

Day with Development

Alexey Filanovskiy - DWGL
Marty Gubar - Big Data PM

Setup Guide

Big Made Simple

Setup: Basics

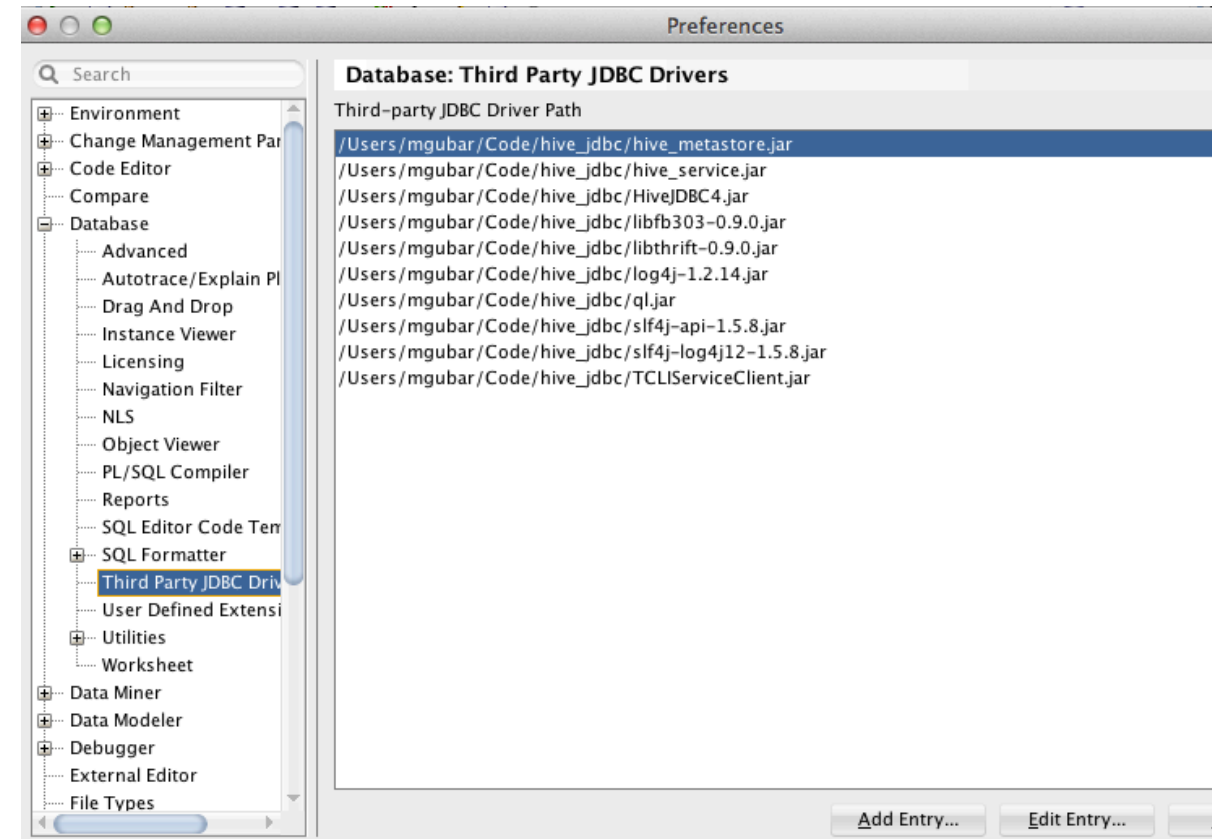
1. Wifi: clear-guest
id/password - guest / <password>
2. Copy following folders from thumb-drive to your computer:
 - setup-files
 - Workshop-HOL
 - Appropriate SQL Developer Version
 - Windows-SQLDev
 - OSX-SQLDev

Setup: Install SQL Developer

- Windows
 - Unzip sqldeveloper-4.1.5.21.78-x64.zip
- OSX
 - Ensure you have JDK 8 (included on Thumb-drive)
 - Unzip sqldeveloper-4.1.5.21.78-no-jre.zip

