



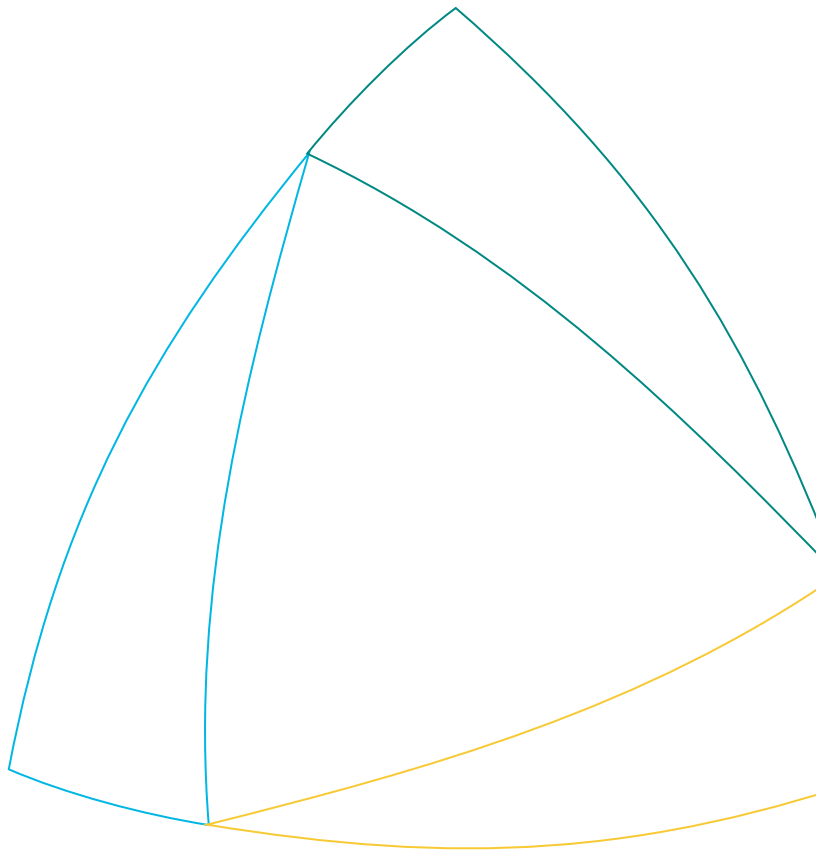
SHEEO

STATE HIGHER EDUCATION EXECUTIVE OFFICERS ASSOCIATION

POSTSECONDARY DATA PARTNERSHIP (PDP) TECHNICAL BRIEF:

USING CROSSWALKS TO FACILITATE DATA
SUBMISSIONS FOR GATEWAY COURSE
REQUIREMENTS AND COURSE CLASSIFICATION
OF INSTRUCTIONAL PROGRAM (CIP) CODES

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.

This paper is based on research funded in part by the Bill & Melinda Gates Foundation. The findings and conclusions contained within are those of the author(s) and do not necessarily reflect positions or policies of the Bill & Melinda Gates Foundation.

TABLE OF CONTENTS

| | |
|--|----|
| Introduction..... | 4 |
| Using Crosswalks to Facilitate Data Submissions | 5 |
| Cohort File Crosswalk for Gateway Course Requirements..... | 5 |
| Course File Crosswalk for Classification of Instructional Program (CIP) Codes | 6 |
| Results..... | 7 |
| Recommendations..... | 7 |
| Appendix A: Sample Gateway Course Crosswalk | 9 |
| Appendix B: Sample SAS Code to Evaluate Gateway Requirements..... | 10 |
| Appendix C: Sample Course CIP-Code Crosswalk..... | 12 |
| Appendix D: Sample SAS Code to Merge Course CIP-Code Information to Course File..... | 13 |

