

Topic Title **Updating the Data Mart: Preparing and Running Jobstreams**

Introduction

This guide is intended to help you learn and follow the process of updating the Institutional Research data mart (DMIR). To “update” the data mart means to **insert new data into the database**. This is accomplished using specialized software (Cognos DecisionStream) to extract data from a source database and load it into a target database. The data may be transformed along the way; e.g. selected fields from multiple tables in the source database may be combined in a single table in the target database. This process is referred to as ETL: Extract, Transform, and Load. The handout you received before this training, “A Brief Primer on Data Marts,” contains more information about data marts and their place within the Santa Clara University reporting environment.

DMIR is updated at regular intervals. Data related to degrees conferred are extracted **annually**, while data related to admissions and enrollment are extracted **twice each term**, once on the census enrollment date and again after the term has ended to capture final enrollment data. Specifying the data to be extracted involves the use of parameters. Parameters are special variables that specify the **scope** of the data to be extracted from the source database. For the IR data mart, the parameters indicate the **population** (i.e. School of Law or all other careers) and **point in time** (i.e. the term and census/final enrollment date). This guide will help you through the process of determining what data to extract and how to do it.

The process of updating the data mart consists of the following general tasks, each of which is broken down into a series of steps and illustrated with screenshots to assist you in learning and following the process:

1. Prepare to update the data mart
2. Connect to the remote computer
3. Review and set the parameters
4. Run the jobstreams and review the results
5. Disconnect from the remote computer

Acronyms in this Guide

DMIR – Institutional Research Data Mart

ETL – Extract, Transform, Load: the process of updating a data mart.

SQL – Structured Query Language: standard language used to perform actions in a database.

Topic Title **Updating the Data Mart: Preparing and Running Jobstreams**

Process

Prepare to update the data mart

1. Determine the type of data to extract: **annual** or **term**. If you are collecting new data for degrees conferred, you are extracting **annual** data. Otherwise you are extracting **term** data.
2. If you are extracting **annual** data, determine the academic year for which you want to extract data. These data are extracted in the Fall term for reporting to various government agencies, college guides, etc. The terms range from the Summer term through the Spring term of the previous academic year. For example, if it is Fall 2005, you are extracting data for the 2004-2005 academic year, from Summer 2004 through Spring 2005. The PeopleSoft codes for terms may be queried from the PS_ACAD_CALTRM_TBL table.
3. If you are extracting **term** data ...
 - a. Determine the term code. The PeopleSoft codes for terms may be queried from the PS_ACAD_CALTRM_TBL table.
 - b. Determine whether you are extracting data for the School of Law or for all other careers. The School of Law is on a semester calendar rather than a quarter calendar, so its term data are extracted at a different time.
 - c. Determine whether you are extracting data for census or final enrollment.
 - d. Determine the effective date for extracting data.
 - If you are collecting data for census enrollment, use the official census date for the term/career. The Student Records Office should be able to provide this date. You may also query the PeopleSoft PS_ACAD_CALTRM_TBL table for the value in the CENSUS_DT field for the appropriate term/career.
 - If you are collecting data for final enrollment, use the current date.

Once you've determined the appropriate values to use in extracting the data, you're ready to connect to the data mart to review and set the parameters.

Topic Title **Updating the Data Mart: Preparing and Running Jobstreams**

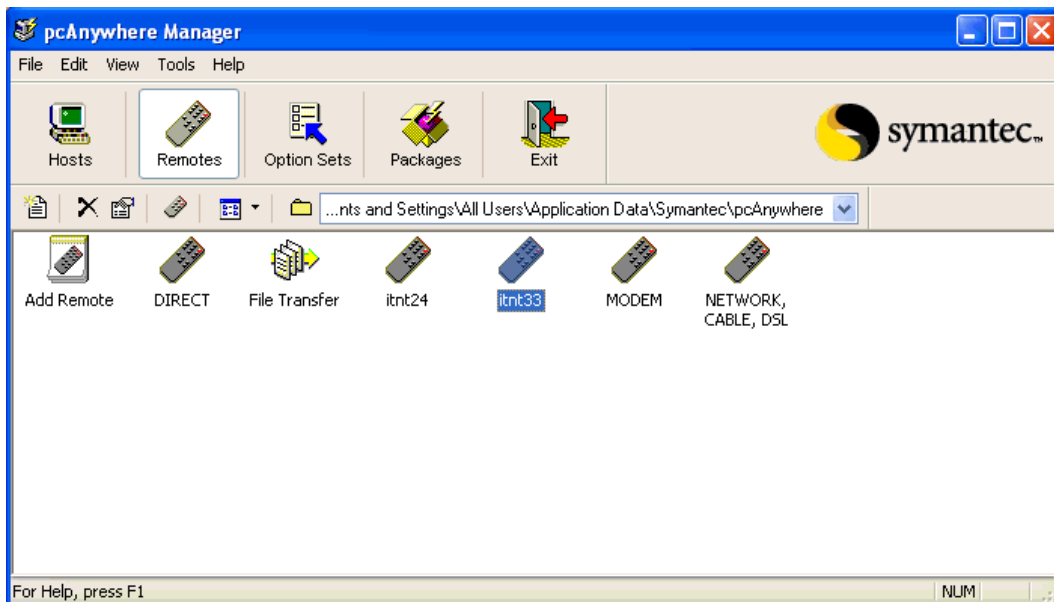
Connect to the remote computer that hosts the data mart

The software used to update the data mart (Cognos DecisionStream) is on a computer behind a firewall in the Information Technology department. Security requires connecting from a computer with an allowed IP address (Rod's computer in Institutional Research). Once you're sitting at that computer, follow the instructions below to connect to the remote computer.

1. Double-click the "pcAnywhere" icon on the desktop to open the software for connecting to a remote computer.



The name of the remote computer that hosts the data mart is "itnt33." In the pcAnywhere window, double-click the "itnt33" icon.



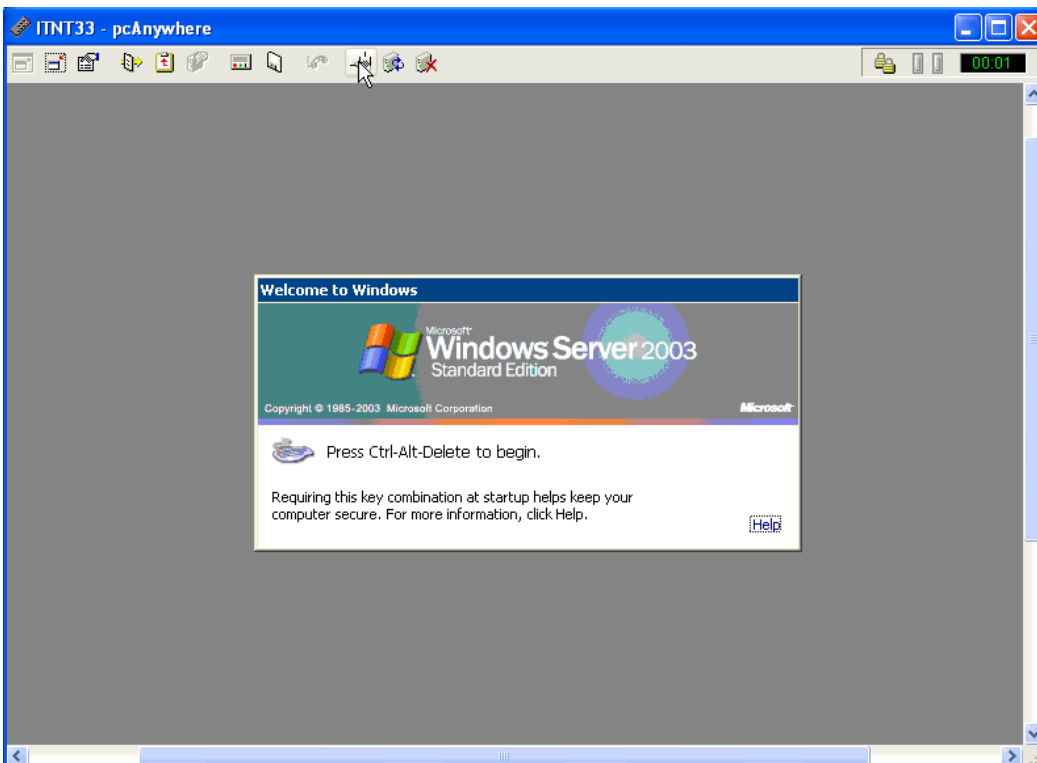
This should establish a connection to the remote computer. If access is denied because the computer is locked, send e-mail to Mai Anh [mky@scu.edu] and Ed [emulliken@scu.edu]. If they are not using the computer, they can unlock it.

