![Institute of Education Sciences logo](IES Institute of Education Sciences)
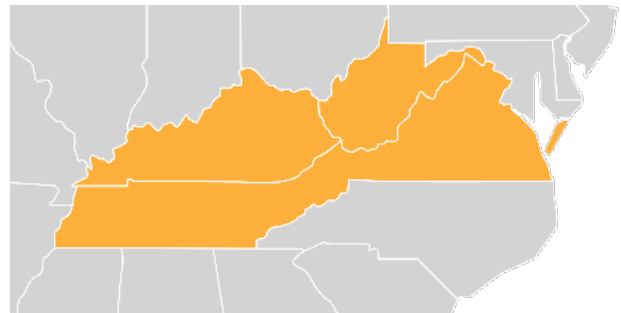
Project 5.2.18: "How-to" Analysis Memo: A Review of High School Mathematics Coursetaking Pathways Among Students Who Completed Algebra I in Grade 7

Regional Educational Laboratory Appalachia at SRI International

August 2021

Plus Alpha Research & Consulting, LLC
Ryoko Yamaguchi, Research Lead

SRI International
Rebecca Schmidt, Research Co-lead

This page intentionally left blank

Contents

Exhibits

Figures

"How-to" Analysis Memo: A Review of High School Mathematics Coursetaking Pathways Among Students Who Completed Algebra I in Grade 7

Overview

The "How-to" analysis memo shows how to create an analytic dataset and conduct descriptive analyses to review high school mathematics coursetaking pathways among students who completed Algebra I in grade 7 in Virginia. The purpose of the memo is to help Virginia database managers and analysts answer two main questions about mathematics coursetaking pathways and proficiency levels on state assessments:

1. For students who completed Algebra I in grade 7, when did they complete their highest level mathematics course? Did the grade vary by student subgroup (students facing socioeconomic barriers or multilingual learners), prior mathematics performance in grade 5 (Advanced proficient, Proficient, or Below proficient), or by diploma type (Standard or Advanced Studies diploma)?

2. For students who completed Algebra I in grade 7, how did they perform on the Geometry and Algebra II assessments? Did performance vary by student subgroup (students facing socioeconomic barriers or multilingual learners), prior mathematics performance in grade 5 (Advanced proficient, Proficient, or Below proficient), or by diploma type (Standard or Advanced Studies diploma)?

With this information, database managers and analysts can support school division leaders and mathematics content experts to use data to inform and improve decisions about policy and practice related to mathematics coursetaking pathways.

Intended users of this memo

This memo is intended for Virginia school division database managers and analysts. Database managers and analysts bring expertise in school and student information data systems, including the configuration, implementation, and operation of those systems, and they are often asked to produce reports for school division leaders on student progress, growth, and outcomes. This memo helps the database managers and analysts produce data reports on mathematics coursetaking and proficiency levels by outlining the steps to retrieve data from the school and student information systems and analyze data in order to graph results.

The knowledge and skills needed to follow the steps outlined here include the ability to retrieve data from school and student information data systems, merge multiple student files into a single dataset using the student testing identifier, recode variables, create new variables based on existing ones, identify duplicate cases in data files, select a subset of cases from a data file, and identify valid records based on the directions given here. This memo does not reference specific software to accomplish these tasks; users may employ their preferred software program. However, users should have basic knowledge of a spreadsheet program such as Microsoft Excel to create graphs. Database managers and analysts can also choose to adapt and refine the steps described in this memo to conduct additional analyses.

This memo was developed as part of the Regional Educational Laboratory (REL) Appalachia Student Success in Mathematics partnership coaching project. The REL Appalachia team acknowledges and thanks the Virginia Department of Education and the Virginia school divisions that took part in this collaborative project. Math specialists and database managers reviewed all higher-level mathematics courses offered in Virginia and helped to sort the mathematics courses into levels; provided feedback on each step in the memo; suggested the addition of an Excel template for documenting results; and tested the steps and the template using data from their school divisions. This memo offers tips generated from database managers who helped to pilot the memo and associated Excel template.

Focus of the analysis

While research shows that taking Algebra I before grade 9 is associated with improved college readiness and high school graduation rates (Allensworth & Easton, 2005, 2007; Matthews & Farmer, 2008; Watson et al., 2010) and higher postsecondary enrollment (LaFave, 2019), much of this research focuses on completing Algebra I in grade 8. There is limited research on students who complete Algebra I in grade 7.

In a study of Algebra I and college preparatory diploma outcomes among Virginia students, researchers found that among students who completed Algebra I in grade 7, students facing socioeconomic barriers and multilingual learners had lower rates of passing Algebra I and graduating with an Advanced Studies diploma than all students (Yamaguchi et al., 2020). Recognizing that about a quarter of Virginia students (24 percent) complete Algebra I in grade 7 (Yamaguchi et al., 2020), the study results prompted Virginia leaders to seek a better understanding of high school coursetaking and academic performance for students who complete Algebra I in grade 7, referred to as hyper-accelerated students (Galanti, 2019).

The focus of the analysis in this memo is on students who completed Algebra I in grade 7. The analysis focuses on a single cohort of students who were in grade 5 in the 2009/10 school year, completed Algebra I in grade 7 in the 2011/12 school year, and graduated high school in the 2016/17 school year. Multiple years of data are needed to look at this single cohort of students from the time they were in grade 5 through high school graduation (eight total years of data). Among this single cohort of hyper-accelerated students, this analysis memo presents how to answer two main questions:

- **Question 1: For students who completed Algebra I in grade 7, when did they complete their highest level mathematics course?**
  The data and results from this question can help Virginia school division leaders and mathematics content experts think about whether high school graduates are underprepared for college-level mathematics despite an early accelerated start in mathematics. For example, a student who completes Algebra I in grade 7 can complete Geometry in grade 8 and Algebra II in grade 9, thus accruing the mathematics credits necessary to graduate with a Standard diploma without completing any mathematics

courses in grade 10, 11, or 12. Likewise, a similar student could go on to complete Trigonometry in grade 10. This student could graduate with an Advanced Studies diploma without completing any mathematics courses in grade 11 or 12. Taking mathematics throughout high school can prepare students for college-level mathematics courses and can support skills and development of various career clusters (Virginia Department of Education Virginia Mathematics Pathways Initiative, n.d.). The analysis can demonstrate whether students are well-prepared for college-level mathematics courses by continuing to complete higher level mathematics courses throughout high school.

- **Question 2: For students who completed Algebra I in grade 7, how did they perform on the Geometry and Algebra II assessments?**

  The data and results from this question can help Virginia school division leaders and mathematics content experts understand whether hyper-accelerated students are mastering mathematics content as expected. Geometry and Algebra II are prerequisites for advanced mathematics courses such as Calculus. Hyper-acceleration may result in students being under-prepared for higher-level mathematics if bypassing the courses traditionally taught in grades 7 and 8 prevents students from fully mastering the fundamentals of algebraic thinking (Galanti, 2019; Stein et al., 2011; Yamaguchi et al., 2020). The analysis can demonstrate whether students are well-prepared for advanced high school mathematics courses by demonstrating proficiency in Geometry and Algebra II.

While this analysis memo focuses on a single cohort of students who completed Algebra I in grade 7 in the 2011/12 school year and graduated from a Virginia high school in the 2016/17 school year, the analyses can be replicated with more recent cohorts of high school graduates, as well as with students who completed Algebra I in grade 8 or other grades. These additional analyses are important for helping school division leaders and mathematics content experts compare outcomes of students in different coursetaking pathways, and to monitor changes over time. A brief description of how to replicate the analyses for other cohorts of students is in the section on Implications for additional analysis.

