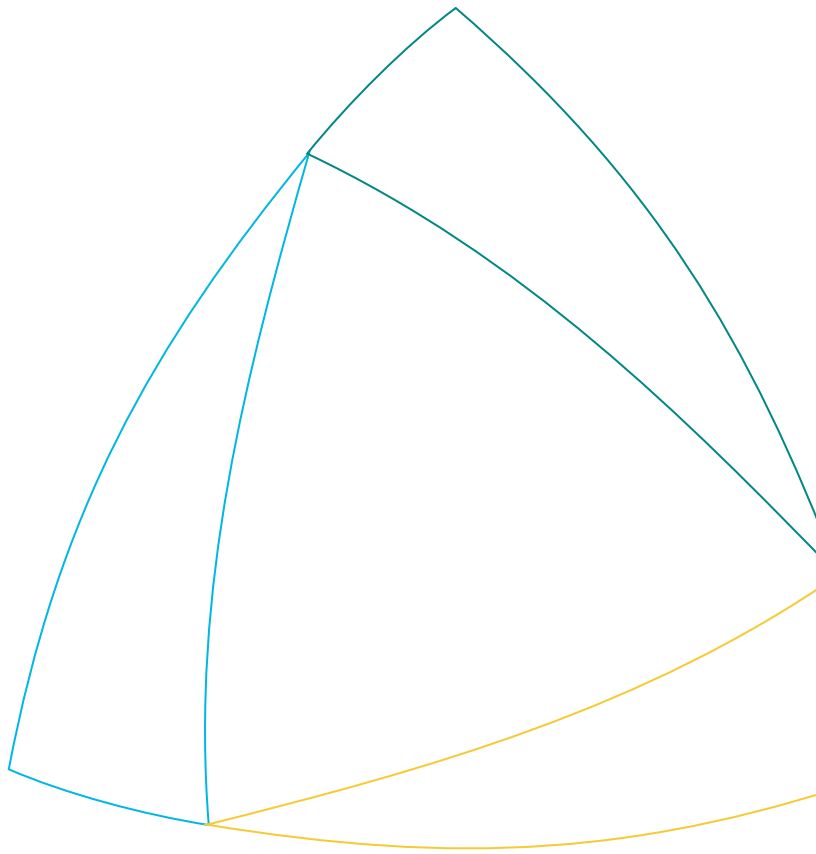# SHEEO

STATE HIGHER EDUCATION EXECUTIVE OFFICERS ASSOCIATION

POSTSECONDARY DATA PARTNERSHIP (PDP) TECHNICAL BRIEF:

LEVERAGING COHORT AND COURSE DATA TO FACILITATE DEVELOPMENTAL COURSE COMPLETION CODING

MARCH 2020

ACKNOWLEDGMENTS

This technical brief was developed in partnership with Achieving the Dream. SHEEO is especially grateful to Bobbie Frye from Achieving the Dream for her technical expertise in creating this document.

SHEEO **PDP TECHNICAL BRIEF:** LEVERAGING COHORT AND COURSE DATA TO FACILITATE DEVELOPMENTAL COURSE COMPLETION CODING

© 2020 by the State Higher Education Executive Officers Association (SHEEO)

2

TABLE OF CONTENTS

SHEEO **PDP TECHNICAL BRIEF:** LEVERAGING COHORT AND COURSE DATA TO FACILITATE DEVELOPMENTAL COURSE COMPLETION CODING

3

INTRODUCTION

This document is part of SHEEO's efforts to support the higher education community with technical assistance[1] related to the National Student Clearinghouse (Clearinghouse) Postsecondary Data Partnership (PDP).[2] The PDP is a national partnership to support institution- and system-level data use and reduce reporting burden. By providing additional data files to the Clearinghouse, institutions and systems have access to Key Performance Indicator (KPI) Tableau dashboards with a specific emphasis on early momentum metrics, an analysis-ready file with PDP submitted and Clearinghouse calculated fields that include Clearinghouse enrollment and degree data, as well as decreased data reporting requirements for institutions and systems participating in initiatives that are also PDP partners.

The objective of this technical brief is to discuss an approach to preparing developmental course completions data for PDP data submission. The data preparation approach was employed by a Central Piedmont Community College's Center for Applied Research (CFAR)[3] analyst during the initial PDP data collection. The analyst used the National Student Clearinghouse's file documentation[4] and sample cohort[5] and course[6] file templates to process files for the cohort and course file extracts. The data preparation approach was also used by multiple institutions during the PDP data collection and yielded successful participation in the PDP.