Topic Title **Updating the Data Mart: Preparing and Running Jobstreams**

2. Log on to the remote computer. The username and password are in the office password folder.

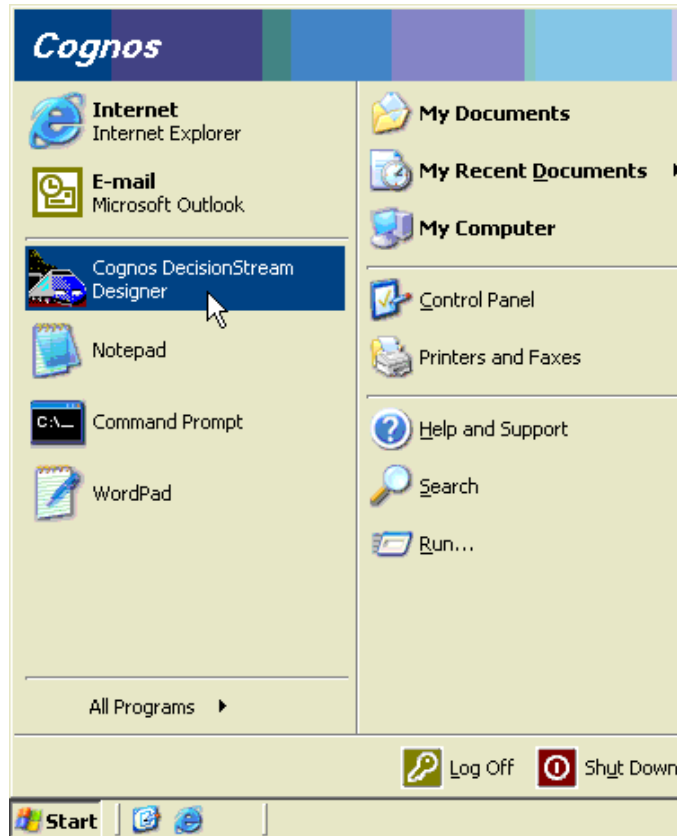


3. The Welcome screen will prompt you to press Ctrl-Alt-Delete to begin. **Don't do it!** That only pertains to someone who is sitting in front of the remote computer. If you press Ctrl-Alt-Delete you will open the Windows Task Manager window on the local computer. Instead click the Ctrl-Alt-Delete icon at the top of the screen as shown below.



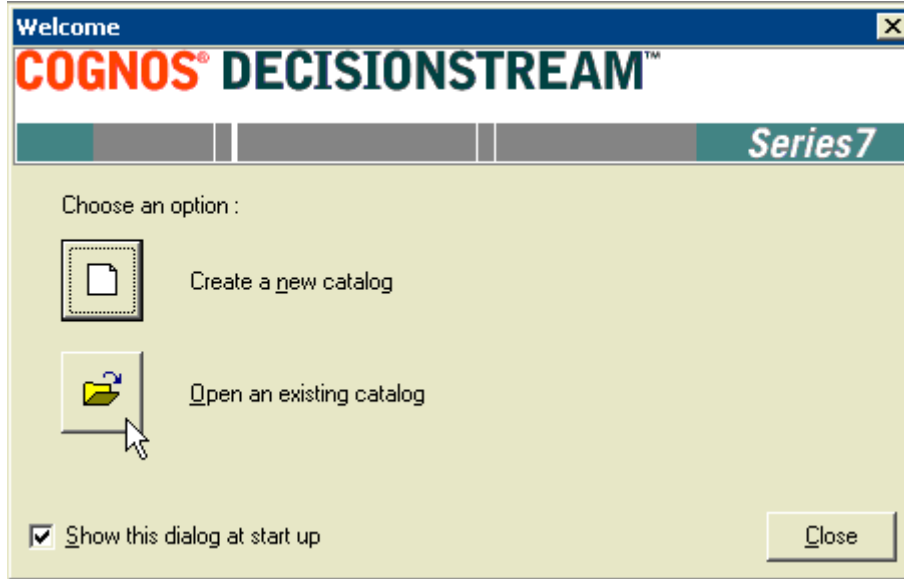
Topic Title **Updating the Data Mart: Preparing and Running Jobstreams**

4. Open the Cognos DecisionStream software.
 - a. On the remote computer, click the Windows “Start” button.
 - b. Click “Cognos DecisionStream Designer.”

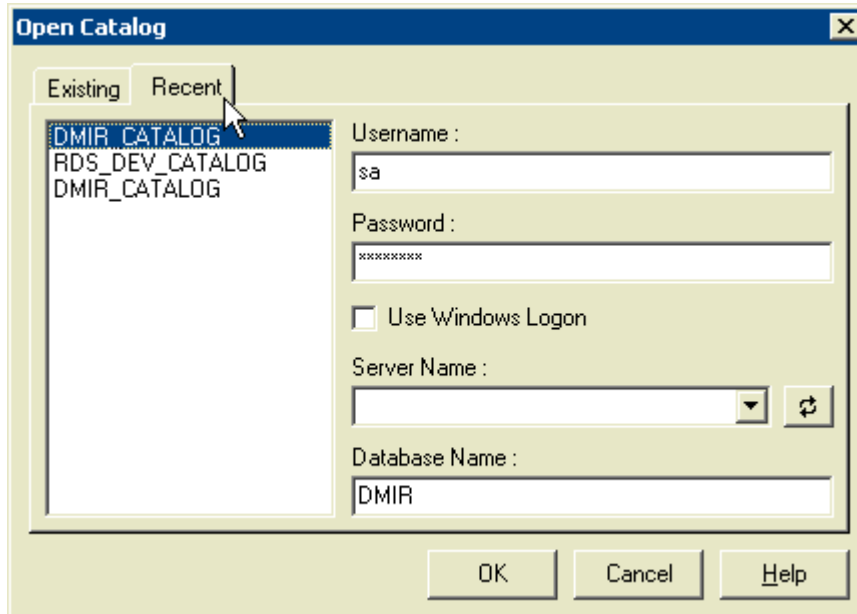


Topic Title **Updating the Data Mart: Preparing and Running Jobstreams**

- c. When DecisionStream starts, it will prompt you to create a new catalog or open an existing catalog. Click on “Open an existing catalog”.



- d. If the value in the “Database Name” field is “DMIR”, click the “OK” button.
- e. Otherwise click the “Recent” tab, select the first instance of “DMIR_CATALOG” in the list, and click the “OK” button.



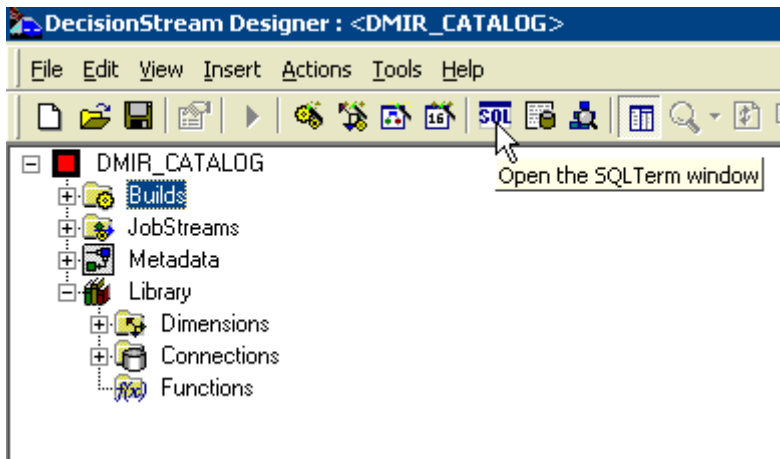
Topic Title **Updating the Data Mart: Preparing and Running Jobstreams**

Review and set parameters for extracting new data

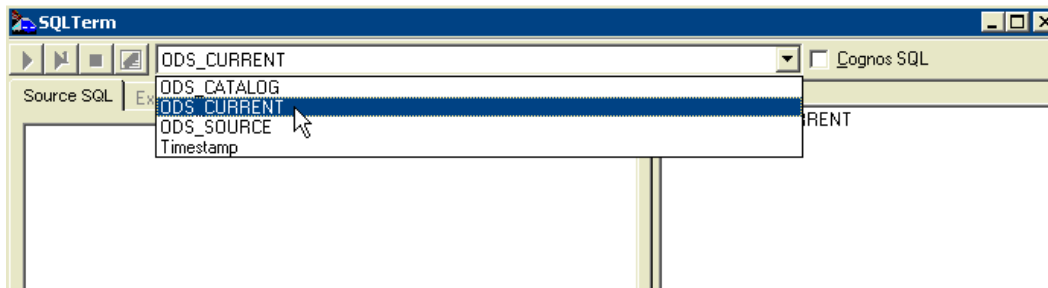
Review parameters

Now that you're in DecisionStream and have opened the DMIR catalog, you should review the current parameters. This isn't strictly necessary, but it may save you time. For example, if you're extracting final enrollment data, the parameters from the previous census enrollment extract may still be set. In that case you may need to change only the "census_run" and "census_date" parameters.