Objective of this memo

The objective of this memo is to provide database managers and analysts with a process for analyzing their data to identify the highest-level mathematics courses and the proficiency levels in subsequent mathematics courses among students who completed Algebra I in grade 7. The memo provides step-by-step instructions and pointers, as well as extension activities to adapt the analyses to include different cohorts of students, different subgroups of students, and different grade levels of Algebra I completion. It describes the steps database managers and analysts in each school division need to take to:

- Identify the population of students to examine.

- Establish the baseline mathematics performance for each student.

- Define the highest-level mathematics course for each student.

- Include Geometry and Algebra II state assessment scaled scores and state assessment proficiency levels in the data descriptions.

- Address missing or conflicting values and other data-cleaning concerns.

- Document decisions made during the data-cleaning and management process, and the implications for results.

The memo has two sections that describe how to create the analytic dataset and how to use the dataset to answer the two primary questions. The first section includes seven steps to create the analytic dataset that includes all the necessary data on students' coursetaking pathways (figure 1). The first step of Section I, represented by the "Student Inputs" box in figure 1, focuses on identifying students in grade 5 in the 2009/10 school year and documenting student demographic and assessment data. The second step, represented by the "Algebra I" box, focuses on creating categories for Algebra I completion and creating a dataset of students who completed Algebra I in grade 7. The third step, represented by the "High school diploma" box, focuses on adding graduation data. Steps 4 through 7, represented by the dashed "Hyper-accelerated students" box, focus on adding Geometry and Algebra II assessment scores, adding high school mathematics courses, and coding mathematics courses into course levels.

Figure 1. Data and analysis overview

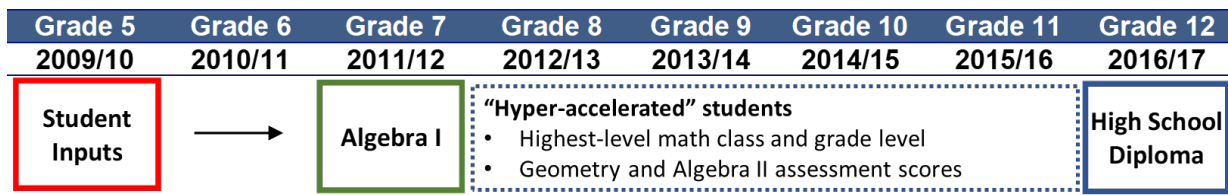| Grade 5 2009/10	Grade 6 2010/11	Grade 7 2011/12	Grade 8 2012/13	Grade 9 2013/14	Grade 10 2014/15	Grade 11 2015/16	Grade 12 2016/17
Student Inputs	→	Algebra I	"Hyper-accelerated" students • Highest-level math class and grade level • Geometry and Algebra II assessment scores				High School Diploma

Figure note: Student inputs in grade 5 include mathematics state assessment scores, English learner indicator, and economically disadvantaged indicator. High school diploma in grade 12 includes the Standard and Advanced Studies diploma.

Section II focuses on using the analytic dataset created in Section I to answer the two main questions. Section II includes two steps to answer the questions and a third step to chart the results. Section II provides table shells database managers and analysts can use to summarize the highest-level mathematics courses completed and proficiency levels on Geometry and Algebra II assessments among students who completed Algebra I in grade 7, disaggregated by student subgroups.

Box 1 defines the key terms of the Virginia educational context that are used throughout the memo.

Box 1. Key terms

Advanced proficient level. The Virginia state assessment has three proficiency levels. Scores of 500–600 are classified as Advanced proficient (Virginia Department of Education, n.d. a).

Advanced Studies diploma. An Advanced Studies diploma is one of two types of diplomas that is available for all Virginia public high school students to earn. It is considered a rigorous college preparatory high school diploma (Holian & Mokher, 2011; Jonas et al., 2012; Jonas et al., 2014) and requires four credits each of English, mathematics, science, and history; three credits of a world language; and one credit of fine arts or career and technical education (Virginia Department of Education, n.d. b).

Below proficient level. The Virginia state assessment has three proficiency levels. Scores of 0–399 are classified as Below proficient and are considered failing scores (Virginia Department of Education, n.d. a).

Coursetaking pathways. For the purposes of this memo, student coursetaking pathways are the series of courses that students take between Algebra I and high school graduation. In Virginia, students may take different mathematics courses in different grade levels. Not all students complete the same courses in the same grade level by the time they graduate.

Course levels. For this analysis, high school mathematics courses offered in Virginia were categorized into course levels, where courses after Algebra I were coded from level 1 (Geometry) through 7 (Multivariate and Differential Calculus). The course levels reflect an increase in difficulty or complexity, based on research (Burkham & Lee, 2003) and discussions with mathematics experts from Virginia state and local education agencies. The full list of high school mathematics courses sorted into course levels is shown in Table 2.

Economically disadvantaged indicator. In Virginia statewide data, students facing socioeconomic barriers are flagged with an "economically disadvantaged" indicator. These students met any of the following criteria in grade 5: was eligible for the National School Lunch Program (NSLP), received Temporary Assistance for Needy Families (TANF), was eligible for Medicaid, was identified as a migrant student, or experienced homelessness at any point during the school year (Virginia Department of Education, 2020). The memo uses the language from the Virginia statewide data and refers to the "economically disadvantaged" indicator.

English learner indicator. In Virginia statewide data, multilingual learners are flagged with an "English learner" indicator. These students met any of the following criteria in grade 5: received English learner services, was eligible to receive English learner services but did not, or was eligible for English learner services but completed the English learner program within the previous two years (Virginia Department of Education, 2020). The memo uses the language from the Virginia statewide data and refers to the "English learner" indicator.

Multilingual learner. For the purposes of this memo, a multilingual learner is defined as a student with a "yes" on the English learner indicator in grade 5. We use the term "multilingual learner" in this memo because it recognizes the strengths of those students, rather than their potential deficits. Thus "English learner" is an indicator in the data system, while "multilingual learner" is a student who was designated with that indicator in grade 5.

PowerSchool. A data storage platform commonly used in Virginia school divisions.

Proficient level. The Virginia state assessment has three proficiency levels. Scores of 400–499 are classified as Proficient (Virginia Department of Education, n.d. a).

Standards of Learning state assessment (SOL). Virginia's state assessments, aligned with the state standards, are called the Standards of Learning assessments (SOL). The scaled scores for the assessments range from 0–600. Virginia assesses mathematics in grades 3–8, in addition to administering assessments after students complete Algebra I, Geometry, and Algebra II, regardless of the grade when students complete these courses.

Standard diploma. A Standard diploma is one of two types of diplomas that Virginia public high school students can earn. It requires four credits of English; three credits of mathematics, science, and history; and two credits of a world language, fine arts, or career and technical education (Virginia Department of Education, n.d. c).

Student facing socioeconomic barriers. For the purposes of this memo, a student facing socioeconomic barriers is defined as a student with a "yes" on the economically disadvantaged indicator in grade 5. As with the term "multilingual learners," we use this term to emphasize the students, rather than the barriers they are facing. Thus "economically disadvantaged" is an indicator in the data system, while "student facing socioeconomic barriers" is a student who was designated with that indicator in grade 5.

Valid assessment scores. An assessment scaled score that is within the scale range for that grade (e.g., in Virginia, the range of the state assessment scale for grade 5 mathematics is from 0 through 600).

Section I: Creating the analytic dataset

Using the information in this section, database managers and analysts can learn to create a single analytic dataset with one student per row. This memo uses the example of students who were in grade 5 in 2009/10, completed Algebra I in grade 7 in 2011/12, graduated from high school in 2016/17, and had complete demographic and assessment data, replicating a statewide analysis previously published by REL Appalachia (Yamaguchi et al., 2020). As a reminder, a brief description of how to replicate the analyses for other cohorts of students is in the section on Implications for additional analysis. Conducting additional analyses, such as students who completed Algebra I in grade 8 or graduated more recently, would provide valuable information for comparison purposes.

