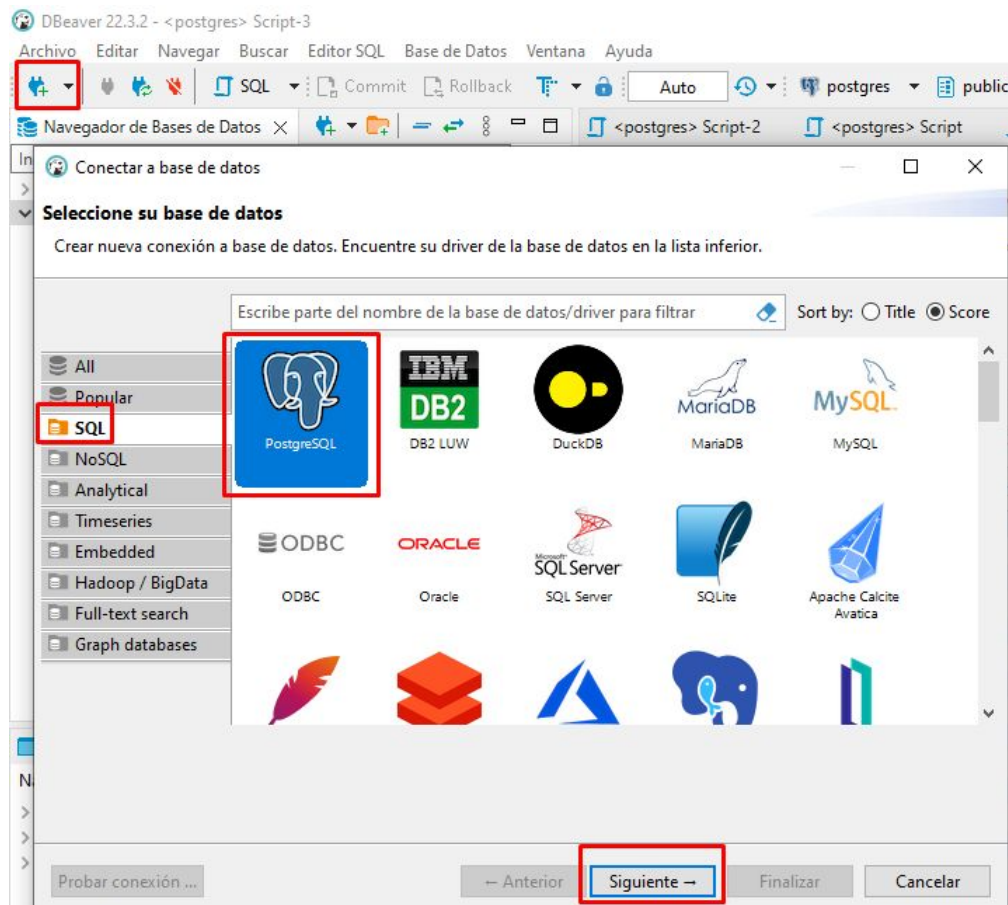
Enterprise Edition [features](#):

- Support of NoSQL databases:
 - [MongoDB](#)
 - [Cassandra](#)
 - [InfluxDB](#)
 - [Redis](#)
 - [Amazon DynamoDB](#)
 - [Amazon DocumentDB](#)
 - [Amazon Keyspaces](#)
 - [Amazon Timestream](#)
 - [Google Bigtable](#)
 - [Couchbase](#)
 - [CouchDB](#)
 - [Yugabyte](#)
- Advanced extensions for:
 - [EnterpriseDB](#)
 - [Informix](#)
 - [MySQL](#)
 - [Nettezza](#)
 - [Oracle](#)
 - [Redshift](#)
 - [Snowflake](#)
 - [SQL Server](#)
 - [SQLite Crypt](#)
 - [SQLite Remote](#)
 - [Teradata](#)
 - [Vertica](#)
- [AWS and GCP cloud explorer](#)



DBeaver

Configuración:





DBBeaver

Configuración:

Pass: edvai

Connect to a database

PostgreSQL

Connection Settings

PostgreSQL connection settings

Main

PostgreSQL

Driver properties

SSH

Proxy

SSL

Server

Connect by: ☒ Host ☐ URL

URL: jdbc:postgresql://localhost:5432/northwind

Host: localhost

Port: 5432

Database: northwind

Authentication

Authentication: Database Native

Username: postgres

Password:

☒ Save password locally

Advanced

Session role:

Local Client: PostgreSQL Binaries

You can use variables in connection parameters.

Connection details (name, type, ...)

Driver name: PostgreSQL

Driver Settings

Driver license

Test Connection ...