1. In DecisionStream, click on the "SQL Term" icon to open the SQL Term window.



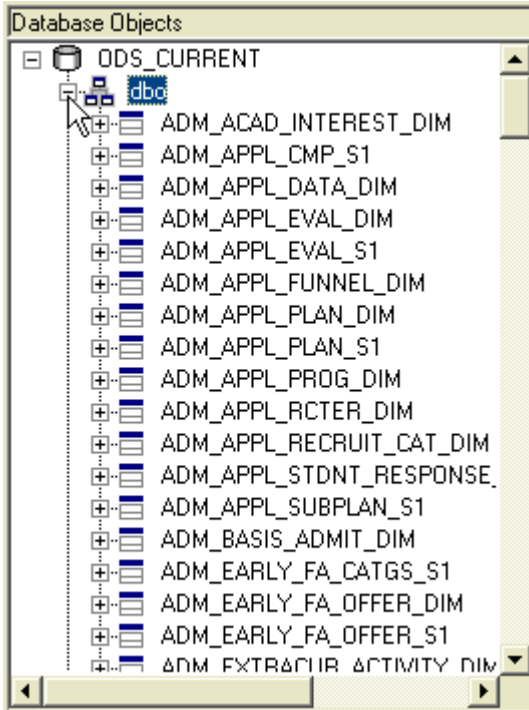
2. At the top of the SQL Term window, use the drop-down menu to select "ODS_CURRENT".



3. In the Database Objects frame, click the "+" (plus) button next to "ODS_CURRENT".

Topic Title **Updating the Data Mart: Preparing and Running Jobstreams**

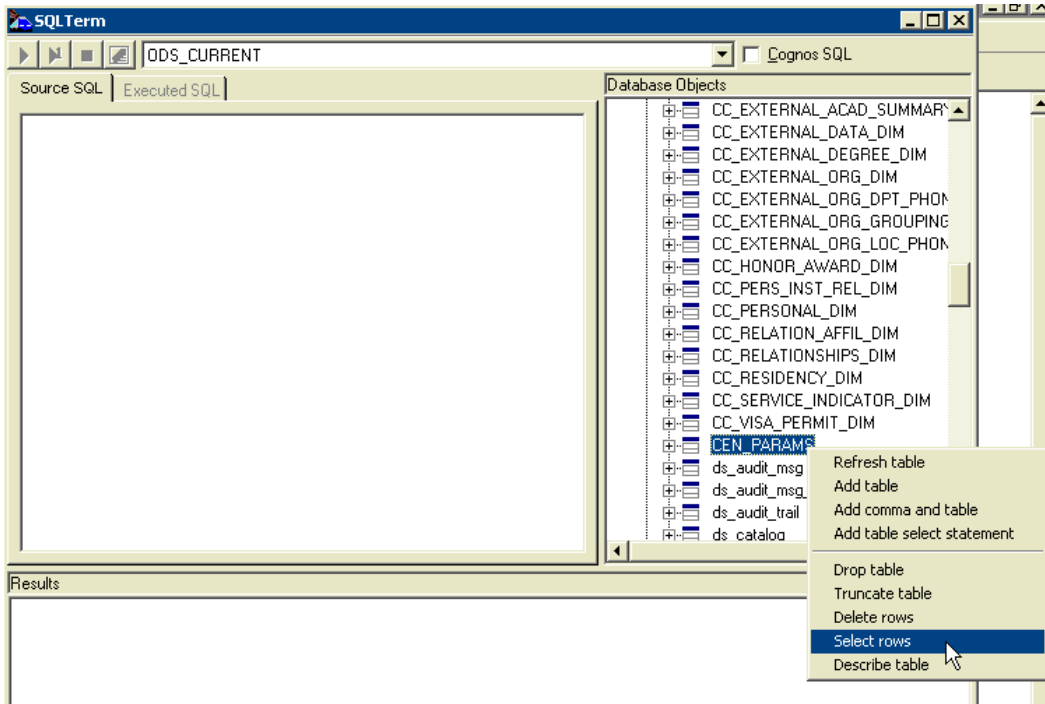
4. Click the “+” (plus) button next to “dbo” to open the database schema and view the tables in the database.



5. Scroll down to the “CEN_PARAMS” table. This table contains the parameters used to extract data from the source database.

Topic Title **Updating the Data Mart: Preparing and Running Jobstreams**

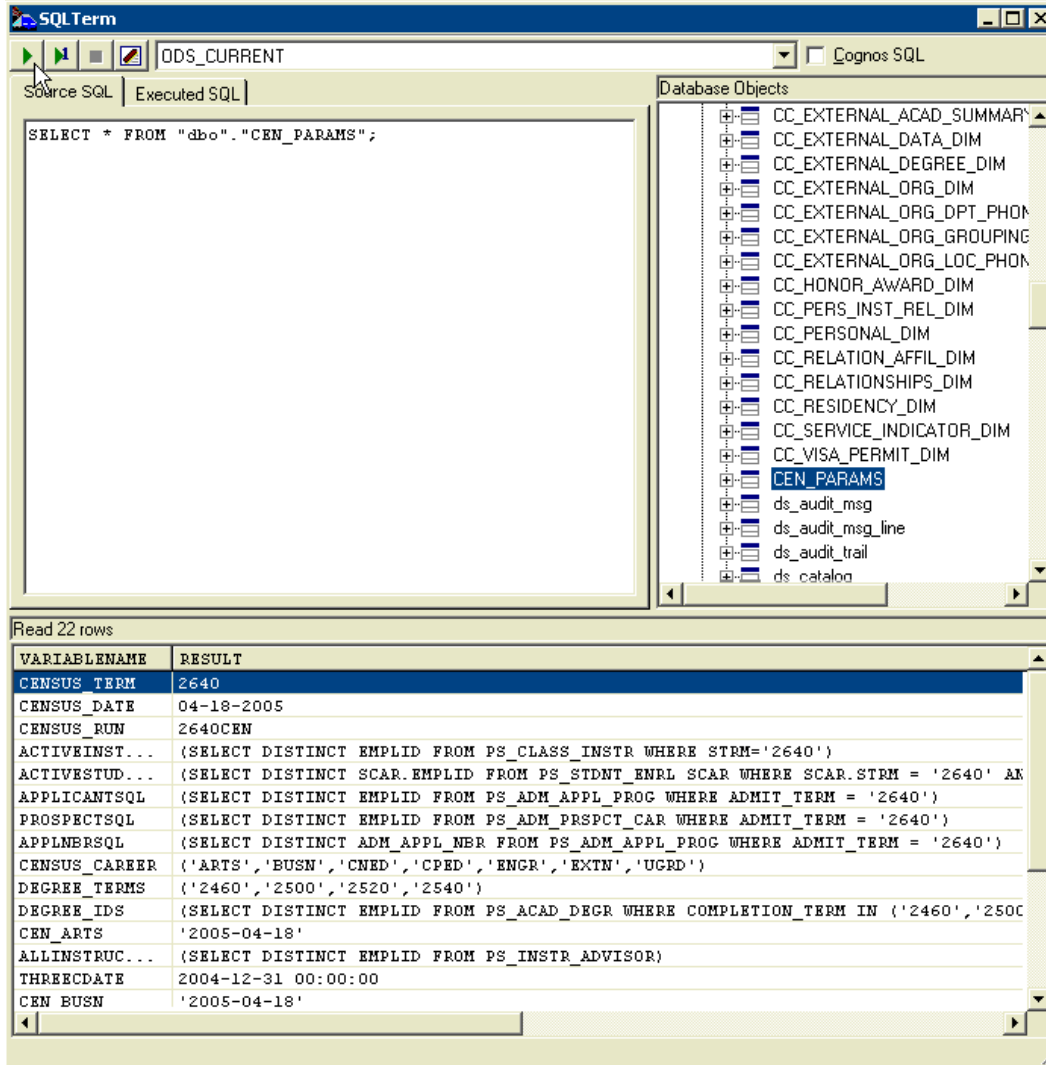
6. Right-click on “CEN_PARAMS” and select “Select rows”.



7. The SQL command ‘SELECT * FROM “dbo”.”CEN_PARAMS”’;’ should appear in the “Source SQL” frame. This command will select all rows from the “CEN_PARAMS” table.