The step-by-step process below (steps I:1 through I:7) demonstrates how to create a single analytic dataset, starting with merging different data files, including grade 5 data, Algebra I in grade 7 data, high school graduation data, mathematics course data, and Geometry and Algebra II state assessment data.

Step I.1: Identify students in grade 5

The first part of this step is to identify a baseline group of students who were in your school division at a timepoint prior to the start of secondary (middle and high) school, and who have a complete high school graduation record. This memo uses students in grade 5 in the 2009/10 school year (students expected to graduate in the 2016/17 school year).[1]

Create a dataset with each student's grade 5 mathematics state assessment scores and grade 5 demographic information. The outline below presents instructions for creating this dataset.

A. Identify students who took a state mathematics assessment in grade 5 in your division in the 2009/10 school year. In Virginia school divisions, these data will come from Pearson test files. Access to the Pearson test files will depend on your school division; some

---

[1] This memo uses students in grade 5 to represent the baseline because it is an important starting point for the different mathematics coursetaking pathways students move through in secondary school. However, use the grade level that corresponds with your school division's starting point for the secondary mathematics coursetaking pathway. For example, some school divisions place students in a specific mathematics coursetaking pathway after grade 4, in which case the baseline may be grade 4.

divisions will have access through a local student information system, such as PowerSchool, and others may have the data stored locally in a spreadsheet.

1. Include in the dataset the following **required** variables:

    *   **Student Identifier (ID).** Identifier variable that is unique for each student but is **the same across years and across files** (e.g., for student test scores and for student demographics). In the Pearson test files, this would be the student test identifier (the STI variable). **Note: This identifier is necessary for you to link student demographics to the student's test score and course outcomes.**
    *   **School year.** Include only rows from the 2009/10 school year.
    *   **Grade level.** Student grade level in 2009/10. Include only rows from grade 5.
    *   **English learner indicator.** This should be a variable with only two categories. If the variable has more than two categories, convert it to two:
        -   **"Yes" or 1** — Was identified as an English learner in 2009/10 (whether or not the parent accepted services) or was identified as an English learner at some point in the previous two school years.
        -   **"No" or 0** — Was not identified as an English learner in 2009/10 and was not identified as an English learner at any point in the previous two school years. Some datasets may only have "Yes" or blank in this variable. In that case, you should code blank values as "No."
    *   **Economically disadvantaged indicator.** This variable will be used to identify students facing socioeconomic barriers. It should be a variable with only two categories. If the variable has more than two categories, convert it to two:
        -   **"Yes" or 1** — Was identified as "economically disadvantaged" in 2009/10. This may be defined as eligible for free or reduced-price meals, received Temporary Assistance for Needy Families (TANF), eligible for Medicaid, identified as either migrant or experiencing homelessness at any point during the school year, or it may include other family or student economic status factors.
        -   **"No" or 0** — Was not identified as "economically disadvantaged" in 2009/10. Some datasets may only have "Yes" or blank in this variable. In that case, you should code blank values as "No."

2. You may also bring in the following **optional** variables:

    *   **Special education indicator.** This should be converted to a variable with only two categories:
        -   **"Yes" or 1** — Had an Individualized Education Program (IEP) or received any special education services in 2009/10.
        -   **"No" or 0** — Did not have an IEP or receive any special education services in 2009/10. Some datasets may only have "Yes" or blank in this variable. In that case, you should code blank values as "No."

- **Race / ethnicity.** This should be a single categorical variable.[2] If stored as two variables, one for race and one for Hispanic ethnicity, combine the two, so that there are seven race/ethnicity categories. Students should be placed in only one category. For example, African American/Black students are those who are African American/Black and non-Hispanic. Likewise, White students are those who are White and non-Hispanic. On the other hand, Hispanic students are students who are any race and are also Hispanic. Finally, include multi-racial (students who are more than one of African American/Black, Asian/Pacific Islander, and White, but are not Hispanic) as a category:
  - American Indian / Alaska Native (non-Hispanic)
  - Asian/Pacific Islander (non-Hispanic)
  - African American/Black (non-Hispanic)
  - Hispanic (any race)
  - White (non-Hispanic)
  - Native Hawaiian of Pacific Islander (non-Hispanic)
  - Two or more races (non-Hispanic)
- **Gender.** For consistency across divisions and because of restrictions in the data available in 2009/10, define this as two categories: "Male" or "Female."[3]

3. Drop duplicate rows (rows with the same student ID and different values for required variables).

   - If a student has two or more rows with conflicting values for the English learner indicator or economically disadvantaged indicator in grade 5, select the row where the value is 1 or "Yes."
     - For example, if a student appears in one row with a 1 or "Yes" for the English learner indicator and in another row with a 0 or "No" for the English learner indicator, select the row where English learner indicator is 1 or "Yes." Prioritize the 1 or "Yes" indicators to ensure you have included any student who may have been identified as an English learner student or economically disadvantaged student at any time prior to starting secondary school (that is, during elementary school).
   - If a student has two or more rows with conflicting values for race or gender, select the later row (the row that was entered at a later date)
   - After removing these duplicates, you will have one row per student, in which that student's demographic information is consistent within their grade 5 year.

B. Also include in the dataset these students' grade 5 mathematics state assessment scores from the 2009/10 school year from the Pearson test file.

---

[2] The categories used here are those used in Virginia's Longitudinal Data System (VLDS). If you are interested in comparisons by student race, you should identify the categories of interest for your school division and define the variable accordingly.

[3] From school year 2020/21, Virginia Department of Education updated its gender codes to include a gender non-binary category. School divisions analyzing data in 2020/21 and beyond should account for the change.

1. Include in the dataset the following **required** variables:

   - **Student ID.** Identifier variable that is unique for each student but is the same **across years and across files** (e.g., for student test scores and for student demographics). **Note: This identifier is necessary for you to link student demographics to the student's test score and course outcomes.**

   - **School year.** Include only rows from the 2009/10 school year.

   - **Grade level.** Student grade level in 2009/10.

   - **Test code or name.** The name or ID of the test the student took, which you will use to identify mathematics tests. For example, in Virginia statewide data, a test code of "072" (Grade 5 Mathematics 2001 Standards) or "082" (Grade 5 Plain English Mathematics) indicates a grade 5 mathematics state assessment.

   - **Scaled score.** A continuous variable with the student's score on the assessment.

   - **Proficiency level.** A categorical variable with the student's proficiency level on the assessment. In Virginia, the three proficiency levels are "Below proficient," "Proficient," and "Advanced proficient."

2. Keep only one record per student:

   - Keep only math test records. For example, in Virginia statewide data, keep test codes 072 and 082.

   - If a student has more than one test score in mathematics, keep the highest recorded score for grade 5 in 2009/10, and remove any other test scores.

   - Keep only rows from when the student was in grade 5 in 2009/10 and had a grade 5 mathematics assessment scaled score that is within the assessment scale range for that grade (e.g., in Virginia, the range of the state assessment scale of this variable is from 0 through 600). These are referred to as valid scores.

C. Keep only students who:

   1. Were in grade 5 in your division in 2009/10.

   2. Had valid grade 5 mathematics state assessment scores.

   3. Had a value (1 or 0) for the economically disadvantaged indicator.

   4. Had a value (1 or 0) for the English learner indicator.[4]

D. **Ensure that there is only one row per student ID** in your dataset. If there are duplicates (more than one row with the same student ID), review the guidelines above (A.3 and B.2) for removing duplicate rows.

