



Bases de Datos Masivas (11088)
Departamento de Ciencias Básicas

GUÍA DE INSTALACIÓN: SUITE PENTAHO

(SO Linux y Windows)

Pentaho BI Suite es un conjunto de aplicaciones libres para generar inteligencia empresarial (Business Intelligence) que emplearemos durante la materia. Utilizaremos la versión **community**, bajo licencia GPL.

INSTALACIÓN PASO A PASO:

1. La Suite Pentaho corre sobre la plataforma Java por lo tanto nos debemos asegurar de tener instalado Java Development Kit (JDK) y Java Runtime Environment (JRE). JDK incluye JRE y herramientas adicionales para el desarrollo de applets y aplicaciones. En el caso de no tener instalado el JDK se recomienda verificar como hacerlo en la sección “*Instalación y Configuración de Java Development Kit (JDK)*” que se encuentra al final de la guía.
2. Una vez que tenemos instalado JAVA y configuradas las variables de entorno, debemos ingresar a la sección de downloads del sitio web de Pentaho community a efectos de descargar el software: <https://community.hitachivantara.com/>. Pentaho es multiplataforma, con lo cual no tendremos que distinguir por Sistema Operativo al momento de realizar la descarga de los instaladores.
3. Una vez en el sitio web, descargaremos las siguientes herramientas:
 - a. Business Analytics Platform,
 - b. Data Integration: también conocida como Kettle,
 - c. Schema Workbench.: también conocida como Mondrian,
4. A continuación, instalaremos **Business Analytics Platform**:
 - a. Descomprimos el archivo *pentaho-server-ce-X.X.X.X-XXX*.



Bases de Datos Masivas (11088)
Departamento de Ciencias Básicas

- b. A continuación, aparecerá la carpeta *pentaho-server*.
- c. En este momento podríamos iniciar Pentaho ejecutando el archivo *start-pentaho.sh* (UNIX) o *start-pentaho.bat* (Windows) y accederlo a través de un navegador mediante la URL: localhost:8080/pentaho/. En el caso de ambientes UNIX debemos cautelar que el archivo tenga permisos de ejecución para el usuario.
- d. No obstante, se podría modificar el puerto mediante 8080 por otro, por ejemplo el 9999. Esto se realiza editando el archivo *server.xml*.
- e. La posición relativa del archivo *server.xml* se encuentra en */pentaho-server/tomcat/conf*. Debemos abrir el archivo, en caso de UNIX con un usuario con permisos de escritura, y modificamos el puerto en la línea:

```
<Connector URIEncoding="UTF-8" port="9999" protocol="HTTP/1.1" connectionTimeout="20000" redirectPort="8443" />
```
- f. Solo falta instalar uno de los Exploradores OLAP que provee Pentaho, en nuestro caso utilizaremos Saiku:
 - Para ello, descargamos el plugin del Marketplace de Pentaho en <https://www.pentaho.com/marketplace/>.
 - Luego, descomprimos la carpeta dentro de la instalación de Pentaho en */pentaho-server/pentaho-solutions/system/*.
 - Por último, tendremos que gestionar una licencia educativa en la web del desarrollador de Saiku, en <http://licensing.meteorite.bi/>, la descargamos, renombramos como *license.lic* y por último la dejamos en el raíz de Saiku en */pentaho-server/pentaho-solutions/system/Saiku*.

5. Luego, instalamos **Mondrian Schema Workbench**:

- a. Descomprimos el archivo *psw-ce-X.X.X.X-XXX*.



Bases de Datos Masivas (11088)
Departamento de Ciencias Básicas

- b. A continuación, aparecerá la carpeta schema-workbench.
- c. Para ejecutar el software, ejecutamos el archivo workbench.sh (UNIX) o workbench.bat (Windows) que se encuentra en la carpeta raíz. En el caso de ambientes UNIX debemos cautelar que el archivo tenga permisos de ejecución para el usuario.
- d. En la cursada, vamos a trabajar con Mondrian interactuando con el SGBD PostgreSQL, para ello debemos descargar el driver del conector JDBC de la web de PostgreSQL, en <https://jdbc.postgresql.org/>, y guardarlo en la carpeta /schema-workbench/drivers/.

6. Ahora, instalamos **Data Integration**:

- a. Descomprimos el archivo pdi-ce-X.X.X.X-XXX.
- b. A continuación, aparecerá la carpeta data-integration.
- c. Para iniciar PDI o Kettle, ejecutamos el archivo Spoon.sh (UNIX) o Spoon.bat (Windows) que se encuentra en la carpeta raíz. En el caso de ambientes UNIX debemos cautelar que el archivo tenga permisos de ejecución para el usuario.



Bases de Datos Masivas (11088)
Departamento de Ciencias Básicas

INSTALACIÓN Y CONFIGURACIÓN DE JAVA DEVELOPMENT KIT (JDK)

1. La instalación de JDK se realiza de la siguiente manera, de acuerdo al entorno que utilice:

a. En UNIX podemos instalarlo a través de los siguientes comandos:

- Agregamos el repositorio para descargar Java8 de Oracle:

```
sudo add-apt-repository ppa:webupd8team/java
```

```
sudo apt-get update
```

- Instalamos desde los repositorios agregados Java:

```
sudo apt-get install oracle-java8-installer
```

- Definimos que por default se utilice Java 8:

```
sudo apt-get install oracle-java8-set-default
```

```
sudo update-alternatives --config java
```

b. En Windows, descargamos el instalador desde la web de Oracle en

<http://www.oracle.com/technetwork/java/index.html>.

2. Una vez descargados e instalados, debemos configurar la variable de entorno JAVA_HOME:

a. En ambientes UNIX:

- Definimos la variable de entorno \$JAVA_HOME:

```
sudo gedit /etc/environment
```

```
JAVA_HOME="/usr/lib/jvm/java-8-oracle"
```

(Verificar el path donde se instaló en esa máquina)

- Volvemos a cargar el archivo environment para que tome los cambios:

```
source /etc/environment
```

b. En ambientes Windows:



Bases de Datos Masivas (11088)
Departamento de Ciencias Básicas

1. Hacemos click derecho sobre “Mi PC” ó “Este equipo”, de acuerdo a la versión del SO, y seleccionamos la opción “Propiedades”. Luego accedemos a “Configuración avanzada / Cambiar configuración -> Opciones avanzadas -> Variables de entorno -> Nueva (Variables del sistema)”.
2. Configuramos la variable JAVA_HOME. Para ello, escribimos en el cuadro de texto lo siguiente:
 - a. **Nombre de variable:** JAVA_HOME,
 - b. **Valor de variable:** ruta en que se haya instalado Java. Por ejemplo “C:\Program Files\Java\jdk1.7.0_51”.
(Verificar el path donde se instaló en esa máquina)