Topic Title Updating the Data Mart: Preparing and Running Jobstreams

- Run the SQL command by clicking the button with a green triangle just above the “Source SQL” frame.



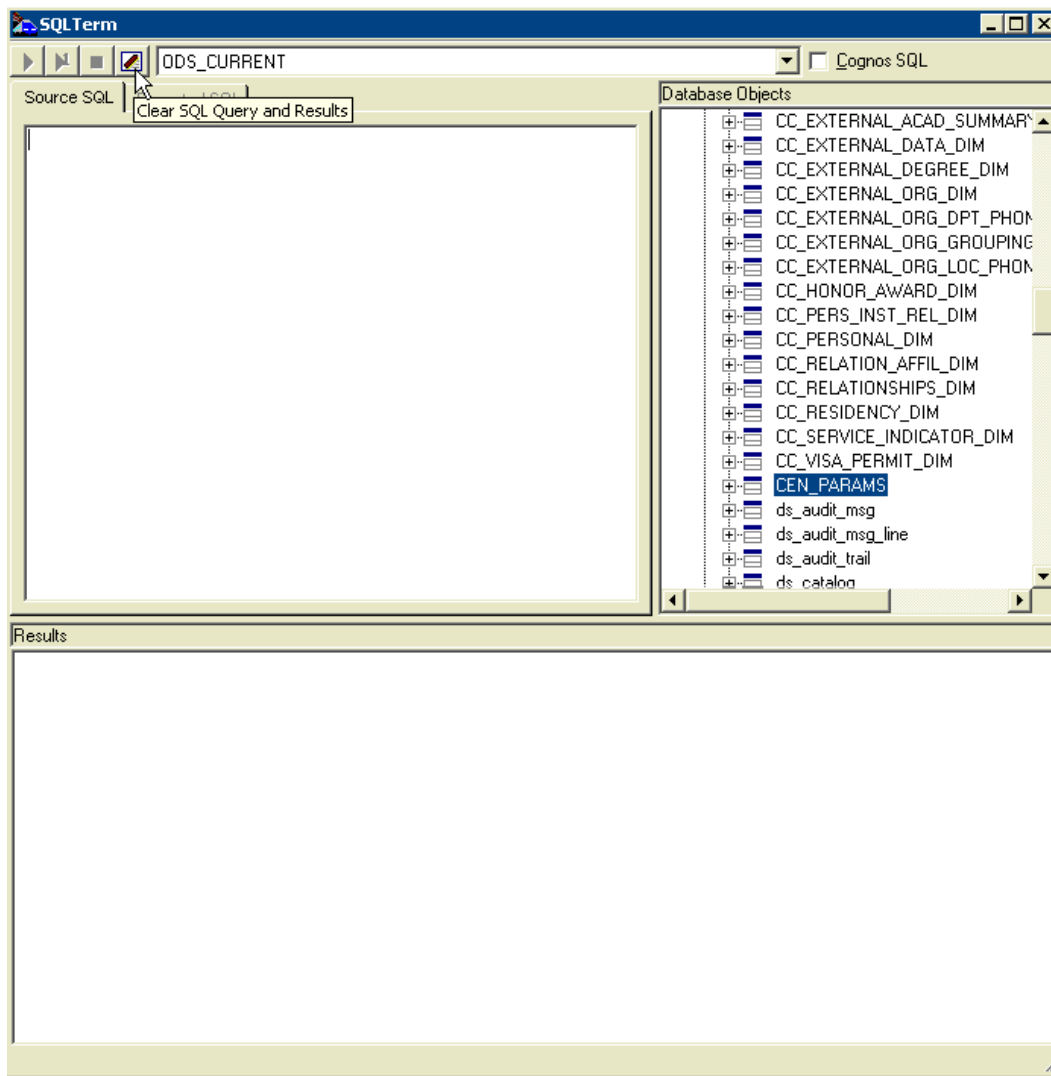
- Running the command should result in rows of data appearing in the “Results” frame at the bottom of the screen. Each row should have two columns: “VARIABLENAME” and “RESULT”. The first column, “VARIABLENAME,” contains the names of variables (i.e. the parameters) used in extracting the new data. The second column, “RESULT,” is the current value associated with the “VARIABLENAME”. For example, the first row contains the value “CENSUS_TERM” in “VARIABLENAME” while the value in “RESULT” is “2640”. When new data are extracted, whenever the “CENSUS_TERM” variable is used, the value “2640” is substituted.

Topic Title **Updating the Data Mart: Preparing and Running Jobstreams**

Set parameters

To set the values for parameters, you must run SQL update statements. All of the necessary SQL statements are in **Appendix A – Modifying Parameters in DMIR**. The easiest approach is to open the digital version of that document, find the SQL statement you want to use, change the values as described below, and copy the statement into DecisionStream’s SQL Term window. However, there is a special process for copying text from your computer and pasting it into the remote computer.

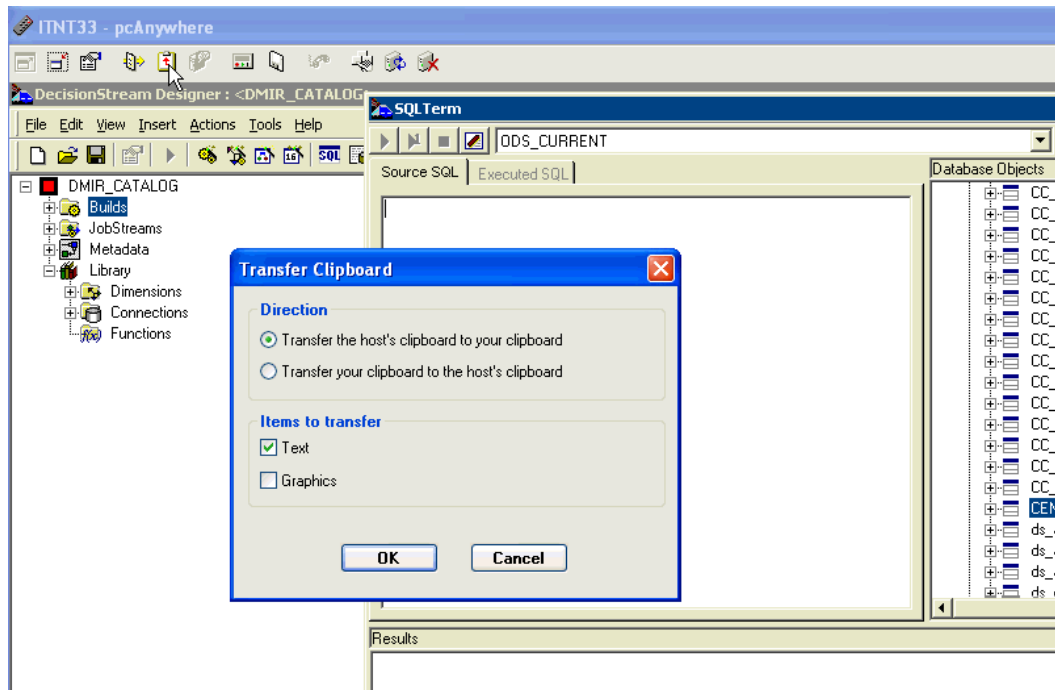
1. The copying part of the process is unchanged. Highlight the text and press CTRL-C to copy.
2. On the remote computer in the SQL Term window, make sure there is nothing in the Source SQL area. If there is text in that window, press the “Clear SQL Query and Results” icon.



3. On the pcAnywhere toolbar, click the “Transfer Clipboard” icon. In the “Direction” section, click the radio button next to “Transfer your clipboard to the host’s clipboard.” In the “Items to

Topic Title **Updating the Data Mart: Preparing and Running Jobstreams**

transfer” section, make sure the “Text” checkbox is checked. Click “OK.”



4. Place your cursor in the Source SQL frame of the SQL Term window and press CTRL-V to paste your text.

Annual extract of degree data

If you are running an annual extract, you are collecting new data for degrees conferred during the previous year. You must set the values for two parameters: “DEGREE_TERMS” and “DEGREE_IDS.” This requires running two SQL update statements. You may choose to run these separately or simultaneously. The steps below will show you how to run them simultaneously.