> Tip: Keep track of the sample sizes as you go from all students in grade 5 in the 2009/10 school year to the number of students with valid grade 5 mathematics scores and indicators for economically disadvantaged and English learner. Keeping track of the

---

[4] You may also choose to keep students with incomplete data. If you do, your population sizes will vary across analyses.

sample sizes will help document missing data, define who is in the group of students you are analyzing, and identify how many students you serve who are not included in the analysis. You will document this information in a flow diagram, as shown in figure 3 at the end of step I.

Step I.2: Add Algebra I assessment data

In this step, you will take the group of grade 5 students you identified in Step I.1, add their Algebra I state assessment results, and identify the grade level in which they completed Algebra I. This analysis defined the grade level in which students completed Algebra I as the first time they took the Algebra I state assessment from the Pearson test files, consistent with prior research (Yamaguchi et al., 2020).

A. Create a dataset with students' Algebra I state assessment results:

1. Include in this dataset the following **required** variables:

   - **Student ID.** Identifier variable that is unique for each student but is the same **across years and across files** (e.g., for student test scores and for student demographics). **Note: This identifier is necessary for you to link student demographics to the student's test score and course outcomes.**
   - **School year.** Include rows only from the 2011/12, 2012/13 and 2013/14 school years.
   - **Grade level.** Student grade level in each school year.
   - **Test code or name.** The name or ID of the test the student took, which you will use to identify Algebra I tests. For example, in Virginia statewide data, a test code of 007 indicates an Algebra I state assessment.
   - **Scaled score.** A continuous variable with the student's score on the assessment. In Virginia, the range of this variable is from 0 to 600.
   - **Proficiency level.** A categorical variable with the student's proficiency level on the assessment ("Below proficient," "Proficient," "Advanced proficient").
   - **Test date.**

2. Keep only scores from the Algebra I assessment. In Virginia statewide data, as of 2021, these are state assessment test codes 007 (Algebra I 1995), 020 (Algebra I 2001), 086 (Algebra I Plain English 2001 Standards), 120 (Algebra I 2009), 129 (Algebra I Plain Language 2009 Standards), 420 (VMAST Algebra I 2009 Standards), or 729 (Algebra I 2009). Keep only valid scores, which range from 0 – 600.

3. Identify students who took the Algebra I state assessment more than once.

- Keep only the row with the earliest date they took the Algebra I assessment and had a valid score. If a student took the assessment more than once in the same year, or in different years, use the earliest date they took the assessment. This should correspond to the first time the student took the assessment.

- For example, if a student took the Algebra I assessment in 2011/12 and also in 2012/13, keep the 2011/12 row (even if they failed the Algebra I assessment in 2011/12). In this case, this student "completed Algebra I" in 2011/12.

B. Create a variable that identifies the students when first took the Algebra I state assessment, and fill in the variable with the correct category:

- If students first took the Algebra I assessment in grade 6 in 2010/11, indicate "Grade 6."

- If students first took the Algebra I assessment in grade 7 in 2011/12, indicate "Grade 7."

- If students first took the Algebra I assessment in grade 8 in 2012/13, indicate "Grade 8," etc.

C. Ensure there is only **one row per student** and the variable for Algebra I completion (created in B) is categorical and mutually exclusive (students can fall into only one category).

D. Link the students from step I.1 to their Algebra I state assessment results. To do this, you will need to merge the dataset you created in step I.1 to the dataset with Algebra I state assessment results you just created. You should be able to link the two datasets using the student ID, which is the unique student identifier included in both datasets (only one row per student ID).

E. Once you have merged the two datasets, keep only students who:

1. Were in the dataset in step I.1 (grade 5 students in your division with all the required variables), AND

2. Had valid Algebra I state assessment scores, AND

3. Completed Algebra I in grade 7 in 2011/12. You may also include students who completed Algebra I in other grades, such as grade 8 in 2012/13, if you would like to include them in the analysis.

   - The remainder of this memo focuses on analyzing the outcomes of students who completed Algebra I in grade 7. If you would like to analyze the outcomes of students who completed Algebra I in other grades, please see Implications for additional analyses in the Next Steps section.

Tip: Conduct a quality check on the data cleaning by running a cross-tabulation to ensure that each student is categorized as taking Algebra I in grade 7 or other (unless you coded it for each grade such as Algebra I in grades 6, 8, 9, 10, 11, and 12).

Step I.3: Add graduation data

In this step, you will add graduation data for the students who were in grade 5 in 2009/10 and completed Algebra I in grade 7 in 2011/12.[5] You will examine data for this group of students through their senior year in the 2016/17 school year, to identify whether they graduated on time and what type of diploma they earned.

A. Create a dataset with graduation data as of the 2016/17 school year. In Virginia school divisions, you should use the federal graduation indicator report data file to identify students expected to graduate in four years.

1. Include in the dataset the following **required** variables:

- **Student ID.** Identifier variable that is unique for each student but is the same **across years and across files** (e.g., for student test scores and for student demographics). **Note: This identifier is necessary for you to link student demographics to the student's test score and course outcomes.**
- **School year.** Include rows only from the 2016/17 school year.
- **Diploma type.** A categorical variable indicating the type of diploma a student earned. This should be converted to a variable with only three categories:
  - "College preparatory diploma," which in Virginia includes the Advanced Studies diploma and the International Baccalaureate (IB) diploma.
  - "Standard diploma," which in Virginia includes the Standard diploma.
  - "Other diploma," which in Virginia includes a modified Standard diploma or other type of diploma.

2. Drop duplicate rows.

- If a student has more than one diploma type, keep the "higher" diploma type.
  - For example, if a student has one row with a Standard diploma and one row with an Advanced Studies diploma, keep the row with the Advanced Studies diploma.
  - For example, if a student has one row for a modified Standard diploma (which you will code as "Other") and a Standard diploma, keep the row with the Standard diploma.

B. Ensure there is only **one row per student** and the variable for diploma type is categorical and mutually exclusive (students can fall into only one category).

C. Link the students from step I.2 to their graduation data. To do this, you will need to merge the dataset you created in step I.2 (which is a combination of demographic and baseline data from grade 5 and Algebra I assessment results) to the dataset with graduation data you just created. You should be able to link the two datasets using the

---

[5] If you are interested in comparing students who completed Algebra in grade 7 with students who completed Algebra I in grade 8, you will also add graduation data of students who completed Algebra I in grade 8 in 2012/13 and graduated in 2017/18.

student ID, which is the unique student identifier included in both datasets (only one row per student ID). Once you have merged the two datasets:

1. Run a cross-tabulation of diploma type and "Algebra I grade."

   - Nearly all the students who completed Algebra I in grade 7 should have earned a Standard or college preparatory diploma. You might have a few students who earned other types of diplomas.

> Tip: As a reminder, keep track of the sample sizes as you go from all students in grade 5 in the 2009/10 school year, to the number of students with valid grade 5 mathematics state assessment scores and other indicators, to those with Algebra I data, and finally to those with diploma data. Keeping track of the sample sizes will help document missing data, define who is in the group of students you are analyzing, and identify how many students you serve who are not included in the analysis. You will document this information in a flow diagram, as shown in figure 3 at the end of step I.

Step I.4: Add Geometry and Algebra II EOC assessment data

In this step, you will link your group of students from steps I.1 through I.3 to their state end-of-course (EOC) assessment scores in Geometry and Algebra II from the Pearson test file. Among students who completed Algebra I in grade 7, most would have taken Geometry and the Geometry state end-of-course assessment in grade 8 and progressed to taking Algebra II in grade 9. However, coursetaking pathways may differ, so in this step you will add all Geometry and Algebra II EOC assessment scores.

A. Create a dataset with students' Geometry and Algebra II state assessment / End-of-Course (EOC) assessment scores.

B. Include in the dataset the following **required** variables:

1. **Student ID.** Identifier variable that is unique for each student but is the same **across years and across files** (e.g., for student test scores and for student demographics). **Note: This identifier is necessary for you to link student demographics to the student's score and course outcomes**

2. **School year.** Include rows only from the 2012/13, 2013/14, 2014/15, 2015/16, and 2016/17 school years.