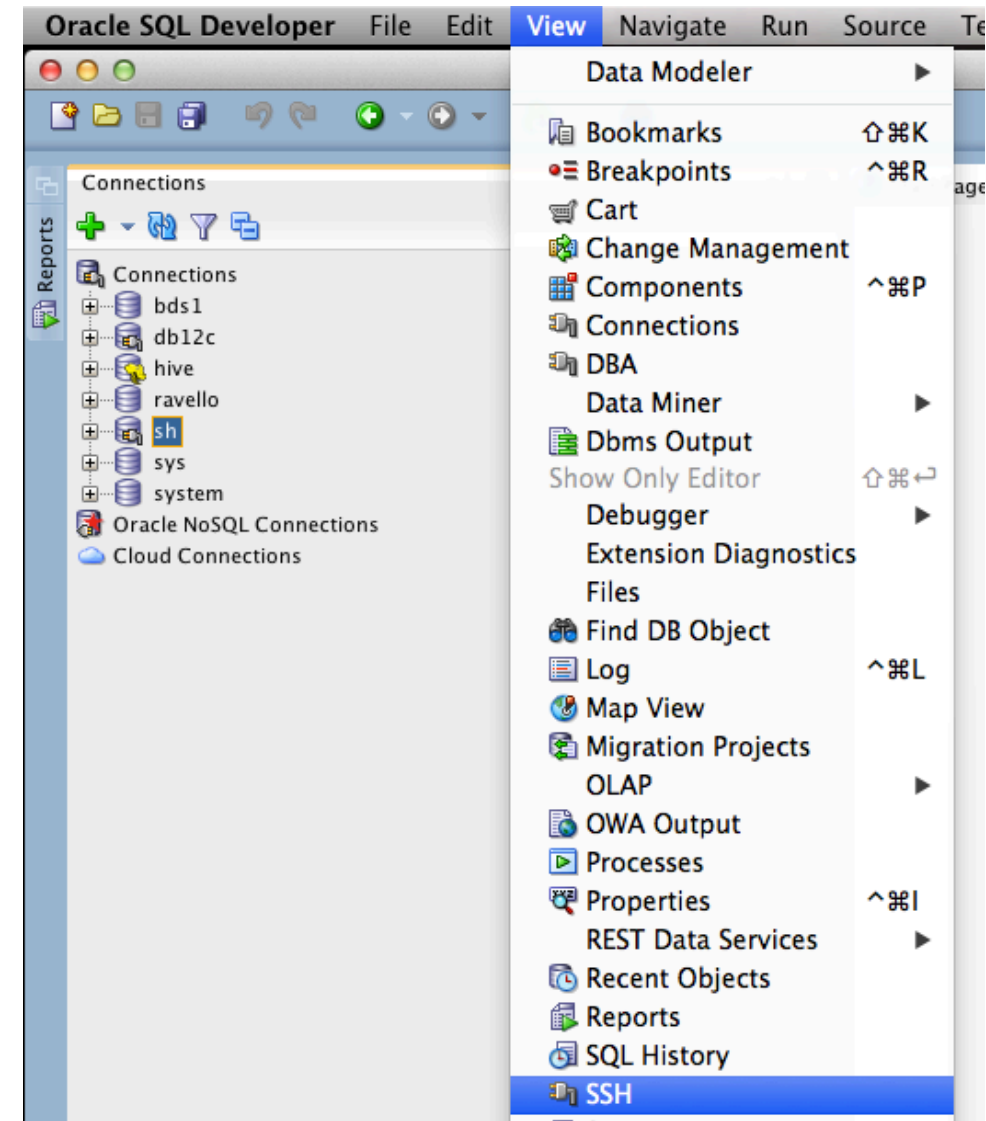
Setup SQL Developer: Configure Hive JDBC Drivers

- Launch SQLDeveloper
- Go to SQL Developer **Preferences** menu
- In **Database->Third Party JDBC ...**
 - **Add Entry**
 - Add each jar found in the **[OSX | Windows]-SQLDev/hive_jdbc** folder (copied from thumb drive)
- Restart SQL Developer