< Back

Next >

Finish

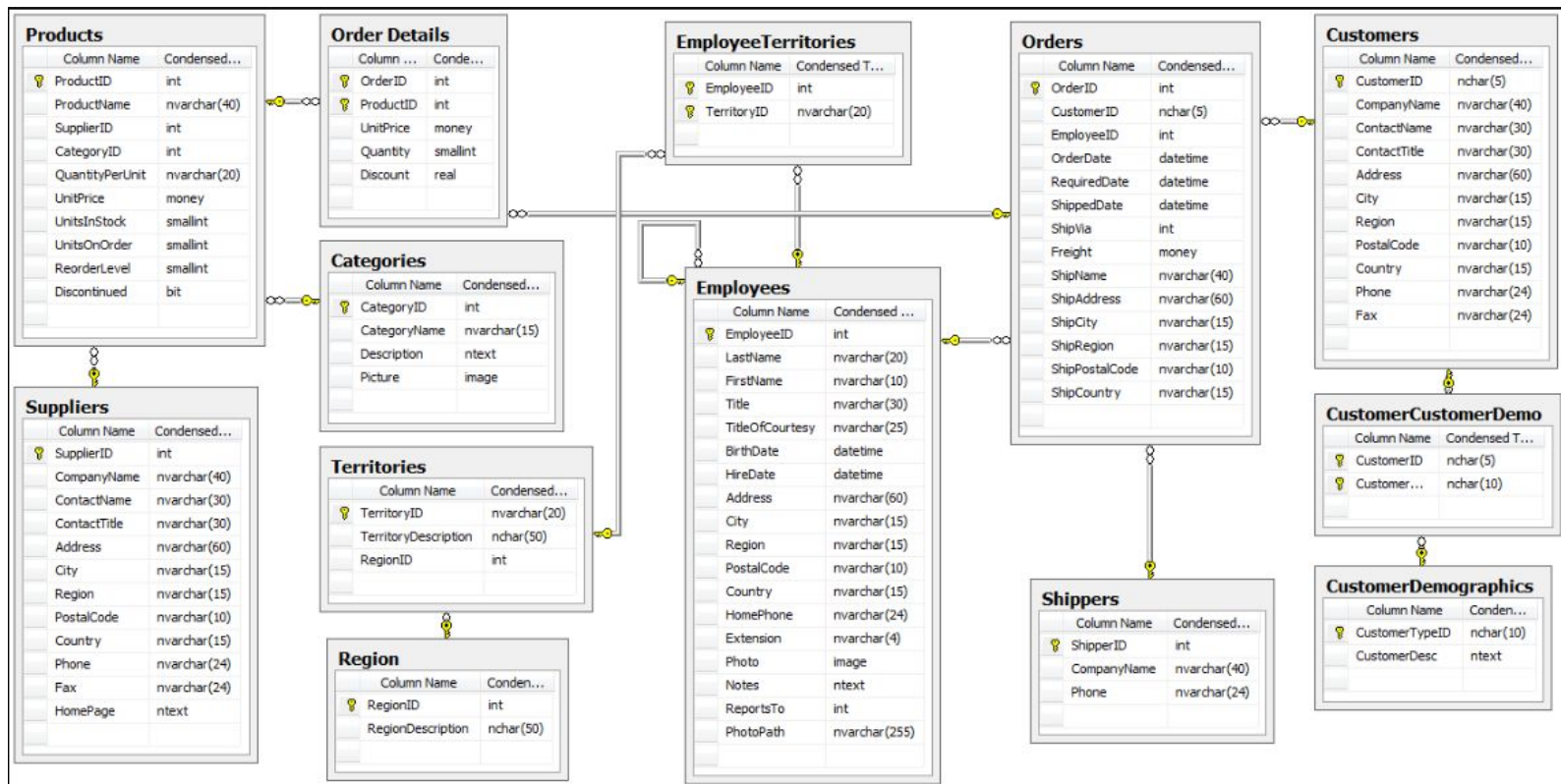
Cancel



Realizar Queries



Esquema de la DB





Ejecutar consultas

SQL

Navegador de Bases de Datos

postgres - localhost:5432

- Bases de Datos
 - northwind
 - Esquemas
 - public
 - Tablas
 - categories 32K
 - customer_customer_demo 8K
 - customer_demographics 16K
 - customers 56K
 - employee_territories 24K
 - employees 32K
 - order_details 192K
 - orders 176K
 - products 48K
 - region 24K
 - shippers 24K
 - suppliers 32K
 - territories 24K
 - us_states 24K
 - Vistas
 - Vistas Materializadas

Project - General

Name DataSource

- Bookmarks
- Diagrams
- Scripts

postgres Script

```
select
  o.order_id ,
  o.order_date ,
  od.unit_price ,
  od.quantity
from orders o inner join order_details od on o.order_id = od.order_id
```

orders(+) 1

select o.order_id, o.order_date, od.uni

Grilla

	order_id	order_date	unit_price	quantity
1	10.248	1996-07-04	14	12
2	10.248	1996-07-04	9,800000191	10
3	10.248	1996-07-04	34,799999237	5
4	10.249	1996-07-05	18,600000381	9
5	10.249	1996-07-05	42,400001526	40
6	10.250	1996-07-08	7,699999809	10
7	10.250	1996-07-08	42,400001526	35
8	10.250	1996-07-08	16,799999237	15
9	10.251	1996-07-08	16,799999237	6
10	10.251	1996-07-08	15,600000381	15
11	10.251	1996-07-08	16,799999237	20
12	10.252	1996-07-09	64,800003052	40

Record

Renovar Save Cancel Exportar datos



Preguntas?