3. **Grade level.** Student grade level in each school year.

4. **Test code or name.** The name or ID of the test the student took, which you will use to identify Geometry and Algebra II tests.

5. **Scaled score.** A continuous variable with the student's score on the assessment. In Virginia, the range of valid scores is 0 to 600.

6. **Proficiency level.** A categorical variable with the student's proficiency level on the assessment (e.g., "Below proficient," "Proficient," "Advanced proficient").

C. Keep only Geometry and Algebra II state assessment scores.

   1. For example, in Virginia statewide data, this is the state assessment test codes 041, 122, 730, 021, 121, 731.

D. Drop duplicate rows.

   1. If a student has two or more rows for the same test in **different years**, keep the earlier test.

   2. If a student has two or more rows with different test scores on the same test in the **same year**, keep the higher test score.

E. **Ensure there is one row per student** with a test score for Geometry and a test score for Algebra II.

F. Link the students from Step I.3 to their Geometry and Algebra II data. To do this, you will need to merge the dataset you created in step I.3 (which is a combination of demographic and baseline data from grade 5, Algebra I assessment results, and graduation data) to the dataset with Geometry and Algebra II scores you just created. You should be able to link the two datasets using the student ID, which is unique in both datasets (only one row per student ID).

G. Once you have merged the two datasets, examine the range of scaled scores in both Geometry and Algebra II to ensure that there are no scores outside of the allowed range. In Virginia, assessment scores are between 0–600. There should be no scores below 0 or above 600. Keep only valid scores (e.g., scores that are between 0–600) and delete the rows with scores outside of the range (e.g., below 0 or above 600).

H. Review the dataset once more. **This dataset will be your "Population" for the analyses.**

   1. **Confirm the dataset includes only one row for each student.** If there is more than one row with the same student ID, check the demographics, test scores, and graduation data, and apply the guidelines above to identify which rows to keep.

   2. **Confirm the dataset includes only students who were in the dataset in step I.3** (grade 5 students in your division with state assessment scores in mathematics, economically disadvantaged indicator, English learner indicator, Algebra I state assessment data, Algebra I grade level, and diploma data) and have Geometry and Algebra II scores.

   3. **Confirm the dataset includes all the required variables.** If your dataset is missing any of these variables, return to the original data sources to locate them and add them to your dataset. The required variables are:

      - Student ID
      - Grade level

- English learner indicator
- Economically disadvantaged indicator
- Grade 5 mathematics scaled score
- Grade 5 mathematics proficiency level
- Algebra I scaled score
- Algebra I proficiency level
- Algebra I grade level
- Diploma type
- Geometry scaled score
- Geometry proficiency level
- Algebra II scaled score
- Algebra II proficiency level

4. You may also keep the following **optional** variables:

- Special education indicator
- Race / ethnicity
- Gender

> Tip: Remember to include variables for both the scaled scores and the state assessment proficiency levels ("Below proficient," "Proficient," or "Advanced proficient"). Check for consistency in the variable values. For example, check the spelling and capitalization for the proficiency level values, and ensure that there is no variation in how the values were entered (i.e., "Below proficient" and "Below Proficient" will be read differently when you get to the analysis stage).

Step I.5: Add high school mathematics courses

In this step, you will add one more type of data to your analysis dataset: the high school mathematics courses students completed.

In steps I.1 through I.4, you created a dataset with one row per student. In this step, you will create a second dataset with one row per student **per year** and eventually reshape and merge the second dataset back to the one you ended with in step I.4.

A. Create a dataset with the mathematics course taken each year after Algebra I from each school year.

1. Include in the dataset the following **required** variables:

- **Student ID.** Identifier variable that is unique for each student but is the same **across years and across all files** (e.g., for student test scores and student course datasets). **Note: This identifier is necessary for you to link student demographics to the student's test score and course outcomes.**

- **School year.** Include rows only from the 2012/13, 2013/14, 2014/15, 2015/16, and 2016/17 school years.

- **Grade level.** Student grade level in each school year.

- **Course name.** Name of the course the student took.

- **Course ID.** Numerical identifier for the course. In Virginia statewide data this is called the "SCED Course Description."

- **Final course grade.** The student's final grade in the course. This can take many forms (e.g., A–F, 0–100, Pass/Fail). You will not use this variable in analysis; you will use it only to verify whether students completed the course and to identify duplicate or multiple rows. Therefore, you can use the data in any form.

2. Include only students who follow the typical on-time grade progression:

- Grade 8 in 2012/13

- Grade 9 in 2013/14

- Grade 10 in 2014/15

- Grade 11 in 2015/16

- Grade 12 in 2016/17

   This excludes students who repeated a grade or did not have data for all five years.

3. Keep only course rows that have a final course grade.

- Drop rows with "withdrawal" and "incomplete" course grades (for example, course grades such as W, WF, WP, NG, I) or missing course grades. These students did not complete the mathematics course in that row, so they should not be included in the counts of students who completed the course.

- Be aware that some students will not have a row for every school year (i.e., if they did not complete a mathematics course in that year or did not have a final grade for their mathematics course).

4. If a student has two or more rows for the same course in the same year with different course grades, keep the higher final course grade value.

B. When you are finished, you should have a dataset with multiple rows per student, one for each mathematics course they took.

> Tip: Students may have multiple rows in the original data because they completed the same course across multiple years. Remember the guideline: if you have multiple rows of data for the same course, take the highest course grade. Your analytic dataset should not have multiple rows for the same high school mathematics course. It should have multiple

> rows only for different high school mathematics courses. For example, if a student took a course in one year and failed it, then took the same course again in a subsequent year and passed, keep the row that corresponds to the year in which they passed the course.

Step I.6: Code high school mathematics courses into course levels

In this step you will create two new variables to categorize the mathematics courses students completed after Algebra I, reshape the dataset to be one row per student again, and link it to the dataset created in step I.4.

A. Using the dataset you created in step I.5 with one row per student per course, categorize each mathematics course into levels. Table 1 shows the levels for each high school mathematics course after Algebra I.

   1. Create a new variable in your dataset called "**Course level**."

   2. For each high school mathematics course in each year, enter in the course level from Table 1. For example, for Geometry, you would enter the number "10" for "Course Level."

   3. Go through each mathematics course to code each course from levels 10 through 71.

      • Note that the "course levels" in table 1 are not continuous. This is because the two digits of the Course Level variable indicate two nested categories.

        - The first digit is a **broad course level**. Courses at level 1 (10, 11, 12, 13, 14, and 15), such as Geometry, are usually taken after Algebra I, while courses at level 2 (20, 21), such as Algebra II, are usually taken after a level 1 course, and so on.

        - Within these broad course levels, the second digit represents course **sublevels**. Students may take sublevel courses within the broad course level in any order.