1. https://sheeo.org/pdp
2. https://www.studentclearinghouse.org/colleges/pdp
3. Central Piedmont Community College, Center for Applied Research: http://cpcc-cfar.com
4. National Student Clearinghouse. *Postsecondary Data Partnership data submission guide. (2019).* https://studentclearinghouse.info/onestop/wp-content/uploads/PDP-Submission-Guide.pdf
5. National Student Clearinghouse. *Postsecondary Data Partnership data submission guide. (2020). PDP cohort template file* https://studentclearinghouse.info/onestop/wp-content/uploads/PDP_Cohort_Template_2020.csv
6. National Student Clearinghouse. *Postsecondary Data Partnership data submission guide. (2020). Course template file* https://studentclearinghouse.info/onestop/wp-content/uploads/PDP_Course_Template_2020.csv

LEVERAGING COHORT AND COURSE DATA TO FACILITATE DEVELOPMENTAL COURSE COMPLETION CODING

The Postsecondary Data Partnership requires institutions to submit two types of files. The first type is a single-record student cohort file that consists of all new students at the institution for the term being reported. The second type of file is a multiple-record course file that consists of students' course records for the term being reported. The cohort file is created for each term and submitted to the Clearinghouse with cohort-identifying information such as cohort term, cohort year, and student ID.

The preparation of the PDP course data files required data fields that needed to be derived by the CFAR analyst. A challenge with the initial course file processing was that information was requested for completion of students' developmental course work. Populating this field requires either manually checking a student's progress in the student information system or applying a 'flag' to the student's course record upon successful completion of the final developmental course requirement in each subject area. The saved cohort file included students' developmental requirement information and provided a reference for initial placement. Subsequent developmental course completions were not stored in the student information system. They were processed utilizing SAS software programming to flag a student's successful completion of the developmental course(s), hence populating these fields appropriately.

COHORT FILES

The PDP cohort files included new students in each term being reported along with other identifying information. These cohort files provided an initial record of developmental placement for math and English.

Saving the cohort files provided a record of developmental placement in all subject areas. The cohort files were then combined to create a record of placement for each student reported to the Clearinghouse. The placement information was gleaned from the aggregate cohort file to derive successful developmental course completion. To begin the process, the analyst combined the cohort files into one SAS cohort data set.

COURSE FILES

The course files included new students in each term being reported. While the data files were able to capture all students in each term being reported, not all of the information was captured in the query process. Information that could not be located in the student information system was left blank, and the analyst used SAS coding to complete the missing values.

The course file is the 'meat' of the longitudinal reporting function of the PDP data and hence includes course grades information, term grade point averages, course CIP codes, credit hours attempted and completed during the term, etc. The main distinction of the course file from the cohort file is that information is associated with the course(s) in which students are enrolled, and multiple instances of student records are provided.

Course grades are typically stored in a student information system utilizing numerical values or letter grades. The PDP uses numerical coding for grading systems on a 4.0 scale. In this example, SAS software was used to translate the letter grades into the appropriate numerical coding. The registrar provided documentation from the institution, which provided the letter grade translations.

In this institutional context, the developmental coursework grades used 'P' for pass and 'F' for not passed. Grading can vary, and 'R' may also be used for attempting and not passing the course. Combining the developmental course attempts and assessing to what extent the student completed the developmental courses required the use of both placement and successful course completions.

The aggregate cohort file was evaluated, and students were flagged if developmental courses were required. The next step was to evaluate the course completion data for each subject area. SAS software was used to create separate files for each developmental course completion, retaining student ID so that the files could be combined. Then the separate files were merged by student ID, retaining the developmental completion fields. Finally, the file was incorporated into the course file, completing the derivation process for the developmental completion fields.

CHALLENGES DURING FILE PREPARATION

One challenge with the initial course file extract was that it included all students enrolled during the term. To create a file that included cohort students, all of the cohort files to date were merged with the term's course file to create a final course file for the term. Since the cohort files included the initial identifying information of cohort term, cohort year, and student ID, this information was incorporated into the course file, and the results captured all cohorts' information for the reported term.

Another challenge with the initial course file extract was that information was requested for completion of students' developmental course work. Populating this field requires either manually checking a student's progress in the student information system or applying a 'C' to the student's course record. Upon successful completion of the final developmental course requirement in each subject area, the student's record was coded as 'C' for completion. The saved cohort file included students' developmental requirement information and provided a reference for initial placement, but subsequent developmental course completions were not stored in the student information system.

Over time, the institution changed the coding of the course prefix from three to four characters, requiring a change in the SAS program to capture the additional character. The analyst used SAS software programming to add the course prefix changes.

RESULTS

The results of using the data collection process outlined here demonstrated an efficient process for handling the PDP data collections. Cohorts of students were reported in the first term of enrollment and also included in subsequent course data collections if enrolled during the reporting term.

The multi-record course files included all of the student cohorts to date, in all courses enrolled during the term. Since course files include the reporting term and term outcomes, the files facilitated longitudinal cohort tracking and metrics measured at points in time along the student's journey.

Students were reported in the cohort files in a single-record instance along with identifying information. The identifying information was stored in a cohort data set and used for subsequent course data collections. Students were reported once in a cohort file, which allowed the analyst to use the information to derive course developmental completion fields by combining the cohort and course information into a series of separate files. The separate files were subsequently incorporated into the final course file and retained the developmental course completion information.