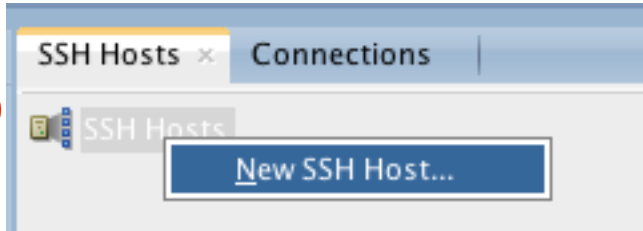
Enable Tunneling

- We'll use an SSH tunnel to allow SQL Developer to connect to Oracle Database and Hive
 - All connections will use “localhost” and the requests will be forwarded
- Open SSH Connection window
- View -> SSH

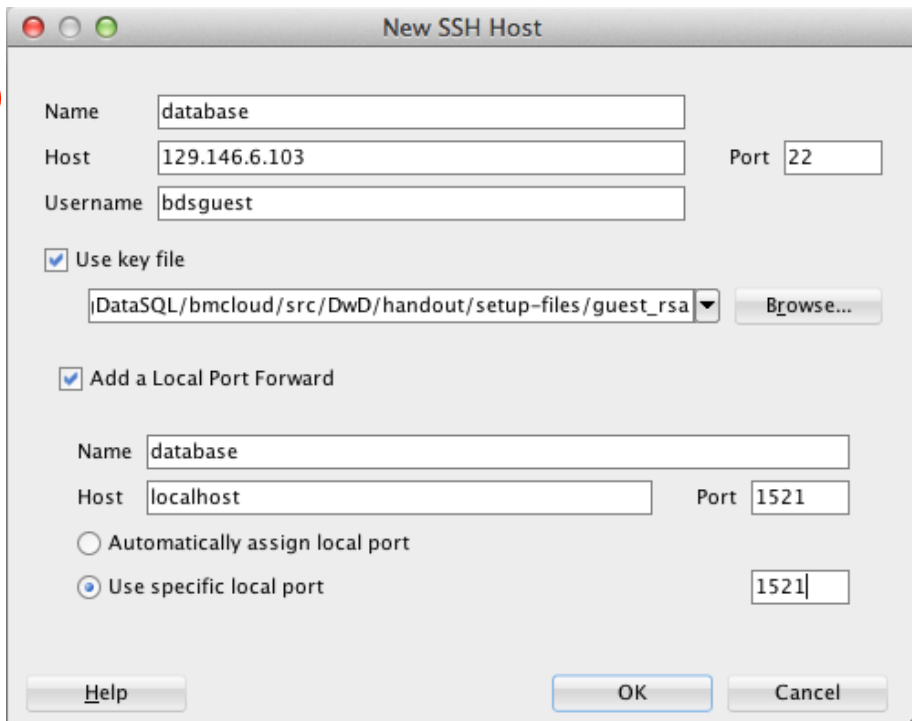


Enable Tunneling: Database

1



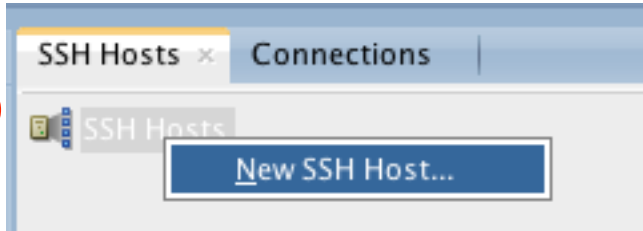
2



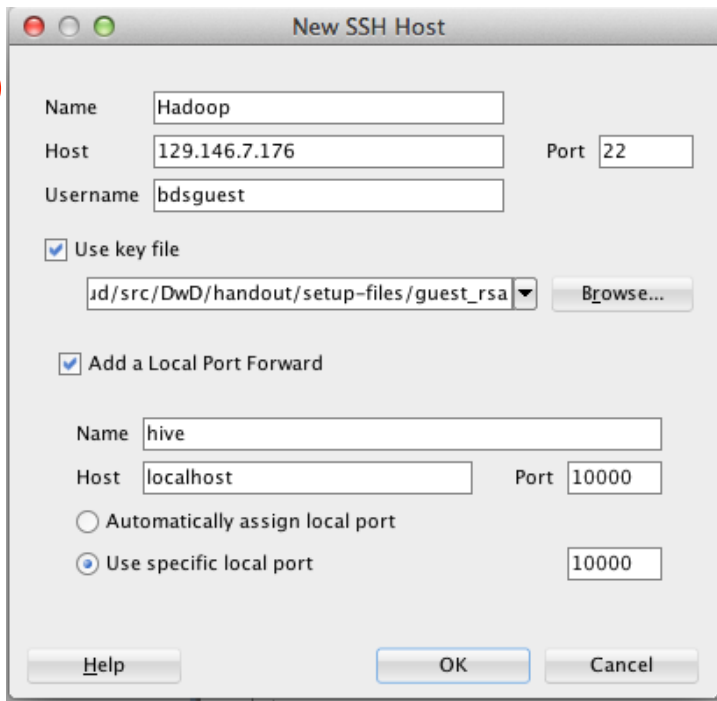
- Right-click SSH Hosts to add new SSH connection and tunnel
- Parameters:
 - SSH
 - Name: Database
 - Host: 129.146.6.103
 - Port: 22
 - Username bdsguest
 - Key file: <topdir>/setup-files/guest_rsa
