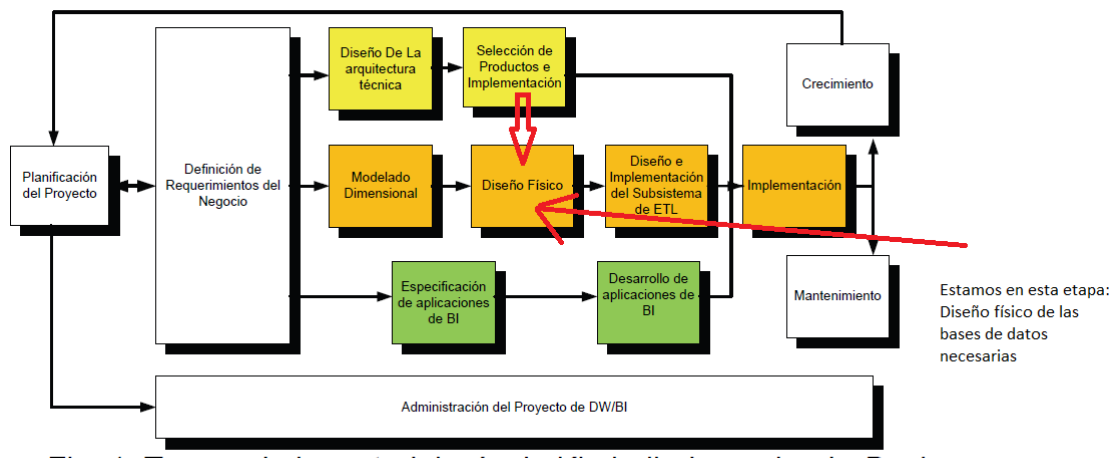


## PASO A PASO CREACIÓN DE TABLAS STAGIN Y DATAMART. Etapa de diseño físico de la metodología Kimball



*Aclaración: si bien en esta etapa estamos creando la base fuente, en una implementación real este primer paso no se realiza ya que la fuente es de un sistema transaccional ya existente. VER TEORÍA!!!*

El presente instructivo tiene como objetivo explicar el paso a paso para la instalación de la base de datos mysql y la creación de las distintas bases de datos a utilizar en la implementación de un sistema de business intelligence (desde ahora BI).

### 1. Instalación de base de datos Mysql y cliente Workbench.

Instalador con Wordbench incorporado: <https://downloads.mysql.com/archives/installer/mysql-installer-web-community-5.7.41.0>.

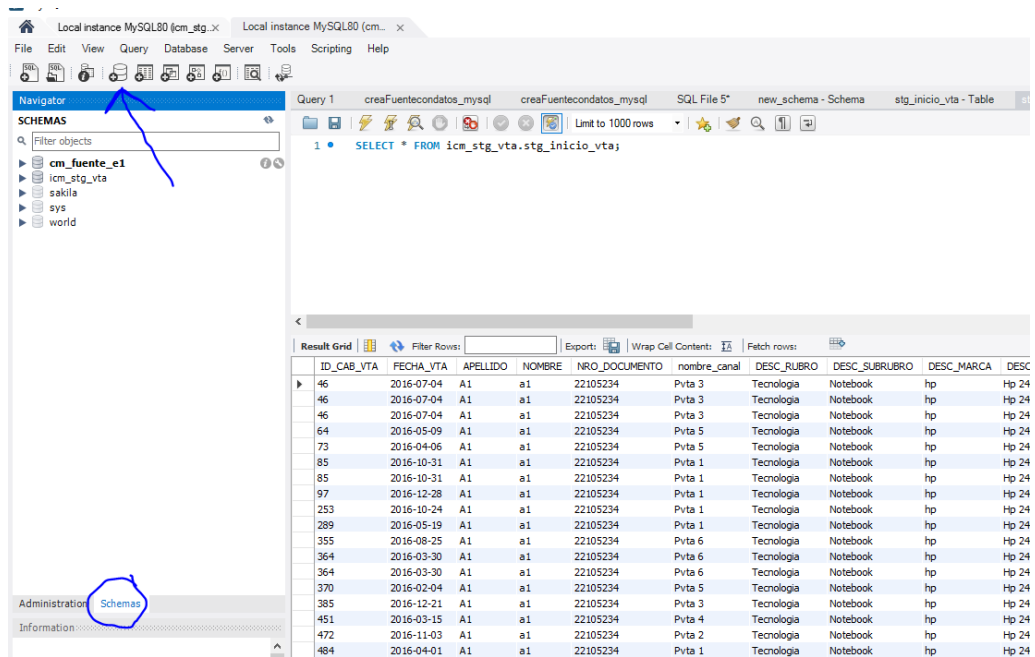
En la configuración del instalador seleccionar Developer.

Recordar la contraseña root que se elige en el proceso de instalación.