RECOMMENDATIONS

- The initial cohort file can be saved for future use and creates a record for each student that can be used in subsequent course files.

- The multiple-record course file adheres to reporting information that is associated with the course(s) with which students are enrolled, with some exceptions. For example, term grade point averages are associated with the student and captured at the end of the term. Additionally, the cohort-identifying information is associated with the student and/or the term being reported. Pay careful attention to the field requirements and carefully read the requirements to conceptualize the file structure needed.

- If the information is required by the Clearinghouse, missing information will create an error when files are uploaded to the Clearinghouse. File preparers will need to adhere to field requirements and ensure that required fields are provided for each student in the file. This also ensures that the PDP reports will provide quality reports that enhance student success efforts locally and across the network.

- The initial data collection is the most time consuming, including up to five years of data, but the process outlined in this document can be replicated for future data collections.

- While the process can be replicated for future data collections, coding changes in the student information system will need to be addressed in the ongoing query extracts. Missing data fields are often an indication of a change in the location of data stored in the student information system or a change in the field coding. In this context, the institution changed the coding of the course prefix from three to four characters, requiring a change in SAS software to capture the additional character.

- The process to determine developmental course completion is complex, requires longitudinal tracking, and within the current changing developmental policy environment, is subject to coding errors.

APPENDIX A: MACRO VARIABLES ARE ADDED AT THE BEGINNING OF SAS SOFTWARE PROGRAM[7]

```
/*Macro variables identify the term for the course file*/


%LET yt01=2019SU; /*Academic term using term naming convention*/
%LET ch01=2018-19; /*Academic year*/
%LET chtm01=Summer; /*PDP term - Fall, Winter Spring or Summer*/


/*Cohort Terms Reported To Date*/


%LET yt1=2013FA;
%LET yt2=2014SP;
%LET yt3=2014SU;
%LET yt4=2014FA;
%LET yt5=2015SP;
%LET yt6=2015SU;
%LET yt7=2015FA;
%LET yt8=2016SP;
%LET yt9=2016SU;
%LET yt10=2016FA;
%LET yt11=2017SP;
%LET yt12=2017SU;
%LET yt13=2017FA;
%LET yt14=2018SP;
%LET yt15=2018SU;
%LET yt16=2018FA;
%LET yt17=2019SP;
%LET yt18=2019SU;
%LET yt19=2019FA;
```

---

7. SAS Institute Inc. The code for this paper was generated using SAS software. Copyright, SAS Institute Inc. SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc., Cary, NC, USA. SAS support website: https://support.sas.com/en/technical-support.html

APPENDIX B: CODING THE DEVELOPMENTAL COMPLETION FIELDS

```
/*Code Developmental Requirements*/


Data DEVGroup&yt01;
set All .Cohort_&yt1
All .Cohort_&yt2
All .Cohort_&yt3
All .Cohort_&yt4
All .Cohort_&yt5
All .Cohort_&yt6
All .Cohort_&yt7
All .Cohort_&yt8
All .Cohort_&yt9
All .Cohort_&yt10
All .Cohort_&yt11
All .Cohort_&yt12
All .Cohort_&yt13
All .Cohort_&yt14
All .Cohort_&yt15
All .Cohort_&yt16
All .Cohort_&yt17
All .Cohort_&yt18
;
proc sort data=DEVGroup&yt01; by id;
run;


/*Begin Processing Developmental Completions*/

Data Course2;
merge  course (in=a)   devgroup&yt01(in=b keep=id MathPlacement EngPlacement
); by id;;
format completedevmath $2. completedevenglish $2.;
completedevmath=' ';
completedevenglish=' ';

if a;
if mathplacement='N' then do;
CompleteDevMath='D';

If course in ('MTH99','MATH904') and grade in ('P')
then CompleteDevMath='C';
end;
if mathplacement='C' then do;
CompleteDevMath='NA';
end;


if engplacement='N' then do;
CompleteDevEnglish='D';

If course in ('ENG99', 'ENGL900') and grade in ('P')
then CompleteDevEnglish='C';
```

```
        end;
        if engplacement='C' then do;
        CompleteDevEnglish='NA';
        end;

    run;


    Data devmath (keep=id completedevmath) deveng (keep=id completedevenglish);
    set course2;
    run;

    proc sort data=devmath; by completedevmath;
    proc sort data=devmath nodupkey; by id;
    proc sort data=deveng; by completedevenglish;
    proc sort data=deveng nodupkey; by id;
    run;

    Data CombDev;
    merge deveng devmath; by id;
    run;

    /*Finish Creating Student Completion Attribute*/

    Data course3;
    merge course2 (in=a drop=completedevenglish completedevmath)  combdev ; by
    id;
    if a;
    if CompleteDevEnglish=' ' then CompleteDevEnglish='NA';
    if CompleteDevMath=' ' then CompleteDevMath='NA';
    run;
```