PENNANT STUDENT RECORDS data collection in the Data Warehouse

Training and resources

Descriptions and long descriptions for codes

Most codes used in Pennant Student Records can be decoded using tables containing all the valid values for the codes, along with a brief description or text name of each code. These validation tables with table names beginning with "V_" are found in the DWNGSS schema in the database and in the "Validation Tables" folder in the Pennant Student Records universe.

For example, division codes and division names are found in the table V_DIVISION. Here are a few of the valid values in that table:

DIVISION_CODE	DIVISION_DESC
AM	LPS Prof Masters
AP	Graduate Arts & Sciences
AU	College of Arts & Sciences
CM	Annenberg Masters
CP	Annenberg PhD

In many of the student and course data tables, the description is provided alongside the code, and there is no need to go to a validation table to get it. For example, in ST_DEGREE_TERM, you will find both the DIVISION and the DIVISION_DESC for the division(s) of the student's curriculum, in any given term. These in-table descriptions always use the descriptions from the Banner tables.

Because of field length restrictions, the Banner table descriptions are limited to 30 characters.

However, at Penn, sometimes we want to use descriptions or text names that are longer than 30 characters. So we have a custom solution: a source table maintained by the Office of the University Registrar, containing long descriptions or long names for some codes. We bring those custom longer descriptions into the Data Warehouse and store them alongside their shorter descriptions and codes. Looking again at the division table, V_DIVISION, example:

DIVISION_CODE	DIVISION_DESC	DIVISION_DESC_LONG	LONG_DESC_START_TERM	LONG_DESC_END_TERM
AM	LPS Prog Masters	College of Liberal and Professional Studies Professional Masters	000000	999999

AP	Graduate Arts & Sciences	School of Arts & Sciences - Graduate Division	000000	999999
AU	College of Arts & Sciences	College of Arts & Sciences	000000	999999
CM	Annenberg Masters	Annenberg School for Communication Masters	000000	999999
CP	Annenberg PhD	Annenberg School for Communication Doctoral	000000	999999

The shorter descriptions come from the Banner tables. The longer descriptions come from Penn's custom table. The long descriptions have a start and end term, so that Penn can easily change a long name and still maintain the older values.

Not all validation tables have long descriptions.

For validation tables that do have long descriptions, because we have the start-end terms, there can potentially be more than one row for any given code and its corresponding long description(s). To facilitate reporting, the warehouse has a view of each of the validation tables with long descriptions, containing only the *most recent* long description for each code. The views have the same name as the underlying table, followed by "_V":

Table name	View with most recent long description values
V_CRSE_PROG_ATTR	V_CRSE_PROG_ATTR_V
V_DEG_DEPT_HONORS	V_DEG_DEPT_HONORS_V
V_DEGREE	V_DEGREE_V
V_DEPARTMENT	V_DEPARTMENT_V
V_DIVISION	V_DIVISION_V
V_INSTITUTION	V_INSTITUTION_V
V_MAJOR_MINOR	V_MAJOR_MINOR_V
V_PROGRAM	V_PROGRAM_V
V_SCHOOL	V_SCHOOL_V
V_SITE	V_SITE_V
V_SUBJECT	V_SUBJECT_V
V_TRANSCRIPT_EVENT	V_TRANSCTIPT_EVENT_V

Let's take a look at a hypothetical example of a made-up division:

CODE	DESC	DIVISION_DESC_LONG	LONG_DESC_START_TERM	LONG_DESC_END_TERM
YX	Pseudo Tech	The old name of the pseudo tech division	000000	202030
YX	Pseudo Tech	The Pseudo-Technical Hypothetical Division of the Graduate School	202110	202130
YX	Pseudo Tech	Pseudo-Technical Post-Undergraduate Division	202210	999999

In this made-up example, prior to the Spring of 2021 there was an old description. During the Spring, Summer, and Fall terms of 2021, a second long description was in effect. Starting in the Spring of 2022, the third and newest version of the long description took effect. There are three rows for the YX code. Joining on just the code would triple the number of rows returned in a query.

To avoid getting back multiple rows when using the validation tables with long descriptions you should do one of the following:

1. Join to the validation table where you find a match on the code *and* where the term you are looking at falls between the start and end term.

Or

2. If it is acceptable in your report to use the most recent long description of the code, use the corresponding view (*V_tablename_*V) to get the long description.

Or

3. If it is acceptable in your report, use the shorter description provided in the originating table. For example, a row in ST_DEGREE_TERM with our made-up division would have the DIVISION = 'YX' and the DIVISION_DESC = 'Pseudo Tech'