 - Port Forward:
 - Name: database
 - Host: localhost
 - Port: 1521
 - Use local port: 1521

Enable Tunneling: Hadoop

1



2



- Right-click SSH Hosts to add second SSH connection and tunnel
- Parameters:
 - SSH
 - Name: Hadoop
 - Host: 129.146.7.176
 - Port: 22
 - Username bdsguest
 - Key file: <topdir>/setup-files/guest_rsa
 - Port Forward:
 - Name: hive
 - Host: localhost
 - Port: 10000
 - Use local port: 10000

Properties on Private Key

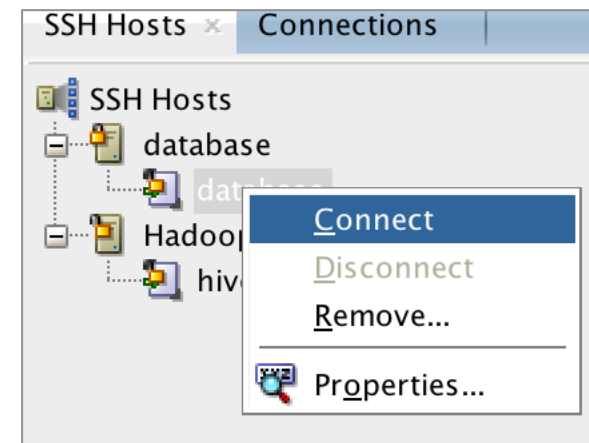
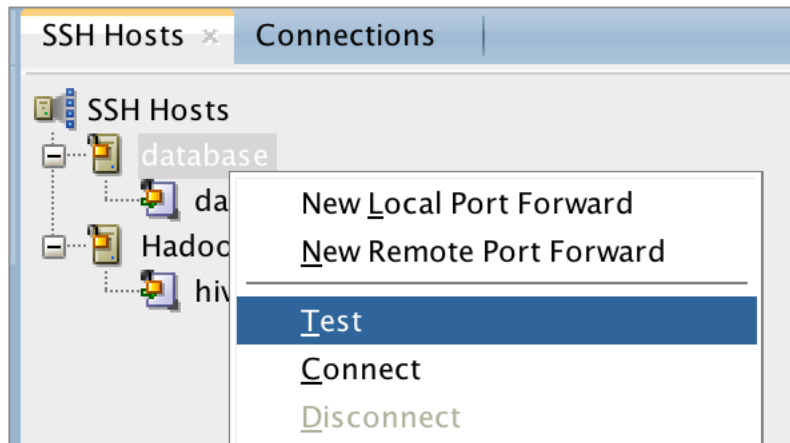
- Ensure that private key **guest_rsa** is owned by you and has the correct ACLs