INTRODUCTION

This document is part of SHEEO's efforts to support the higher education community with technical assistance¹ 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 explain an approach to prepare PDP data collections that includes the development of term cohort files and course files utilizing crosswalks for 1) gateway course requirements, and 2) course Classification of Instructional Program (CIP) codes. The process was developed by a Central Piedmont Community College's Center for Applied Research (CFAR)³ analyst to capture information from four PDP participating institutions in North Carolina. The institutions share a common course library⁴ (Combined Course Library) and Colleague student information system through the North Carolina Community College System (NCCCS). Crosswalks were developed to bridge the local information found in the Colleague system to the NCCCS information. The data preparation approach leveraged institutional crosswalks 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. North Carolina Community College System. Combined Course Library
<https://www.ncccommunitycolleges.edu/academic-programs/combined-course-library>

USING CROSSWALKS TO FACILITATE DATA SUBMISSIONS

The Postsecondary Data Partnership requires two types of submission files. The first is a single-record student cohort file that consists of all new students at the institution. The second is a multiple-record course file that consists of students' course records. To facilitate the preparation of cohort and course files, the Clearinghouse provides documentation for file formatting and field length requirements.⁵

Beginning in the fall of 2017, the CFAR analyst began working with Frontier Set colleges in North Carolina to prepare PDP files for three Frontier Set institutions: Wake Technical Community College, Guilford Technical Community College, and Davidson County Community College. Central Piedmont Community College joined the PDP data collection in the summer of 2018. Through a previous Bill & Melinda Gates-funded partnership, known as Completion by Design (CBD),⁶ the CFAR analyst shared SAS software programs to create common SAS data sets. The participating community colleges shared the same Colleague student information system, data-mart structure, and files. Utilizing SAS software, the PDP data fields were extracted from the common data sets using a set of SAS software programs specific to the PDP requirements.

COHORT FILE CROSSWALK FOR GATEWAY COURSE REQUIREMENTS

The PDP cohort submission files include new students in each term being reported. While the SAS software programs developed for use by the colleges in North Carolina were able to capture new students in each term being reported, not all the information was captured in the process. Each institution was required to provide gateway course requirements for both math and English. Since institutions map and align gateway course requirements to specific programs, catalog years, and degree type, a gateway course crosswalk was developed to capture the local information (see *Appendix A: Sample Gateway Course Crosswalk*). The institutional researcher at each institution completed the crosswalk with the required data fields by manually adding the information. In some instances, an Informer report generated the gateway course information. The crosswalks were compiled in an Excel spreadsheet, and SAS software was used to evaluate the gateway course requirements for each institution and to code them appropriately (see *Appendix B: Sample SAS Code to Evaluate Gateway Requirements*).

To conduct data quality checks, the CFAR analyst added validation checks to ensure gateway course requirements were captured. The cohort files were processed in the SAS environment and scanned for missing data fields. Data fields that were not populated in the institution's files were flagged, and error reports were generated to notify the institutional researcher of a problem with the crosswalk. Corrections and or program additions were subsequently handled in the crosswalks and processed using SAS software coding. The process required collaboration with the institutional researcher to ensure that the fields were accurately coded. The cohort files were processed in the SAS environment using appropriate field formatting and prepared for uploading to the Clearinghouse. The cohort files were saved for future processing of the course files.

5. National Student Clearinghouse. *Postsecondary Data Partnership data submission guide (2019)*. <https://studentclearinghouse.info/onestop/wp-content/uploads/PDP-Submission-Guide.pdf>

6. Bill & Melinda Gates Foundation. Completion by Design. <https://www.completionbydesign.org/s>

Two challenges to creating and updating gateway crosswalks are:

- Gateway course requirements are specific to local institutions, and gateway crosswalks need to be quality checked and validated by a cross-functional team.
- Determining gateway course requirements for transfer-in students is challenging without collecting specific transfer-in credits. Students may be flagged with gateway course requirements that were completed prior to transfer.