1. For “DEGREE_TERMS” update the “RESULT” field so that its value is a comma-separated list of PeopleSoft term codes. For example, to collect data on degrees conferred during the 2003-04 academic year, your “RESULT” value should be: ('2460','2500','2520','2540').
2. For “DEGREE_IDS,” update the “RESULT” field so that it contains the same comma-separated list of PeopleSoft term codes that you used in the previous step.
3. Copy the following code into the Source SQL frame:


```
UPDATE CEN_PARAMS
SET RESULT=('2460','2500','2520','2540')
```

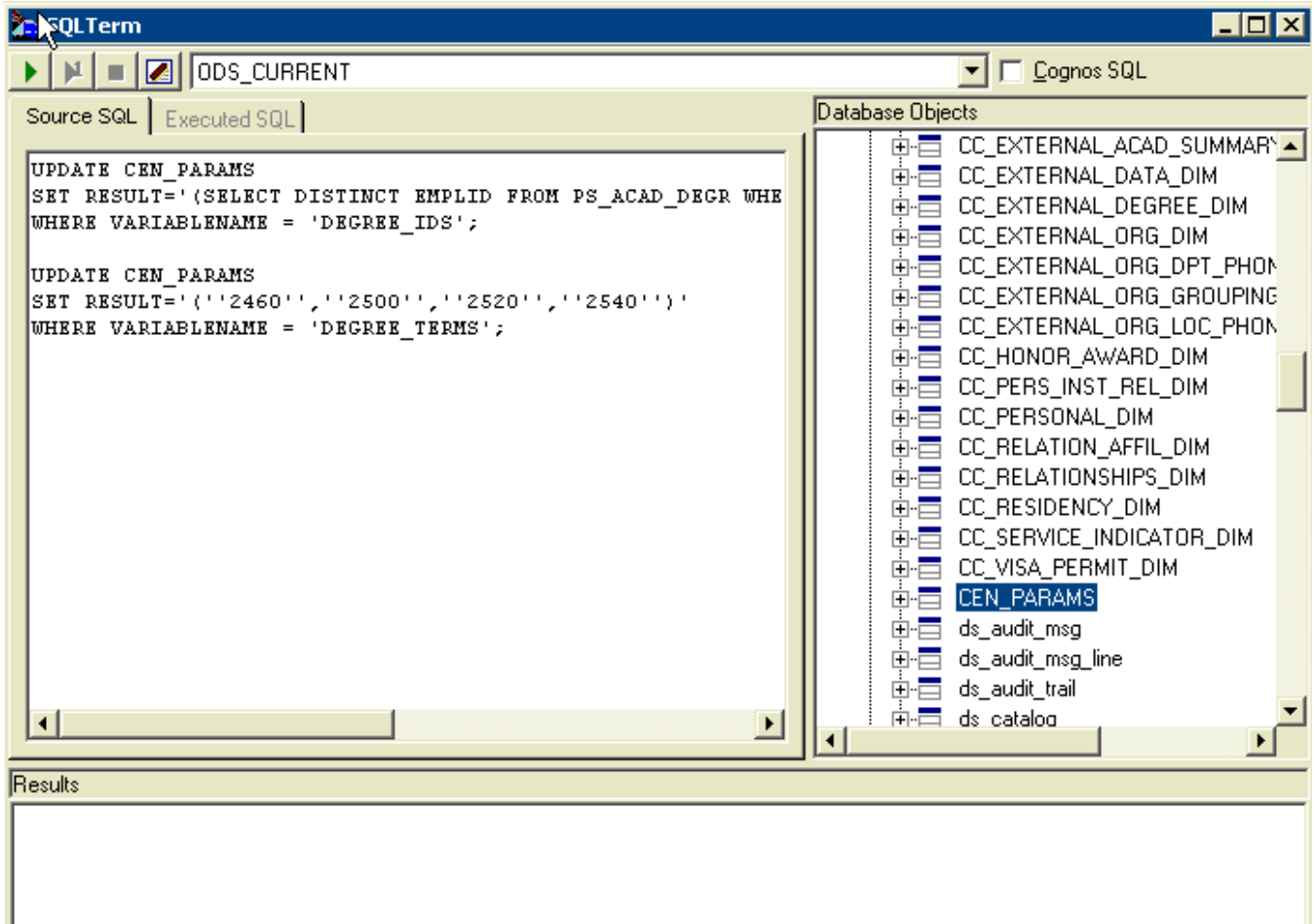
Topic Title **Updating the Data Mart: Preparing and Running Jobstreams**

```
WHERE VARIABLENAME = 'DEGREE_TERMS';
```

```
UPDATE CEN_PARAMS
```

```
SET RESULT='(SELECT DISTINCT EMPLID FROM PS_ACAD_DEGR WHERE COMPLETION_TERM  
IN (''2460'', ''2500'', ''2520'', ''2540''))'
```

```
WHERE VARIABLENAME = 'DEGREE_IDS';
```



4. Run the SQL command by clicking the button with a green triangle just above the Source SQL frame. In the Results frame, you should see a copy of the command you ran and the phrase "1 row updated." Once you've followed the preceding steps, you're ready to run the jobstream to update DMIR. Proceed to **Annual extract of degree data jobstream** (p. 18).

Topic Title **Updating the Data Mart: Preparing and Running Jobstreams**

Term extract of data

If you are running a term extract, you are extracting new data for a given term. You must set the values for many parameters related to admissions, student records, faculty, and classes.

1. For “CENSUS_TERM” update the “RESULT” field so that it contains the PeopleSoft code for the appropriate term.
2. For “CENSUS_DATE” update the “RESULT” field so that it contains the effective date for extracting data. If you are extracting census enrollment data, use the census date obtained from either the Student Records Office or the PS_ACAD_CALTRM_TBL table. If you are extracting final enrollment data, use the current date.
3. The “CENSUS_RUN” variable should contain a unique identifier for the data being extracted. The format for the value associated with this variable is: [term][‘CEN’ or ‘FIN’][‘LAW’ or nothing]. For example, the census enrollment extract for the School of Law in Fall 2004 would use the value “2600CENLAW” while the final enrollment extract for all careers except Law in Winter 2005 would use the value “2620FIN”.
4. Copy the following code into the “Source SQL” frame:

```
UPDATE CEN_PARAMS  
SET RESULT = '2600'  
WHERE VARIABLENAME='CENSUS_TERM';
```

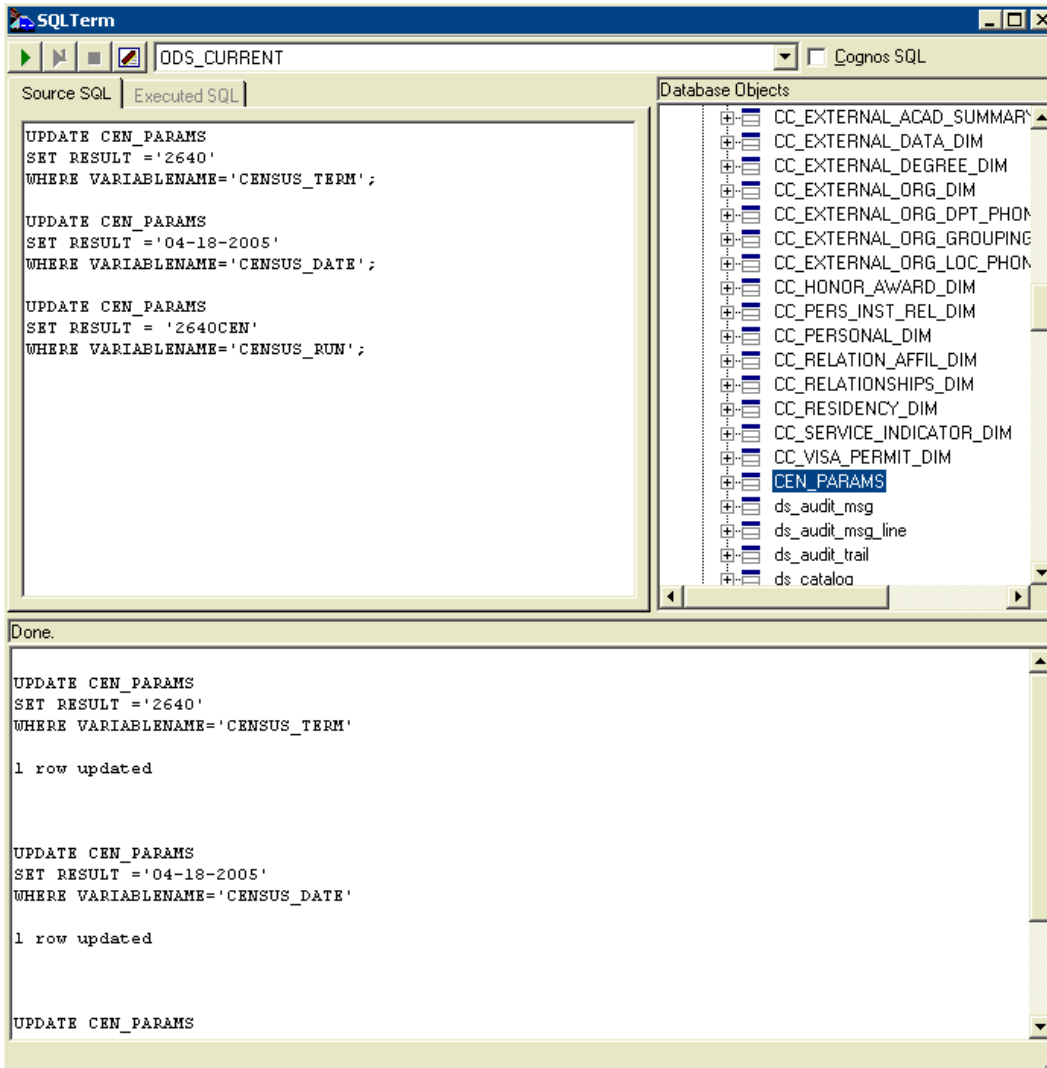
```
UPDATE CEN_PARAMS  
SET RESULT = '09-27-2004'  
WHERE VARIABLENAME='CENSUS_DATE';
```

```
UPDATE CEN_PARAMS  
SET RESULT = '2600CEN'  
WHERE VARIABLENAME='CENSUS_RUN';
```