*\* Si hay dudas buscar tutoriales en you tube.*

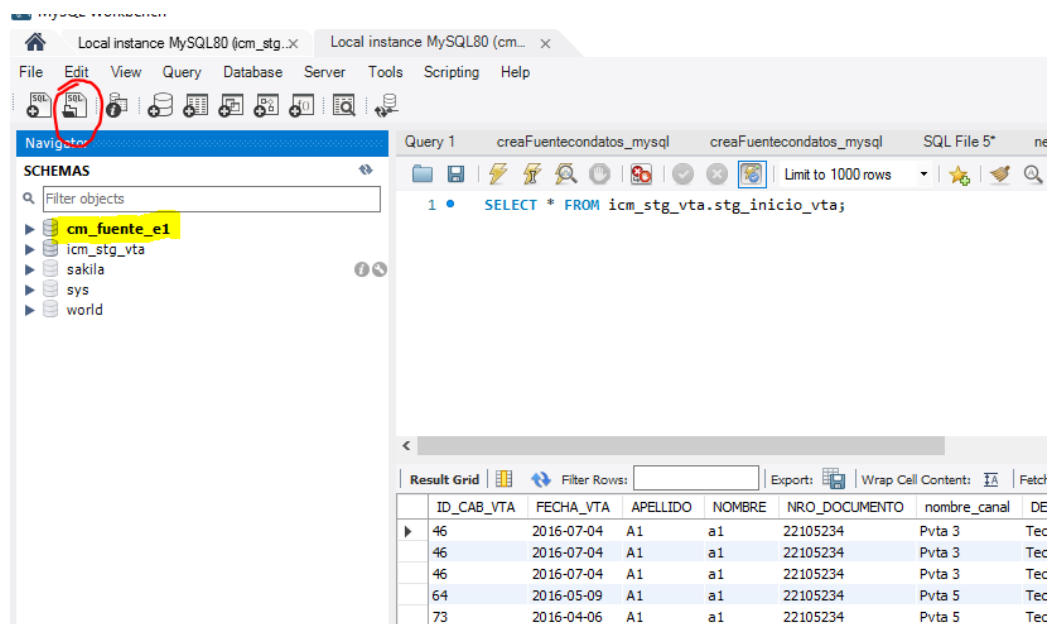
### 2. Creación de base de datos fuente (del sistema transaccional)

- Ejecutar Workbench
- Crear un nuevo esquema (base de datos) denominado *icm\_fuente\_e1* (esta es la base de datos del sistema transaccional que será la fuente de datos del datawarehouse).



- Una vez creado el esquema se debe ejecutar el script que se pueden descargar desde la plataforma o desde este enlace:  
[https://drive.google.com/file/d/1LJLPRUr20LOVqaFuFutOFELpiBum\\_xO/view?usp=sharing](https://drive.google.com/file/d/1LJLPRUr20LOVqaFuFutOFELpiBum_xO/view?usp=sharing)
- Para correr el script, primero se debe seleccionar el schema (el esquema debe quedar en negrita como se ve en la siguiente imagen resaltado) en el panel izquierdo de

Workbench, luego importar el archivo mediante el botón que se muestra en la siguiente pantalla:



\* Para seleccionar el esquema se debe hacer doble click sobre el mismo y debe quedar en negrita. Ver captura anterior.

Hecho esto en el panel de la derecha aparecerá el script el cual deberá ejecutarse para que se cree la estructura de tablas del sistema fuente (transaccional) junto con sus datos.

### 3. Creación de base de datos stagin con tablas intermedias y tablas stg con misma estructura que tendrá el datamart.

Para entender estos conceptos leer los apuntes subidos a la plataforma.

- Crear un nuevo esquema (base de datos) denominado *icm\_stg\_vta* (Esta es una base intermedia donde se cargarán los datos tomados de la fuente antes de pasar a la base definitiva). El script de creación de las tablas creará en esta misma base dos modelos iguales pero con distintos nombres: unos que empiezan con INT (intermedias) y otro que empieza con STG de Stagin)
- El script de creación de esta estructura se encuentra en <https://drive.google.com/file/d/1e3Aai12DDqpvq3OTZ0dX8l6sa8auLq5u/view?usp=sharing> o en la plataforma del instituto.
- Para ejecutar hacer el mismo procedimiento anterior seleccionando antes el esquema *icm\_stg\_vta* recién creado. Este script solo crea la estructura de las tablas, no carga los datos ya que la carga se realizará mediante los procesos ETL que desarrollaremos con la herramienta Pentaho tomando los datos de la base fuente (*icm\_fuente\_e1*) creada en el punto 2.