Table 1. High school mathematics courses after Algebra I

SCED code	Virginia course code	Course name	Course level
2072	3143	Geometry	10
2079	3144	Geometry - Two Year Sequence Part I	11
2071		Informal Geometry	12
2156	3184	Computer Mathematics	13
10019	33186	AP Computer Science Principles	14

SCED code	Virginia course code	Course name	Course level
99001	3134	Algebra, Functions, & Data Analysis	15
2056	3135	Algebra II	20
99002	3136	Grade12 Mathematics Capstone	21
2106	3137	Algebra II & Trigonometry	30
2103	3150	Trigonometry (One Semester)	31
2105	23162	Trigonometry/Mathematics Analysis	32
10157	3185	AP Computer Science A	40
2102	3154	Discrete Mathematics	41
2057	23155	Algebra III	42
2104	3162	Mathematical Analysis/Pre-Calculus	43
2110	23164	Pre-Calculus	44
2201	3190	Probability & Statistics	45
2209	23190	Probability &Statistics - Other	46
2149		Analytic Mathematics - Other	47
2134	Part I (33171) Part II (33172)	IB Further Mathematics SL	48
2203	3192	AP Statistics	49
2073	3176	Analytic Geometry	50
2121	2121	Calculus	51
2111	93165W	Linear Algebra	52
2124	3177	AP Calculus AB	53
2131	33170	IB Mathematics Studies SL	54
2125		AP Calculus BC	60
2132	Part I (33181) Part II (33182)	IB Mathematics HL	61
2122	3178	Multivariate Calculus	70
2123	93180W	Differential Calculus	71

Table note: The course levels were adapted from past research (Burkham & Lee, 2003) and discussions with mathematics leaders from Virginia local and state education agencies. The course names are affiliated with the School Courses for the Exchange of Data (SCED) codes in the U.S. Department of Education SCED Finder database (https://nces.ed.gov/scedfinder). The course name may be different in your school division, including if students took courses virtually (such as through Virtual Virginia) or at one of Virginia's Governor's Schools. AP refers to courses in the Advanced Placement program. IB refers to the courses in the International Baccalaureate program.

B. Reshape your dataset so that you have only one row per student.

1. First, narrow down to one row per student per year:

   - Keep the highest-level mathematics course the student took in that year.
     - For example, if a student took both Linear Algebra (level 52) and AP Calculus BC (level 60) in 2015/16, keep the row with AP Calculus BC (level 60) in 2015/16.

2. Next, turn the data so that the rows for school year become columns (figure 2). The specific command needed to complete this step will vary depending on the software program you are using.

   - When complete, each student should have a column for the highest-level mathematics course they took in each school year.

   - Figure 2 shows an example of what this reshaping looks like before (the image above the arrow) and after reshaping (the image below the arrow)

Figure 2. Sample data reshaping

Student ID	Grade level	School year	Course Name	Level
10000	8	2012	Geometry	10
10000	9	2013	Algebra II	20
10000	10	2014	Trigonometry/Math Analysis	32
10000	11	2015	AP Statistics	49
10000	12	2016	AP Calculus AB	53

Student ID	Grade level 2012	Course name 2012	Level 2012	Grade level 2013	Course name 2013	Level 2013	Grade level 2014	Course name 2014	Level 2014	Grade level 2015	Course name 2015	Level 2015	Grade level 2016	Course name 2016	Level 2016
10000	8	Geometry	10	9	Algebra II	20	10	Trigonometry/ Math Analysis	32	11	AP Statistics	49	12	AP Calculus AB	53

Note: The figure illustrates an example of one student (student 10000). Your resulting data should have one student per row.

C. Identify the highest-level mathematics course each student completed across all school years in high school.

1. In the "Reshaped data" sheet, create a new variable called "**Highest course level.**"

   - Fill in the "Highest-level" variable with the value for the highest-level course each student completed in any school year in high school.
     - For example, the student in figure 2 completed courses at levels 10, 20, 32, 49, and 53; their "highest course level" is a 53.

2. Create a new variable called "**Highest course name**."

   - Fill in the variable with the name of the course that corresponds to the highest course level reflected in the "Highest course level" column.

3. Create a new variable called "**Highest course grade level.**"

   - Fill in the variable with the grade level of the student when they completed their highest-level mathematics course.

D. Link the students from step I.4 to their course data. To do this, you will need to merge the [dataset you created in step I.4](#) (which is a combination of demographic and baseline data from grade 5, Algebra I assessment results, graduation data, and Geometry and Algebra II assessment results) to the course dataset you just created and reshaped. You should be able to link the two datasets using the student ID, which is the unique student identifier in both datasets (only one row per student ID).

> Tip: Students may have multiple rows in the original data because they took different mathematics courses within the same year. Remember the guideline: Take the highest-level mathematics course based on the prioritized list of mathematics courses in Table 1. If the student took the same course more than once (within or across years), take the record in which the student earned the highest course grade for that course.

Step I.7: Finalize the dataset for analysis

In this step, you will finalize your dataset in preparation for analysis.

A. Ensure that your dataset has the following **required** variables:

- Student ID
- English learner indicator in grade 5 (0/1)
- Economically disadvantaged indicator in grade 5 (0/1)
- Grade 5 state mathematics assessment scaled score
- Grade 5 state mathematics assessment proficiency level
- Algebra I state assessment scaled score
- Algebra I state assessment proficiency level
- Diploma type
- Algebra I grade level
- Geometry scaled score
- Geometry proficiency level
- Algebra II scaled score
- Algebra II proficiency level
- Course name 2012/13 (2012/13 school year)
- Course name 2013/14 (2013/14 school year)
- Course name 2014/15 (2014/15 school year)
- Course name 2015/16 (2015/16 school year)
- Course name 2016/17 (2016/17 school year)
- Course level 2012/13

- Course level 2013/14
- Course level 2014/15
- Course level 2015/16
- Course level 2016/17
- Highest course level
- Highest course name
- Highest course grade level

B. Your dataset may also have the following optional variables:

- Special education indicator in grade 5 (0/1)
- Race/ethnicity in grade 5 (categorical)
- Gender in grade 5 (Male/Female)

C. Ensure that your dataset has one row per student. If there are still duplicate or multiple rows, only keep the data that meet the criteria below:

- The highest grade 5 state mathematics assessment score
- The earliest Algebra I state assessment score
- The highest mathematics course taken in each year
- High school mathematics courses as columns for each year, not rows
- The highest course grade within each high school mathematics course

D. Ensure that you have the correct students. The dataset should include students who:

- Were in grade 5 in 2009/10
- Had a grade 5 mathematics score in 2009/10
- Completed Algebra I (first took the Algebra I state assessment) in grade 7 in 2011/12
- Had Algebra I state assessment results in in grade 7 in 2011/12
- Took any mathematics course in grade 8 in 2012/13, grade 9 in 2013/14, grade 10 in 2014/15, grade 11 in 2015/16, or grade 12 in 2016/17
- Had diploma type data in 2016/17

E. Note: If any students in your dataset do not meet all of these criteria, drop them from the dataset.

F. Create a flow diagram of the population dataset.

1. From step I.1 through step I.7, you merged student data from grade 5 through grade 12 and high school graduation. A flow diagram shows the number of students at each step from step I.1 through step I.7 (figure 3).

2. It is important to start at grade 5 and track students forward to create this graphic. Each section of the flow diagram should reflect the count of students remaining in your dataset as you add additional years of data (for example, grade 5 mathematics assessment scores, Algebra I assessment scores).

3. In the flow diagram, you will enter the number of students included and the number of students missing at each step from step I.1 through step I.7.