5. Run the SQL command by clicking the button with a green triangle just above the Source SQL frame. In the Results frame, you should see a copy of each command you ran and the phrase “1 row

Topic Title Updating the Data Mart: Preparing and Running Jobstreams

updated.”



- Following the same process as above, modify the values for the following parameters so that each contains the appropriate term code:

VARIABLENAME	RESULT
ACTIVEINSTRUCTORSQ	UPDATE CEN_PARAMS SET RESULT = '(SELECT DISTINCT EMPLID FROM PS_CLASS_INSTR WHERE STRM="2600")' WHERE VARIABLENAME='ACTIVEINSTRUCTORSQ';
ACTIVESTUDENTSQ	UPDATE CEN_PARAMS SET RESULT = '(SELECT DISTINCT SCAR.EMPLID FROM PS_STDNT_ENRL SCAR WHERE SCAR.STRM = "2600" AND

Topic Title **Updating the Data Mart: Preparing and Running Jobstreams**

VARIABLENAME	RESULT
	<pre>SCAR.STDNT_ENRL_STATUS = "E")' WHERE VARIABLENAME='ACTIVESTUDENTSQ';</pre>
APPLICANTSQ	<pre>UPDATE CEN_PARAMS SET RESULT = '(SELECT DISTINCT EMPLID FROM PS_ADM_APPL_PROG WHERE ADMIT_TERM = "2600")' WHERE VARIABLENAME='APPLICANTSQ';</pre>
PROSPECTSQ	<pre>UPDATE CEN_PARAMS SET RESULT = '(SELECT DISTINCT EMPLID FROM PS_ADM_PRSPCT_CAR WHERE ADMIT_TERM = "2600")' WHERE VARIABLENAME='PROSPECTSQ';</pre>
APPLNBRSQL	<pre>UPDATE CEN_PARAMS SET RESULT = '(SELECT DISTINCT ADM_APPL_NBR FROM PS_ADM_APPL_PROG WHERE ADMIT_TERM = "2600")' WHERE VARIABLENAME = 'APPLNBRSQL';</pre>

7. The “CENSUS_CAREER” variable should contain a comma-separated list of careers for which you want to extract data. Copy the following code into the Source SQL frame:

- If you are running an extract for the School of Law:

```
UPDATE CEN_PARAMS
SET RESULT = ('LAWS')
WHERE VARIABLENAME = 'CENSUS_CAREER';
```

- If you are running an extract for non-Law data:

```
UPDATE CEN_PARAMS
SET RESULT = ('ARTS','BUSN','CNED','CPED','ENGR','UGRD')
WHERE VARIABLENAME = 'CENSUS_CAREER';
```

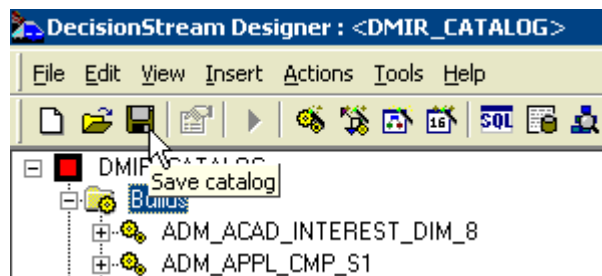
8. Run the SQL command by clicking the button with a green triangle just above the Source SQL frame. In the Results frame, you should see a copy of the command you ran and the phrase “1 row updated”.
9. Finally, set an extract date for each career. Use the following command with each of these variables: CEN_ARTS, CEN_BUSN, CEN_CNED, CEN_CPED, CEN_ENGR, CEN_LAWS. Use the same date that you used as the value for “CENSUS_DATE” above.

Topic Title **Updating the Data Mart: Preparing and Running Jobstreams**

```
UPDATE CEN_PARAMS  
SET RESULT = "'2004-10-10'"  
WHERE VARIABLENAME = 'CEN_ARTS';
```

10. When you're finished setting all of the necessary parameters, close the SQL Term window and save the changes you made to the catalog.

- Click on the "Save" icon.
- Or, navigate through the menu to File > Save Catalog.



Topic Title **Updating the Data Mart: Preparing and Running Jobstreams**

Run and review the jobstreams

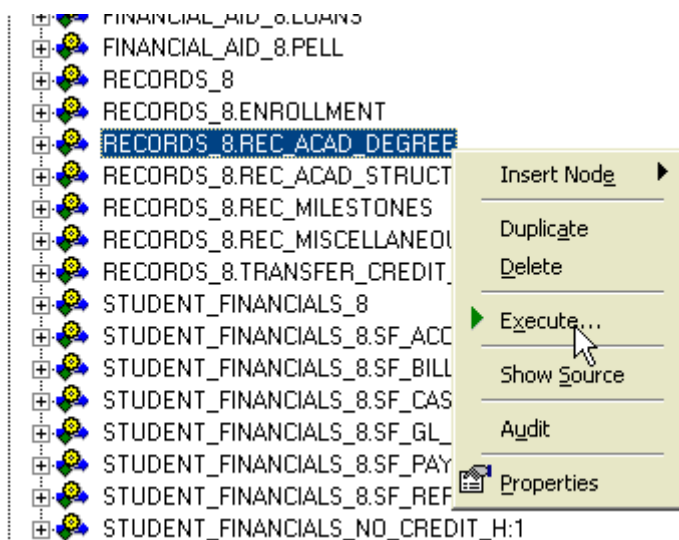
Jobstreams are processes that use the parameters you set to extract data from the source database (PeopleSoft) and insert the data into the target database (DMIR). In the main DecisionStream window, run the appropriate jobstreams depending on whether you’re doing a term or an annual update.

Note: If there are errors during the jobstream process, a pop-up window will alert you and provide further information. Techniques for identifying and handling jobstream errors are covered in **Session 2** of your training.

Annual extract of degree data jobstream

If you are running an annual extract, you are collecting new data for degrees conferred during the previous year.

1. In the DMIR_CATALOG (left) frame of the main window, click the “+” (plus) button next to “JobStreams” to open this list of jobstreams.
2. Scroll down to the “RECORDS_8.REC_ACAD_DEGREE” jobstream.
3. Right-click the “RECORDS_8.REC_ACAD_DEGREE” jobstream and select “Execute” from the popup menu. A command window will show the progress of the jobstream.



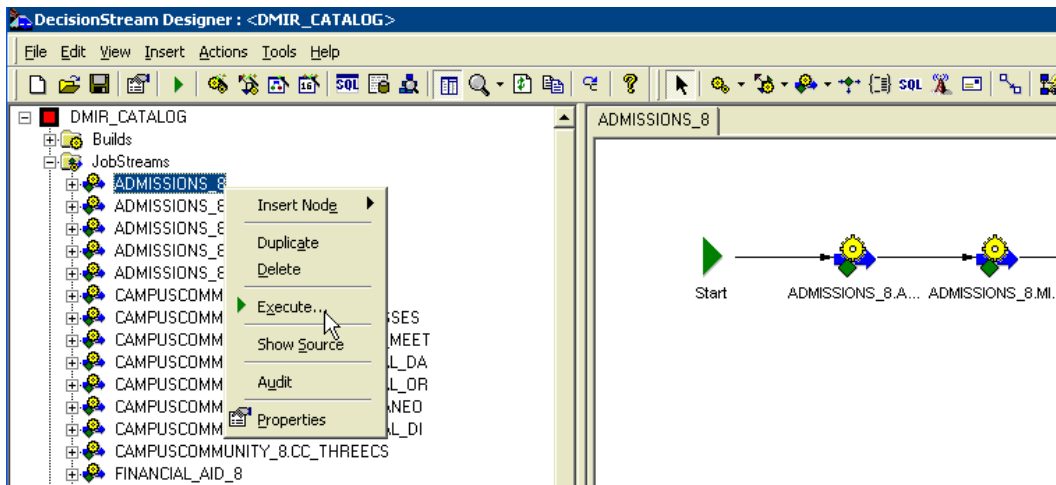
4. When the jobstream is finished running, the command window will alert you and prompt you to close the command window.
5. Proceed to **Review the results of the jobstream(s)**.