Example:

```
mgubar-mac:setup-files mgubar$ ls -l guest_rsa
-rw----- 1 mgubar staff 1675 Mar 1 16:18 guest_rsa
```

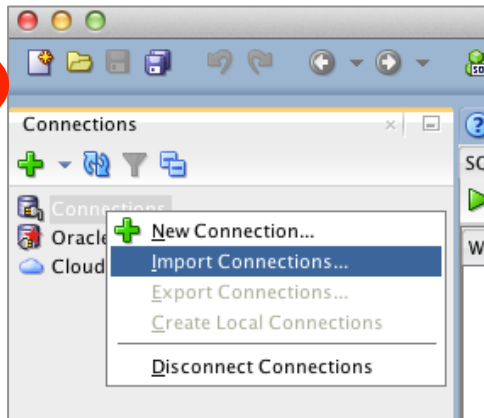

Test the SSH Connection & then Connect!

- Test that the connection works for both the Database and Hadoop
- When successful, connect using the Local Port Forward for both the Database and Hive

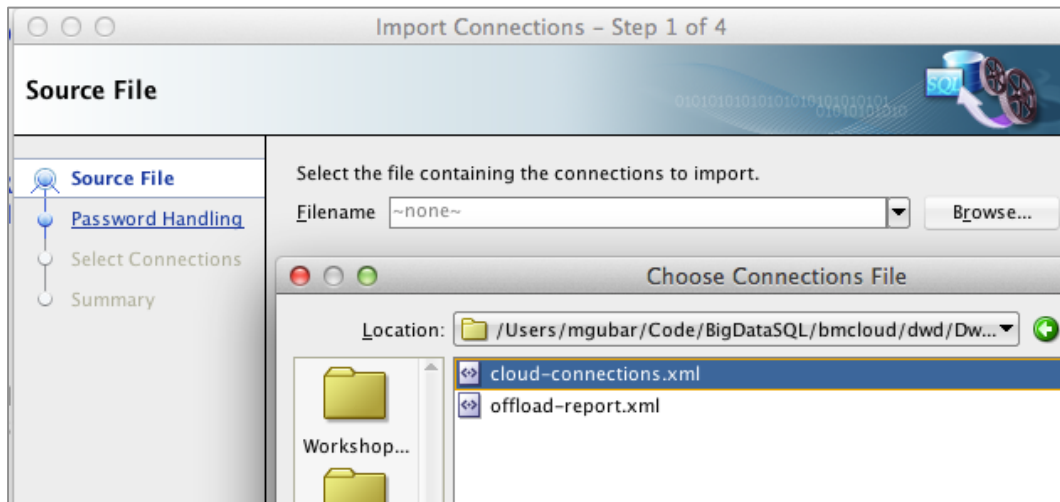