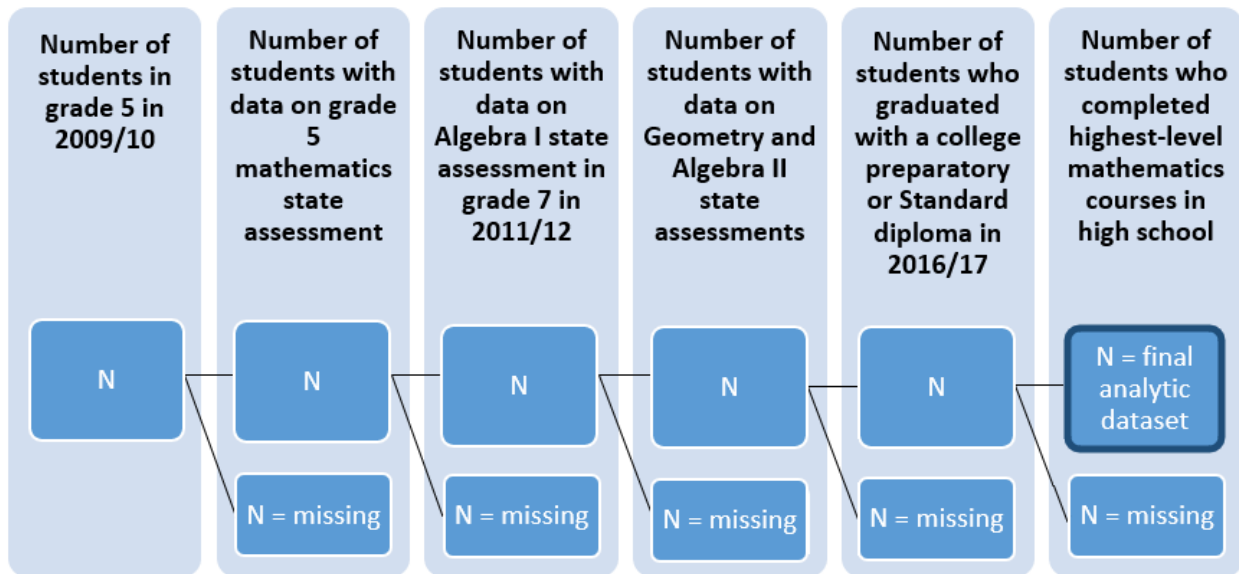**Figure 3. Sample flow diagram for the population dataset**

Figure note: Fill out the number of students in each column with the final count of students who represent the analytic dataset in the bolded box to the right. This type of flow chart diagram is helpful in seeing the representativeness of the student population in the final analytic dataset.

Tip: The flow diagram is a valuable visual because it clearly presents which students are included in your final dataset and which students are missing, to remind you that the results of your analysis may not represent all students in your school division. In displaying the number of students missing from your dataset, you may decide to document and note the reasons for missing data, such as the number of students who left the division or the number of students exempted from taking the state assessment due to their Individualized Education Plans (IEP).

Section II: Analyzing data to answer questions

In this section, the goal is to answer the two main questions using the analytic dataset created in the first section. The accompanying Excel spreadsheet ("RELAP_5.2.18_How-to Analysis Graph Template") will help you graph the results. The step-by-step process below demonstrates how to identify the highest-level mathematics course and the proficiency levels for Geometry and Algebra II assessments among students who completed Algebra I in grade 7.

Step II.1: In what grade level did students who completed Algebra I in grade 7 complete their highest-level mathematics course?

Now that your dataset is complete and clean, you are ready to begin answering questions about the coursetaking and performance of your students. In this step, you will create a table showing the number of students who took their highest-level mathematics course at each grade level. The table shell is provided below. Creating these tables will prepare you to fill in the "Q1 Results" sheet of the accompanying Excel template. The template will then automatically create bar graphs that you can use to visualize the data.

A. Create a table in Excel or Word with the number of students who took their highest-level mathematics course at each grade level.

1. You should create a table similar to the example shown in table 2. This table will show the number (n) of your students who took each course as their highest-level mathematics course high school.

2. First, fill in the cells for each grade level column and sublevel row. These cells represent the number of students who took their highest-level course at each sublevel, in each grade.

   - Each student should appear **only once** in these cells, indicating the grade level in which they completed their highest-level mathematics course. For example, if a student completed Geometry (sublevel 10) in grade 8, Algebra II (sublevel 20) in grade 9, Algebra III (sublevel 42) in grade 10, and AP Calculus AB (sublevel 53) in grade 11, this student would be counted only once in the cell in the grade 11 column and the sublevel 53 row. This student should not appear in four different cells in the table.

3. **Second, fill in the totals for each level.** You should create a summative total row for each broad course level. The "Total Level" rows represent the total number of students who completed their highest-level mathematics course at the broad course level. For example, the "Total Level 1" row is the sum of the students who completed their highest-level mathematics course at sublevel 10, 12, 13, 14, and 15.

4. **Third, fill in the totals by column.** You should also create a summative total column on the far right side of the table. The "Total" column represents the total number of students who, regardless of the grade level, took their highest-level mathematics course in a specific course sublevel. For example, the "n" on the second line of table 2 in the "Total" column represents the total number of students who took their highest-level mathematics course at sublevel 10, Geometry, across all four grade levels.

5. **Fourth, fill in the "Total ALL LEVELS" row.** This row represents the total number of students who took their highest-level mathematics course in each grade, regardless

of broad course levels. These are the numbers you will enter into the accompanying Excel spreadsheet in the "Q1 Results" tab.

Table 2. Frequency count of the number of students taking each course as their highest-level mathematics course and grade level

Level	Course name	Grade 9	Grade 10	Grade 11	Grade 12	Total
10	Geometry	n	n	n	n	n
12	Informal Geometry	n	n	n	n	n
13	Computer Mathematics	n	n	n	n	n
14	AP Computer Science Principles	n	n	n	n	n
15	Algebra, Functions, & Data Analysis	n	n	n	n	n
Total Level 1		**n**	**n**	**n**	**n**	
20	Algebra II	n	n	n	n	n
21	Grade12 Mathematics Capstone	n	n	n	n	n
Total Level 2		**n**	**n**	**n**	**n**	
30	Algebra II & Trigonometry	n	n	n	n	n
Total Level 3		**n**	**n**	**n**	**n**	
40	AP Computer Science A	n	n	n	n	n
41	Discrete Mathematics	n	n	n	n	n
42	Algebra III	n	n	n	n	n
43	Mathematical Analysis/Pre-Calculus	n	n	n	n	n
44	Pre-Calculus	n	n	n	n	n
45	Probability & Statistics	n	n	n	n	n
49	AP Statistics	n	n	n	n	n
Total Level 4		**n**	**n**	**n**	**n**	
50	Analytic Geometry	n	n	n	n	n
51	Calculus	n	n	n	n	n
52	Linear Algebra	n	n	n	n	n
53	AP Calculus AB	n	n	n	n	n

Level	Course name	Grade 9	Grade 10	Grade 11	Grade 12	Total
Total Level 5		**n**	**n**	**n**	**n**	
60	AP Calculus BC	n	n	n	n	n
Total Level 6		**n**	**n**	**n**	**n**	
70	Multivariate Calculus	n	n	n	n	n
71	Differential Calculus	n	n	n	n	n
Total Level 7		**n**	**n**	**n**	**n**	
Total ALL LEVELS		**n**	**n**	**n**	**n**	

Note: In the rows for specific courses, enter the number of students (n). The "Total Level" rows are the summary results for each level where you will enter the total number of students (n) for that course level. In the "Total ALL LEVELS" row, add up all the students who took courses in all levels.

B. Create a similar table to the example in table 2 for the six subgroups of students listed below. For each subgroup, run the analysis only for students in that group. You should have variables in your dataset for each of these subgroups:

- Students with a "Yes" or "1" on the English learner indicator in grade 5 (multilingual learners)
- Students with a "Yes" or "1" on the Economically disadvantaged indicator in grade 5 (students facing socioeconomic barriers)
- Students who scored Advanced proficient on the grade 5 state mathematics assessment
- Students who scored Proficient on the grade 5 state mathematics assessment
- Students who earned a Standard diploma in grade 12
- Students who earned a college preparatory diploma (Advanced Studies diploma) in grade 12

C. Complete the "Q1 Results" sheet of the Excel template

1. Fill in cells C3 through C6 ("Group 1: Number of all students") with the counts from the "Total ALL LEVELS" row of your table that included all students. These cells represent the number of students who completed their highest-level mathematics course in each grade level. For example, the number of students in grade 9 means that these students completed their highest-level mathematics course in grade 9 and did not take any other mathematics course in grades 10 through 12.