COURSE FILE CROSSWALK FOR CLASSIFICATION OF INSTRUCTIONAL PROGRAM (CIP) CODES

Similar to the PDP cohort submission file, the PDP course submission files also include new students in each term being reported, and not all the information was initially captured in the SAS software programs developed for use by the colleges in North Carolina. Information that could not be located in the student information system was left blank. If the information was required by the Clearinghouse, missing information created an error when files were uploaded to the Clearinghouse. File preparers at your institution will need to adhere to field requirements and ensure that required fields are provided for each student in the file.

The course file is the ‘meat’ of the longitudinal reporting function of the PDP data and includes course grades, 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.

For the North Carolina colleges, a crosswalk was developed to code course CIP codes that leveraged the National Center for Education Statistics (NCES) CIP code classifications for instructional programs⁷ and connected the information to the NCCCS coding in the Combined Course Library (see *Appendix C: Sample Course CIP-Code Crosswalk*). The Course CIP-Codes Crosswalk leverages NCCCS course prefix/subject area and course number, which ensures consistent reporting for the institutions.

The CFAR analyst added validation checks to ensure that CIP codes were not missing for any courses, which would lead to errors in the Clearinghouse submission process and require corrective action. To prevent uploading files with errors, courses with ‘missing courses report’ were processed in the SAS environment, and courses with missing CIP codes were displayed in an Excel spreadsheet (see *Appendix D: Sample SAS Code to Merge Course CIP-Code Information to Course File*). Corrections to course CIP codes and/or course additions were subsequently handled in the crosswalk and processed using SAS software coding. The CFAR analyst maintained the master course CIP-code crosswalk, and when course CIP codes were added or changed, the master crosswalk was updated and sent to the reporting institutions.

7. U.S. Department of Education National Center for Education Statistics. 2020. List of Classifications of Instructional Programs (CIP). <https://nces.ed.gov/ipeds/cipcode>

Two challenges to creating and updating CIP-code crosswalks are:

- NCES CIP codes are associated with classifications of instructional programs. The course CIP-code mapping to individual courses reflects the courses associated with the instructional program that best describes program content. In general, program credentials that are awarded at different levels have the same CIP code, but there are some exceptions.
- CIP codes are updated by NCES every ten years. Therefore, CIP codes will need to be updated and maintained in the crosswalks, which include system/source documentation.

RESULTS

The results of using the data collection process outlined here demonstrate 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. Students were reported in the cohort files in a single-record instance along with identifying information. The gateway course information was captured in the gateway crosswalk for each institution and translated in the cohort files. Students' gateway course information was reported once in a cohort file.

Course CIP codes were captured during the course file preparation using course CIP-code crosswalks. Since course files include the reporting term and term outcomes, the files facilitate longitudinal cohort tracking and metrics measured at points in time along the student's journey. To ensure data integrity, CIP codes were built into the SAS coding software process and also prevented missing CIP-code errors in the Clearinghouse reports.

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 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.
- The multiple-record course file adheres to reporting information that is associated with the course(s) in 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.
- Gateway crosswalks can be used to code gateway course requirements in the cohort files and course CIP codes in the course files. While crosswalks provide efficiency in the process, the crosswalks require maintenance and validation checks. Master crosswalks can accommodate consistent reporting of information for single/multiple institutions and systems.

- While the process can be replicated for future data collections, coding changes in the student information system or at the system level will need to be addressed in the ongoing query extracts. Customizations to source data systems can lead to access and data interpretation issues. The process developed in North Carolina evolved and improved over time as missing data fields and quality checks were handled using crosswalks.
- The crosswalk processes developed at CFAR provide potential multi-institutional and system solutions for handling the PDP data collections and increasing participation across the ATD network of colleges and beyond.