Topic Title **Updating the Data Mart: Preparing and Running Jobstreams**

Term extract of data jobstream

If you are running a term extract, you are collecting new data for a given term.

1. In the DMIR_CATALOG (left) frame of the main window, click the “+” (plus) button next to “JobStreams” to open this list of jobstreams.
2. Scroll down to the “ADMISSIONS_8” jobstream. This jobstream contains these subordinate jobstreams, which are run when you execute the “ADMISSIONS_8” jobstream:
 - ADMISSIONS_8.APPLICATION
 - ADMISSIONS_8.MISCELLANEOUS
 - ADMISSIONS_8.PROSPECT
 - ADMISSIONS_8.RECRUITER
3. Right-click the “ADMISSIONS_8” jobstream and select “Execute” from the popup menu. A command window will show the progress of the jobstream. When the jobstream is finished running, the command window will alert you and prompt you to close the command window.



4. Follow the same process for the following jobstreams (each of which has its own subordinate jobstreams listed below it):
 - CAMPUSCOMMUNITY_8
 - CAMPUSCOMMUNITY_8.CC_ADDRESSES
 - CAMPUSCOMMUNITY_8.CC_CAMPUS_MEET
 - CAMPUSCOMMUNITY_8.CC_EXTERNAL_DA
 - CAMPUSCOMMUNITY_8.CC_EXTERNAL_OR
 - CAMPUSCOMMUNITY_8.CC_MISCELLANEO

Topic Title Updating the Data Mart: Preparing and Running Jobstreams

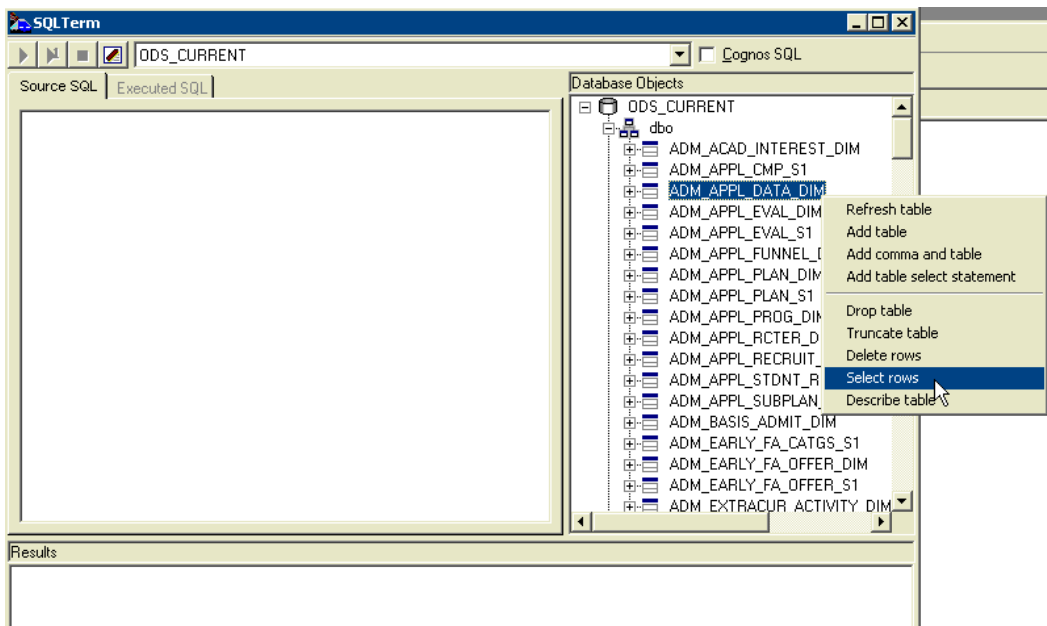
- CAMPUSCOMMUNITY_8.CC_PERSONAL_DI
- CAMPUSCOMMUNITY_8.CC_THREECS
- RECORDS_8
 - RECORDS_8.ENROLLMENT
 - RECORDS_8.REC_ACAD_STRUCTURE_DIM
 - RECORDS_8.REC_MILESTONES
 - RECORDS_8.REC_MISCELLANEOUS
 - RECORDS_8.TRANSFER_CREDIT_EVALUA

Topic Title **Updating the Data Mart: Preparing and Running Jobstreams**

Review the results of the jobstream(s)

After you've run your jobstream(s), review the results to make sure DMIR was updated.

1. Open the SQL Term window as described in **Review and set parameters for extracting new data** (p. 7).
2. In the Database Object frame, select one of the tables that were updated, such as the ADM_APPL_DATA_DIM table if you extracted term data.



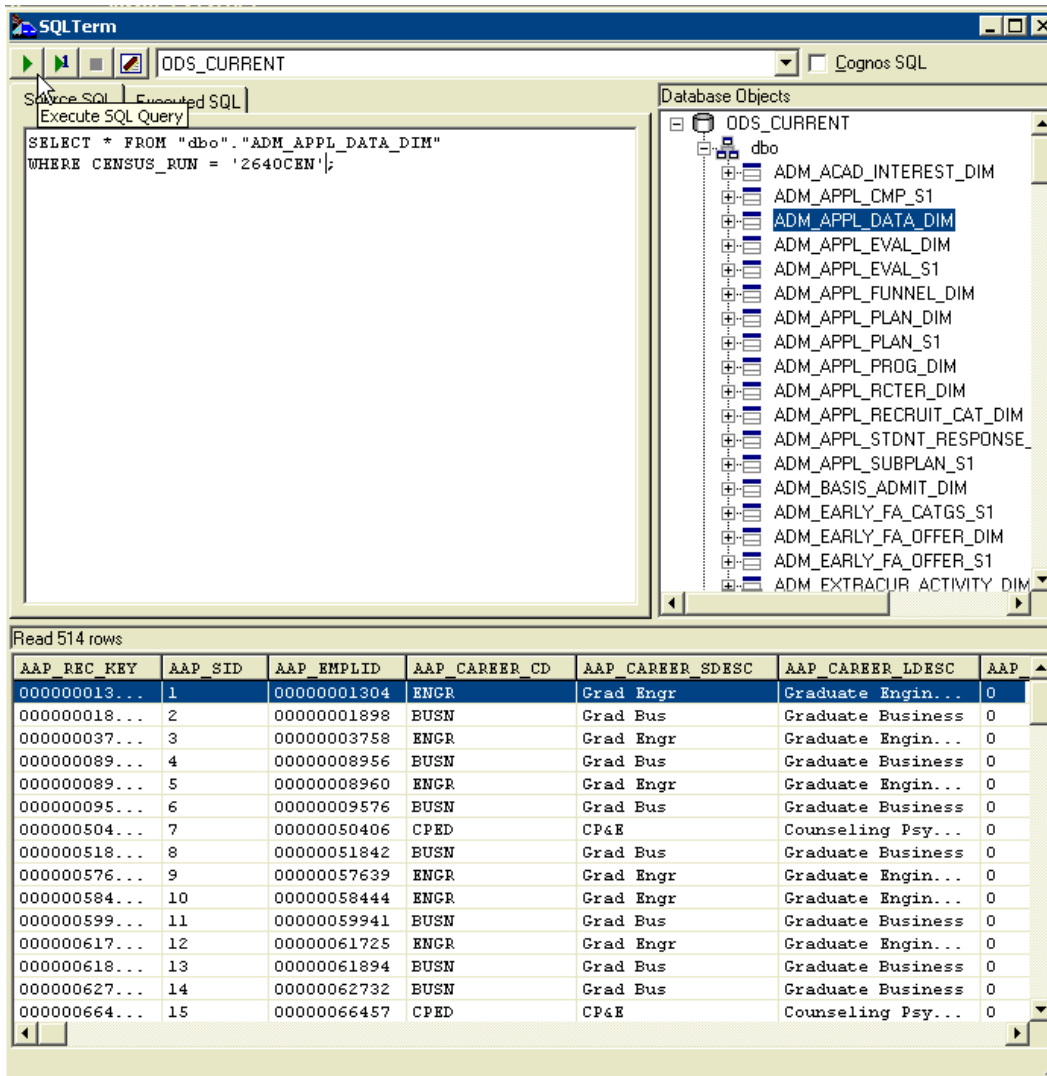
3. Modify the SQL statement in the Source SQL frame by adding a “WHERE” clause. This is used to specify a condition for filtering the results. If you extracted census enrollment data for the summer term of 2004, for example, the SQL statement would look like this:

```
SELECT * FROM "dbo"."ADM_APPL_DATA_DIM"
WHERE CENSUS_RUN = '2640CEN';
```

Note that the semi-colon is at the end of the statement. If you add the “WHERE” clause after the semi-colon, it will be ignored by the SQL engine that runs your query.