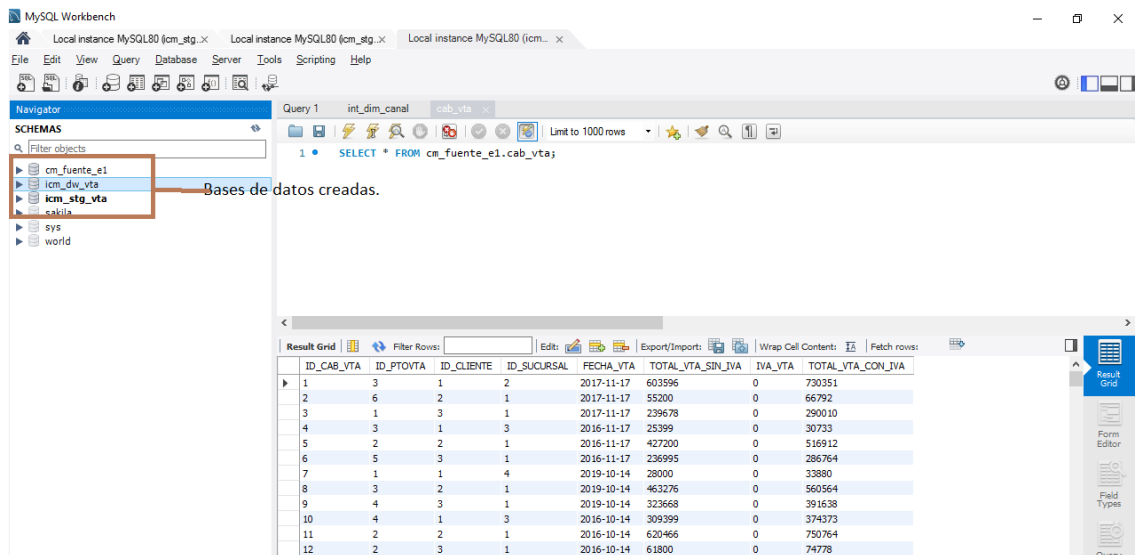
### 4. Creación de la base de datos del datamart ventas (parte del datawarehouse)

- Este procedimiento aún no hicimos en clases, pero deben crear este esquema ya que estamos en la etapa de Diseño Físico de la metodología de Kimball.

- Crear un nuevo esquema (base de datos) denominado *icm\_dw\_ventas* (Esta es una base definitiva del datamart Ventas).
- El script de creación de esta estructura se encuentra en:  
<https://drive.google.com/file/d/1A3aYB3dnO-P3knMPpOY7HIX87bl-UWzH/view?usp=sharing> o en la plataforma del instituto.
- Para ejecutar hacer el mismo procedimiento anterior seleccionando antes el esquema *icm\_dw\_ventas* recién creado. Este script solo crea la estructura de las tablas, no carga los datos ya que la carga se realizará mediante los procesos ETL que desarrollaremos con la herramienta Pentaho tomando los datos de la base de stagin creada en el punto 3.

## 5. Resumen de bases creadas

- *icm\_fuente\_e1* (es la fuente de donde se toman los datos para la creación del datawarehouse). En una implementación real esta base será preexistente ya que corresponde al sistema transaccional u operacional que utiliza la empresa en su operación diaria.
- *icm\_stg\_vta*. Es la base de stagin o de pruebas donde se las tablas del modelo están duplicadas, una con el sufijo INT (intermedia) y otras con el sufijo STG (stagin). Los procesos ETL cargan primero estas tablas para recién después cargarlos al datamart final que se encontrará en producción.
- *icm\_dw\_ventas*. Esta es la base definitiva que va a estar en producción y a la cual se conectará power bi para armar las visualizaciones.



MySQL Workbench interface showing the SCHEMAS list on the left and a query result grid on the right. The SCHEMAS list includes *cm\_fuente\_e1*, *icm\_dw\_vta*, *icm\_stg\_vta*, *sakila*, *sys*, and *world*. The query result grid shows data from the *cm\_fuente\_e1.cab\_vta* table.

Query 1: `SELECT * FROM cm_fuente_e1.cab_vta;`

	ID_CAB_VTA	ID_PTOVTA	ID_CLIENTE	ID_SUCURSAL	FECHA_VTA	TOTAL_VTA_SIN_IVA	IVA_VTA	TOTAL_VTA_CON_IVA
1	3	1	2		2017-11-17	603596	0	730351
2	6	2	1		2017-11-17	55200	0	66792
3	1	3	1		2017-11-17	239678	0	290010
4	3	1	3		2016-11-17	25399	0	30733
5	2	2	1		2016-11-17	427200	0	516912
6	5	3	1		2016-11-17	236995	0	286764
7	1	1	4		2019-10-14	28000	0	33880
8	3	2	1		2019-10-14	463276	0	560564
9	4	3	1		2019-10-14	323668	0	391638
10	4	1	3		2016-10-14	309399	0	374373
11	2	2	1		2016-10-14	620466	0	750764
12	2	3	1		2016-10-14	61800	0	74778