Setup SQL Developer: Add Database Connections

1

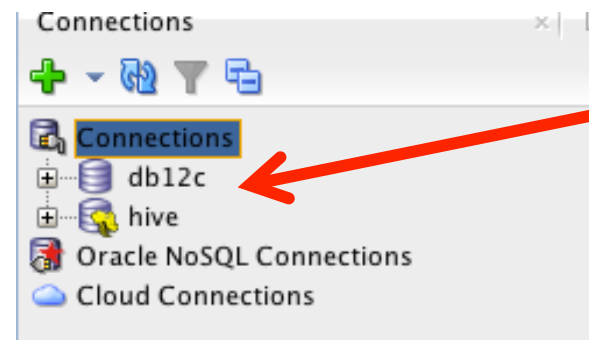


2

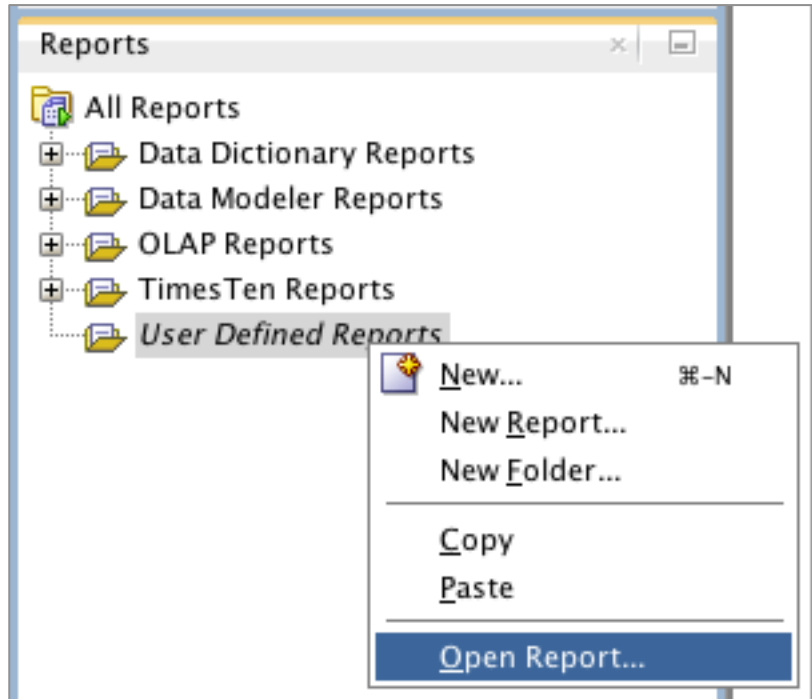


- Right-click **Connections**. Select **Import Connections**
- Select setup-files/cloud-connections.xml. Next.
- Remove All Passwords. Next.
- Select All Connections. Finish

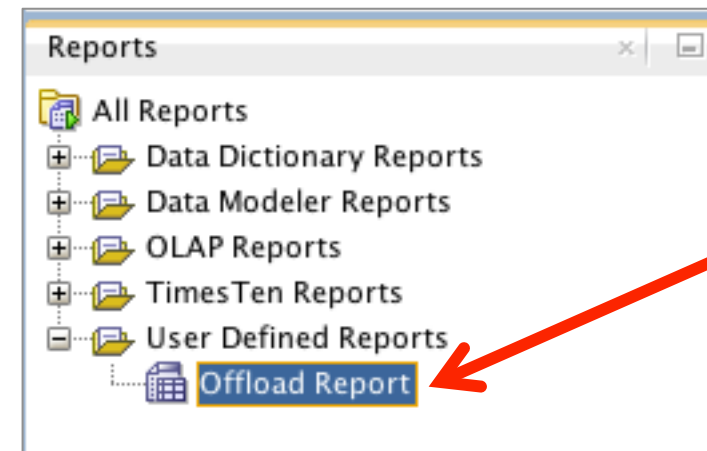
Result:



Setup SQL Developer: Add User Defined Report

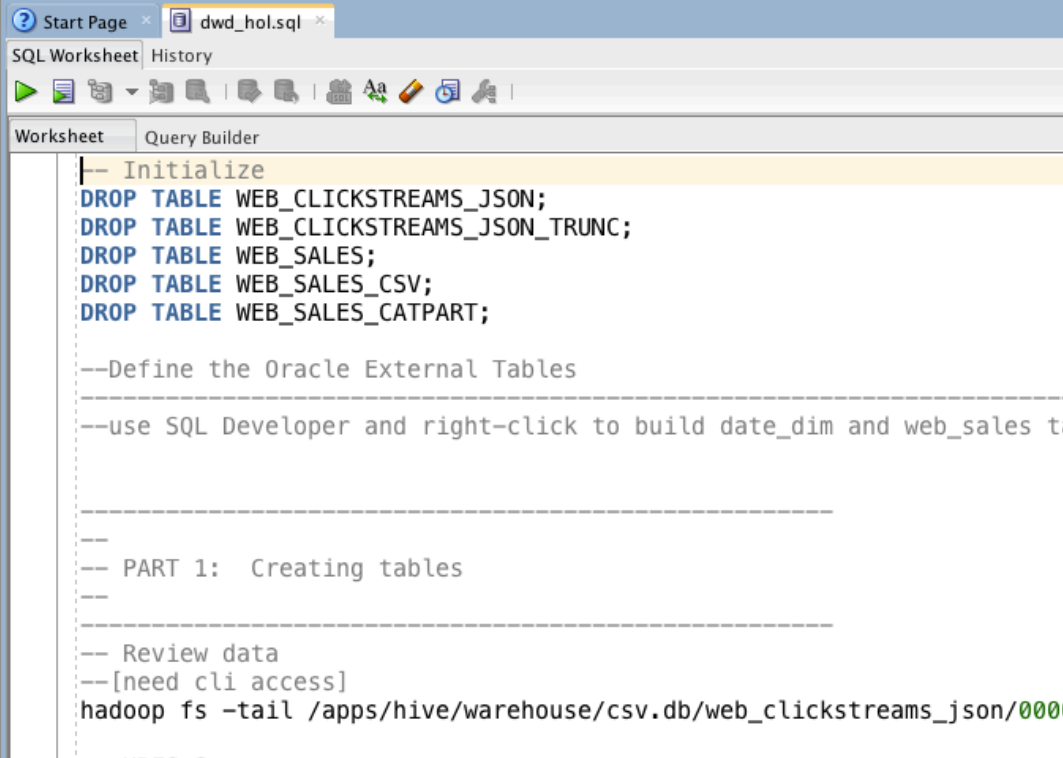


- Right click on **User Defined Reports** and select **Open Report...**
- **Select setup-files/offload-report.xml**



SQL Developer: Open HOL Script

- File->Open
 - Workshop-HOL/dwd_hol.sql



```
-- Initialize
DROP TABLE WEB_CLICKSTREAMS_JSON;
DROP TABLE WEB_CLICKSTREAMS_JSON_TRUNC;
DROP TABLE WEB_SALES;
DROP TABLE WEB_SALES_CSV;
DROP TABLE WEB_SALES_CATPART;

--Define the Oracle External Tables

--use SQL Developer and right-click to build date_dim and web_sales t

-----

-- PART 1: Creating tables
--

-----

-- Review data
--[need cli access]
hadoop fs -tail /apps/hive/warehouse/csv.db/web_clickstreams_json/000

HDFS Source
```

Connect!

- You will be assigned a number - your user id will be **bds[i]**
- All passwords are **bigdata**

The screenshot shows the 'New / Select Database Connection' dialog box. On the left, there is a table with two columns: 'Connection Name' and 'Connection Details'. The table contains two rows: 'db12c' with details 'bds[n]@//129.146.0...' and 'hive' with details 'opc@//129.146.0.9...'. The 'db12c' row is selected. On the right, the details for the selected connection are shown. The 'Connection Name' is 'db12c'. The 'Username' is 'bds1'. The 'Password' is masked with dots. There are checkboxes for 'Save Password' and 'Connection Color'. Below these are tabs for 'Oracle' and 'Hive'. The 'Oracle' tab is active, showing 'Connection Type' as 'Basic' and 'Role' as 'default'. The 'Hostname' is '129.146.0.64', 'Port' is '1521', and 'SID' is 'orcl'. There are radio buttons for 'SID' and 'Service name'. At the bottom, there are checkboxes for 'OS Authentication' and 'Kerberos Authentication', and an 'Advanced...' button. At the very bottom of the dialog are buttons for 'Help', 'Save', 'Clear', 'Test', 'Connect', and 'Cancel'.

Connection Name	Connection Details
db12c	bds[n]@//129.146.0...
hive	opc@//129.146.0.9...

Connection Name: db12c
Username: bds1
Password:
☐ Save Password ☒ Connection Color
Oracle Hive
Connection Type: Basic Role: default
Hostname: 129.146.0.64
Port: 1521
☒ SID: orcl
☐ Service name
☐ OS Authentication ☐ Kerberos Authentication Advanced...
Status :
Help Save Clear Test Connect Cancel