APPENDIX A: SAMPLE GATEWAY COURSE CROSSWALK

| College | CatalogYear | English Gateway | Math Gateway | ADCSTuCurPri ProgCode | Title |
|---------|-------------|-----------------|--------------|-----------------------|-----------|
| _GTCC | 2017-18 | 1 | 1 | A10100 | Associate |
| _GTCC | 2018-19 | 1 | 1 | A10100 | Associate |
| _GTCC | 2017-18 | 1 | 1 | A10100AE | Associate |
| _GTCC | 2018-19 | 1 | 1 | A10100AE | Associate |
| _GTCC | 2018-19 | 1 | 1 | A10100C1 | Associate |
| _GTCC | 2018-19 | 1 | 1 | A10100E1 | Associate |
| _GTCC | 2018-19 | 1 | 1 | A10100H1 | Associate |
| _GTCC | 2018-19 | 1 | 1 | A10100P1 | Associate |
| _GTCC | 2018-19 | 1 | 1 | A10100P2 | Associate |
| _GTCC | 2018-19 | 1 | 1 | A10100P3 | Associate |
| _GTCC | 2017-18 | 1 | 1 | A10100PS | Associate |
| _GTCC | 2018-19 | 1 | 1 | A10100PS | Associate |
| _GTCC | 2018-19 | 1 | 1 | A10100R1 | Associate |
| _GTCC | 2018-19 | 1 | 1 | A10100S1 | Associate |
| _GTCC | 2018-19 | 1 | 1 | A10100S2 | Associate |
| _GTCC | 2017-18 | 1 | 1 | A1010A | Associate |
| _GTCC | 2018-19 | 1 | 1 | A1010A | Associate |
| _GTCC | 2017-18 | 1 | 1 | A1010B | Associate |
| _GTCC | 2018-19 | 1 | 1 | A1010B | Associate |
| _GTCC | 2017-18 | 1 | 1 | A1010C | Associate |
| _GTCC | 2018-19 | 1 | 1 | A1010C | Associate |
| _GTCC | 2017-18 | 1 | 1 | A1010D | Associate |
| _GTCC | 2018-19 | 1 | 1 | A1010D | Associate |
| _GTCC | 2017-18 | 1 | 1 | A1010E | Associate |
| _GTCC | 2018-19 | 1 | 1 | A1010E | Associate |

APPENDIX B: SAMPLE SAS CODE TO EVALUATE GATEWAY REQUIREMENTS⁸

```

OPTIONS PS=5000 LS=256 NONUMBER NOCENTER;

%LET PathRoot=P:\CPCCSASDatasets\A_XNC_Datatel; /*Change to your Drive

LIBNAME CBDUP "&PathRoot\Downloads\SASDatasets\12_CBD\Uploads";
LIBNAME TRANSFER "&PathRoot\Downloads\SASDatasets\04_TransferCredits";
LIBNAME ALL "&PathRoot\Downloads\SASDatasets\00_All";
LIBNAME CBD "&PathRoot\Downloads\SASDatasets\12_CBD";
LIBNAME STUDENT "&PathRoot\Downloads\SASDatasets\02_Student";
LIBNAME ISIR "&PathRoot\Downloads\SASDatasets\06_FinAidISIR";

%LET In01=_GTCC; /*Change Institutional ID*/
%LET rp02=_ET; /*End of Term Data Point*/
%LET yt01=201902; /*Change for each academic term*/

%LET GW01=Gateway_crosswalk.csv; /*Enter the Name of Gateway CSV Crosswalk*/
%LET CatalogYear=2017-18; /*Enter the Catalog Year in this format XXXX-XX if
applicable-Enter the word None if Catalog Year not-Applicable*/

/*Add Gateway Mapping Program Code Dataset*/

data WORK.Gateways ;
%let _EFIERR_ = 0; /* set the ERROR detection macro variable */
infile "&PathRoot\Downloads\SASDatasets\99_Crosswalks\&GW01" delimiter =
',' MISOVER DSD lrecl=32767 firstobs=2 ;
informat College $5. ;
informat CatalogYear $20.;
informat EnglishGateway best32. ;
informat MathGateway best32. ;
informat ADCSTuCurPriProgCode $12. ;
informat Title $98. ;
informat PreviousCipCode best32. ;
informat CurrentCipCode best32. ;
format College $5. ;
format CatalogYear $20.;
format EnglishGateway best12. ;
format MathGateway best12. ;
format ADCSTuCurPriProgCode $12. ;
format Title $98. ;
format PreviousCipCode best12. ;
format CurrentCipCode best12. ;
input
College $
CatalogYear $
EnglishGateway