4. Run the SQL command by clicking the button with a green triangle just above the “Source SQL” frame. In the Results frame, you will see all of the rows of data which were extracted from the source database and inserted into DMIR by the jobstream(s).

Topic Title Updating the Data Mart: Preparing and Running Jobstreams



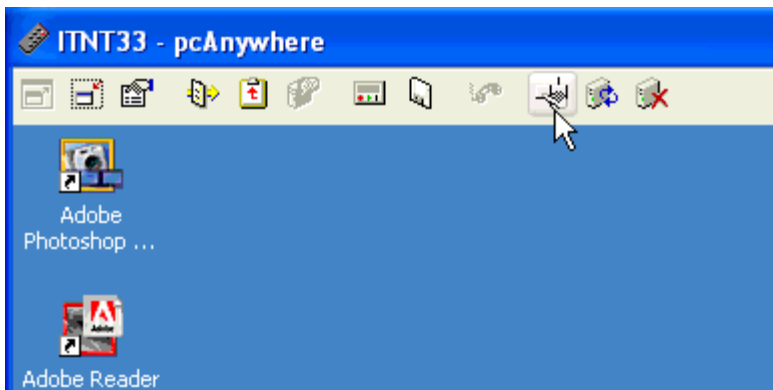
The fact that there are results in the Results frame indicates that there are data in DMIR with a “CENSUS_RUN” value of ‘2640CEN.’

- Query tables from other jobstreams using the same “WHERE” clause to ensure that they received new data as well.

Topic Title **Updating the Data Mart: Preparing and Running Jobstreams**

Disconnect from the remote computer

1. When you're finished reviewing the data, close DecisionStream. Then click the "disconnect" icon in the pcAnywhere toolbar.



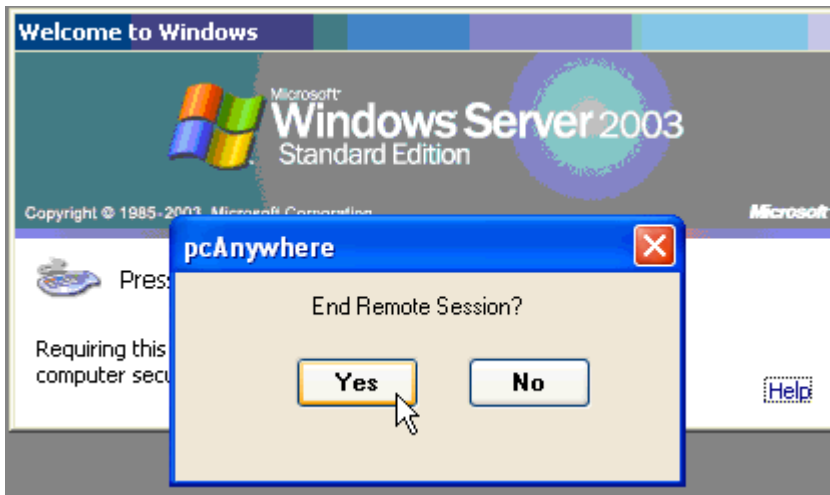
2. When the Windows Security window appears, click the "Log Off" button.



3. A Windows dialog box will ask "Are you sure you want to log off?" If you are sure, click the "Log Off" button. Otherwise click the "No" button.

Topic Title **Updating the Data Mart: Preparing and Running Jobstreams**

4. And finally, you will be prompted to end the remote session. Click “Yes.”



Topic Title **Updating the Data Mart: Preparing and Running Jobstreams**

Appendix A – Modifying Parameters in DMIR

The SQL statements below are used to update the parameters in Cognos DecisionStream that determine what data to extract from the PeopleSoft Student Administration system.

If you're extracting annual data, you'll be running the REC_ACAD_DEGREE jobstream. Change the values for the DEGREE_IDS and DEGREE_TERMS parameters to indicate the terms for which you want to extract data. Then cut and paste these statements into the SQL Term window.

```
UPDATE CEN_PARAMS
SET RESULT='(SELECT DISTINCT EMPLID FROM PS_ACAD_DEGR WHERE
COMPLETION_TERM IN ("2460","2500","2520","2540"))'
WHERE VARIABLENAME = 'DEGREE_IDS';
```

```
UPDATE CEN_PARAMS
SET RESULT='("2460","2500","2520","2540")'
WHERE VARIABLENAME = 'DEGREE_TERMS';
```

If you're extracting term data, you'll be running **all jobstreams except REC_ACAD_DEGREE**. The first set of parameters define the term and date to use. The "CENSUS_RUN" parameter consists of the term (e.g. '2600') and a string to distinguish among extracted data within a term (e.g. 'CEN' for the census run, 'FIN' for the final run, 'TMP' for a temporary run that will later be deleted).

```
UPDATE CEN_PARAMS
SET RESULT = '2600'
WHERE VARIABLENAME='CENSUS_TERM';
```

```
UPDATE CEN_PARAMS
SET RESULT = '09-27-2004'
WHERE VARIABLENAME='CENSUS_DATE';
```

```
UPDATE CEN_PARAMS
SET RESULT = '2600TMP'
WHERE VARIABLENAME='CENSUS_RUN';
```

These next parameters define various populations. For most of them you need to modify only the term being queried. For "census_career" we run 'LAWS' on one date and all other careers on another date.

```
UPDATE CEN_PARAMS
SET RESULT = '(SELECT DISTINCT EMPLID FROM PS_CLASS_INSTR WHERE STRM="2600")'
WHERE VARIABLENAME='ACTIVEINSTRUCTORSQL';
```

```
UPDATE CEN_PARAMS
SET RESULT = '(SELECT DISTINCT SCAR.EMPLID FROM PS_STDNT_ENRL SCAR WHERE
SCAR.STRM = "2600" AND SCAR.STDNT_ENRL_STATUS = "E")'
WHERE VARIABLENAME='ACTIVESTUDENTSQL';
```

```
UPDATE CEN_PARAMS
```

Topic Title **Updating the Data Mart: Preparing and Running Jobstreams**

```
SET RESULT = '(SELECT DISTINCT EMPLID FROM PS_ADM_APPL_PROG WHERE  
ADMIT_TERM = "2600")'  
WHERE VARIABLENAME='APPLICANTSQL';
```

```
UPDATE CEN_PARAMS  
SET RESULT = '(SELECT DISTINCT EMPLID FROM PS_ADM_PRSPCT_CAR WHERE  
ADMIT_TERM = "2600")'  
WHERE VARIABLENAME='PROSPECTSQL';
```

```
UPDATE CEN_PARAMS  
SET RESULT = '(SELECT DISTINCT ADM_APPL_NBR FROM PS_ADM_APPL_PROG WHERE  
ADMIT_TERM = "2600")'  
WHERE VARIABLENAME = 'APPLNBRSQL';
```

```
UPDATE CEN_PARAMS  
SET RESULT = '(SELECT DISTINCT EMPLID FROM PS_INSTR_ADVISOR)'  
WHERE VARIABLENAME = 'ALLINSTRUCTORSQL';
```

```
UPDATE CEN_PARAMS  
SET RESULT = '2003-12-31 00:00:00'  
WHERE VARIABLENAME='THREECDATE';
```

```
UPDATE CEN_PARAMS  
SET RESULT = '("ARTS","BUSN","CNED","CPED","ENGR", "UGRD")'  
WHERE VARIABLENAME = 'CENSUS_CAREER';
```

```
UPDATE CEN_PARAMS  
SET RESULT = '("LAWS")'  
WHERE VARIABLENAME = 'CENSUS_CAREER';
```

Set the census dates for each career. Below is an example for "CEN_ARTS":

```
UPDATE CEN_PARAMS  
SET RESULT = "'2004-10-10'"  
WHERE VARIABLENAME = 'CEN_ARTS';
```

Also CEN_BUSN, etc.