

Data Warehouse

Notes about the Data

Sharon Schmitz
Institutional Research - X3087
Sharon.Schmitz@wwu.edu

Regular Student Tables vs. Peak Tables: The Data Warehouse now has two student models. The main model is “live” student data that builds nightly using current Banner data. When you click the button labeled “Peak” on the main model, you get to the Data Warehouse Peak Tables. The Peak Tables are essentially a copy of the main student model (does not include the HISTORY tables). They are copied at the end of business on the peak enrollment day for each term. These peak tables are used by the Registrar’s Office for official reporting purposes. These peak tables contain snapshots of each term since Fall 2001. On the day of peak enrollment for each new term, another snapshot is added to the peak tables.

Table Access: If you attempt to run a query using a table that you don’t have access to, you will get a message something like this: “ORA-0942 - Object or view does not exist”.

Metadata: text description of data

A short table description shows up when you mouse-over a table on the roadmap. A description of each field is available by clicking the question mark next to the attribute. Click the ‘Load Comment’ button. When you’re finished viewing the comment, click ‘Cancel’ so that you will be reminded to ‘Load Comment’ next time (in case it changes).

Tables:

- STUDENT – one record for each student. The term associated with the student is the last term that they were eligible to enroll.
- REG_SUMMARY & REG_DETAIL – this is the registration info for a student by term such as student status and type, add/drop dates, billing info, major info, etc. REG_SUMMARY has one record per term per student, and REG_DETAIL has one record per course per student. These tables only go back about two terms into the past.
- HIST_TERM & HIST_DETAIL – these tables are similar to above, but are more a picture of the end of the term. HIST_TERM has term grade information, and HIST_DETAIL has grades by course. HIST_TERM has some data as far back as 1935, but realistic enrollment numbers beginning in 1984.
- SCHEDULE – one record for each course offered – includes off-campus, correspondence, courses with zero enrollment, etc. Data in this table goes back to Fall 1999. The SCHEDULE table may be used in conjunction with the MEETING and INSTRUCTOR tables for detailed course information (described further below).
- DEGREE – one record associated with each degree. However, each record does not necessarily represent an “Awarded” degree. See DEGREE_STATUS and DEGREE_STATUS_DESC to determine whether the degree has been awarded.
- HIST_CUM_HRS & HIST_CUM_GPA store cumulative credits and grades. HIST_CUM_HRS is the same as HIST_CUM_GPA except for GPA info.
- MAJOR_CODES gives information about major codes
- PREVC – One record per previous college attended – Registrar’s record of current credits applied toward degree.
- The ADMITPREVC, TESTSCORES, HIGHSCHOOL, S_ADMITS and S_HS_SUBJECT tables contain student Admissions information.

SELECTING A SAMPLE OF STUDENTS: If you want current enrollment or information pertaining to the beginning of a recent term (registration), use REG_SUMMARY. If you want information that pertains to end-of-term information, use HIST_TERM. Do not use TERM in the STUDENT table to select students. If you need information from the STUDENT table, first select the students using TERM in either REG_SUMMARY or HIST_TERM, then link to the student table (not selecting TERM). You can however, use COHORT to select students from the STUDENT table. If you need all students who started as matriculated students in a certain term, you can use COHORT in the STUDENT table. COHORT is also in the HIST_TERM & REG_SUMMARY tables. If you use COHORT from one of these tables, you'll want to also select a TERM. COHORT is 6-digit term followed by student type at admission (see the values for STUDENT_TYPE). So, if you want all freshmen that started in 200040, select COHORTs 200040F and 200040X (F=Freshmen X=Running Start).

HIST_TERM table: When using this table to get all students that were enrolled for a specific term, your query may return more students than you want. You may or may not want to use these qualifiers depending on your purposes.

CAMPUS1=M *removes other off-campus (O) and SPAN (S)
CLASS != 00 *leaves out non-matriculated students
LEVEL_CODE=UG *gets undergrads only

REG_SUMMARY table: The same general idea as above applies – look at variables, and be sure you're getting the sample you want.

PERSON table: DEAD_IND – Exclude deceased students when appropriate.

PERSON table: CONFID_IND – Consider excluding students who have indicated that they do not wish to be identified as associated with Western in any way.

In the SCHEDULE table, many of the fields begin with PRIMARY. For example, if there are two instructors associated with a course, only the PRIMARY one will be included here. If you want all the instructors that are associated with a course, you identify the course(s) in the SCHEDULE table and select the instructor information from the INSTRUCTOR table. It works the same for the MEETING table. You can get the primary meeting time/place from the SCHEDULE table, but if you want all the meeting times or places, select them from the MEETING table. See the attached screen print.

MEETING table: value of O- in BUILDING means course meets off campus. If you need to know which off-campus site this is, select the SITE variable.

SCHEDULE table: Some courses are offered for variable credit. That is, students may register for a course at different credit levels. For example, a course may be available for registration at one credit OR three (1,3). A different course may be available for registration at 5 TO 7 credits (5,6,7). The variable CREDIT_HRS_IND has values OR and TO. When used in conjunction with CREDIT_HRS_LOW and CREDIT_HRS_HIGH, the allowed values of credit level are defined. BILLING_HRS_IND works the same way.