```

8. 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>

```

        MathGateway
        ADCStuCurPriProgCode $
        Title $
        PreviousCipCode
        CurrentCipCode
    ;
    if _ERROR_ then call symputx('_EFIERR_',1); /* set ERROR detection macro
variable */

    run;

/*Select the College and Catalog year if Applicable*/
Data gateways2;
set gateways;
if College EQ "&IN01" and CatalogYear EQ "&CatalogYear";
run;
proc sort data=static11&yt01; by ADCStuCurPriProgCode;
proc sort data=gateways2 nodupkey; by ADCStuCurPriProgCode;

run;

/*Add Gateway Variables*/

Data static11a&yt01;
merge static11&yt01 (in=a) gateways2; by ADCStuCurPriProgCode;
if a;
format gatewaymath $2. gatewayenglish $2.;
GatewayMathStatus=MathGateway;
if MathExam=-3 then do; /*Transfer already completed requirement*/
if MathGateway=1 then do;
GatewayMathStatus=0;
end;
end;

If GatewayMathStatus=. then GatewayMathStatus=-1;

GatewayEnglishStatus=EnglishGateway;
if EngExam=-3 then do; /*Transfer already completed requirement*/
if EnglishGateway=1 then do;
GatewayEnglishStatus=0;
end;
end;

If GatewayEnglishStatus=. then GatewayEnglishStatus=-1;

if GatewayMathStatus=1 then GatewayMath='R';
else if GatewayMathStatus=0 then GatewayMath='N';
else GatewayMath='UK';

if GatewayEnglishStatus=1 then GatewayEnglish='R';
else if GatewayEnglishStatus=0 then GatewayEnglish='N';
else GatewayEnglish='UK';

proc sort data=static11a&yt01 nodupkey; by adcpersonid;
run;

```

APPENDIX C: SAMPLE COURSE CIP-CODE CROSSWALK

| CourseCip | StateNumber | StateTitle |
|-----------|-------------|---------------------------|
| 51.2311 | AAI-110 | Animal Assisted Intro |
| 51.2311 | AAI-120 | Animals in Human Lives |
| 51.2311 | AAI-130 | Animal Handling Skills |
| 51.2311 | AAI-210 | Interaction Methodology |
| 51.2311 | AAI-220 | Interaction Documentation |
| 32.0101 | ABE-01112 | ABE/GED Program |
| 37.0101 | ACA-085 | Improving Study Skills |
| 37.0101 | ACA-090 | Study Skills |
| 37.0101 | ACA-111 | College Student Success |
| 37.0101 | ACA-112 | Intro to Dist Learning |

APPENDIX D: SAMPLE SAS CODE TO MERGE COURSE CIP-CODE INFORMATION TO COURSE FILE

```
proc sort data = course1&yt01&rp02; by StateNumber;
run;
proc sort data=CBD .NCCCS_crse_cipcodes_mod20190214 nodupkey; by
StateNumber;
run;

data course2&yt01&rp02 csun&yt01&rp02;
merge course1&yt01&rp02 (in=a) CBD .NCCCS_crse_cipcodes_mod20190214 (in=b
); by StateNumber;
if a and b then output course2&yt01&rp02;
else if a and not b then output csun&yt01&rp02;
run;

data course3&yt01&rp02;
set course2&yt01&rp02 Csun&yt01&rp02

;

run;

data missingcodes;
set csun&yt01&rp02
;
Title1 'Check for courses with missing cip codes';
proc sort nodupkey; by CourseNumber;
run;

ods html file='test.html';

proc print data=missingcodes;
Title1 'Check for courses with missing cip codes';
run;

ods html close;
```

STATE HIGHER EDUCATION EXECUTIVE OFFICERS

3035 CENTER GREEN DRIVE, SUITE 100, BOULDER, COLORADO, 80301
303.541.1600 • SHEEO.org