2. Fill in cells E3 through E6 ("Group 2: Number of multilingual learners from grade 5") with the counts from the "Total ALL LEVELS" row of your table that includes only students with a "Yes" or "1" on the English learner indicator.

3. Fill in cells G3 through G6 ("Group 3: Number of students facing socioeconomic barriers from grade 5") with the counts from the "Total ALL LEVELS" row of your

table that includes only students with a "Yes" or "1" on the economically disadvantaged indicator.

4. Continue with the remaining subgroups.

5. The raw numbers of students will be automatically converted into percentages and graphed in the Excel template in the tab labeled "Q1 Graphs" (see step II.3).

Step II.2: How did students who completed Algebra I in grade 7 perform on the Geometry and Algebra II assessments?

In this step, you will create descriptive statistics of student end-of-course (EOC) assessment scores for Geometry and Algebra II. You will calculate the average scaled score (in Virginia, valid scores range from 0–600) and the number of students who scored Advanced proficient, Proficient, and Below proficient for the Geometry and Algebra II EOC scores. The table shell is provided below and includes columns for all students and for each subgroup of students. Completing this table will prepare you to fill in the "Q2 Results-Geometry" and "Q2 Results-Algebra II" sheets of the accompanying Excel template. The Excel template will automatically calculate percentages and create bar graphs that you can use to visualize the data.

A. Describe the **Geometry test scores** of your students. A sample shell is shown in Table 3. The purpose of this table is to see how students performed on the end-of-course assessment, regardless of when they completed Geometry.

1. First, calculate the number of students who scored Advanced proficient, Proficient, and Below proficient using the "Geometry proficiency level" variables in your dataset. Enter these values into the corresponding rows of the "All students" column. Enter the sum of the number of students who scored Advanced proficient, Proficient, and Below proficient in the "Total N" row of the "All students" column. Check that the Total N in the table matches the total sample size in your dataset.

2. Second, calculate the average scaled score for Geometry by taking the mean (average) of the "Geometry scaled score" variable in your dataset. Enter this mean value into your table in the "Average Geometry scaled score" row in the "All students" column.

3. Once you filled out the cells for the "All students" column, you will also calculate the number of students who scored Advanced proficient, Proficient, and Below proficient and the average scaled score for the following subgroups of students:

   • Students with a "Yes" or "1" on the English learner indicator in grade 5 (multilingual learners)

   • Students with a "Yes" or "1" on the Economically disadvantaged indicator in grade 5 (students facing socioeconomic barriers)

- Students who scored Proficient on the grade 5 state mathematics assessment
- Students who scored Advanced proficient on the grade 5 state mathematics assessment
- Students who earned a Standard diploma in grade 12
- Students who earned a college preparatory diploma (Advanced Studies diploma) in grade 12

Table 3. Geometry state assessment performance by subgroup

	All students	Multilingual learners in grade 5	Students facing socio-economic barriers in grade 5	Scored Advanced proficient in grade 5 mathematics	Scored Proficient in grade 5 mathematics	Earned a Standard diploma	Earned an Advanced Studies diploma
Number Advanced proficient	n	n	n	n	n	n	n
Number Proficient	n	n	n	n	n	n	n
Number Below proficient	n	n	n	n	n	n	n
Average Geometry scaled score	xxx.x	xxx.x	xxx.x	xxx.x	xxx.x	xxx.x	xxx.x
Total N	n	n	n	n	n	n	n

Note: If this table includes data on students who took Geometry in different grade levels, you may want to add that information to a table note. You may also choose to enter notes in the Excel template.

B. Complete the "Q2 Results-Geometry" sheet of the Excel template by filling in the raw numbers. These raw numbers of students will be automatically converted into percentages and graphed in the Excel template in the tab labeled "Q2 Graphs-Geometry."

- Fill in cells C3 through C5 of the "Group 1: Number of all students" column of the template with the counts from the first column, "All students," and the first three rows of your version of table 3 (Number Advanced proficient, Number proficient, Number Below proficient).
- Fill in cells E3 through E5 of the "Group 2: Number of multilingual learners from grade 5" column of the template with the counts from the second column ("Multilingual learners in grade 5") and the first three rows of your version of table 3 (Number Advanced proficient, Number proficient, Number Below proficient).
- Fill in cells G3 through G5 of the "Group 3: Number of students facing socioeconomic barriers from grade 5" column of the template with the counts from the third column ("Students facing socioeconomic barriers in grade 5") and the first

three rows of your version of table 3 (Number Advanced proficient, Number proficient, Number Below proficient).

- Continue with the remaining subgroups.

C. Describe the **Algebra II end-of-course scores** of your students. A sample shell is shown in table 4. The purpose of this table is to determine how the students performed on the end-of-course assessment, regardless of when they completed Algebra II.

1. First, calculate the number of students who scored Advanced proficient, Proficient, and Below proficient using the "Algebra II proficiency level" variables in your dataset. Enter these values into the corresponding rows of the "All students" column. Enter the sum of the number of students who scored Advanced proficient, Proficient, and Below proficient in the "Total N" row of the "All students" column. Check that the Total N in the table matches the total sample size in your dataset.

2. Second, calculate the average scaled score for Algebra II by taking the mean (average) of the "Algebra II scaled score" variable in your dataset. Enter this mean value into your table in the "Average Algebra II scaled score" row in the "All students" column.

3. Once you filled out the cells for the "All students" column, you will also calculate the number of students who scored Advanced proficient, Proficient, and Below proficient and the average scaled score for the following subgroups of students:

- Students with a "Yes" or "1" on the English learner indicator in grade 5 (multilingual learners)
- Students with a "Yes" or "1" on the Economically disadvantaged indicator in grade 5 (students facing socioeconomic barriers)
- Students who scored Proficient on the grade 5 state mathematics assessment
- Students who scored Advanced proficient on the grade 5 state mathematics assessment
- Students who earned a Standard diploma in grade 12
- Students who earned a college preparatory diploma (Advanced Studies diploma) in grade 12

Table 4. Algebra II state assessment performance by subgroup

	All students	Multilingual learners in grade 5	Students facing socio-economic barriers in grade 5	Scored Advanced proficient in grade 5 mathematics	Scored Proficient in grade 5 mathematics	Earned a Standard diploma	Earned an Advanced Studies diploma
Number Advanced proficient	n	n	n	n	n	n	n
Number Proficient	n	n	n	n	n	n	n
Number Below proficient	n	n	n	n	n	n	n
Average Algebra II scaled score	xxx.x	xxx.x	xxx.x	xxx.x	xxx.x	xxx.x	xxx.x
Total N	n	n	n	n	n	n	n

Note: If this table includes data on students who took Algebra II in different grade levels, you may want to add that information to a table note. You may also choose to enter notes in the Excel template.

D. Complete the "Q2 Results-Algebra II" sheet of the Excel template by filling in the raw numbers. These raw numbers of students will be automatically converted into percentages and graphed in the Excel template in the tab labeled "Q2 Graphs-Algebra II."

1. Fill in cells C3 through C5 of the "Group 1: Number of all students" column of the template with the counts from the first column, "All students," and the first three rows of your version of table 4 (Number Advanced proficient, Number proficient, Number Below proficient).

2. Fill in cells E3 through E5 of the "Group 2: Number of multilingual learners from grade 5" column of the template with the counts from the second column ("Multilingual learners in grade 5") and the first three rows of your version of table 4 (Number Advanced proficient, Number proficient, Number Below proficient).

3. Fill in cells G3 through G5 of the "Group 3: Number of students facing socioeconomic barriers from grade 5" column of the template with the counts from the third column ("Students facing socioeconomic barriers in grade 5") and the first three rows of your version of table 4 (Number Advanced proficient, Number proficient, Number Below proficient).

4. Continue with the remaining subgroups.

Step II.3: Chart the results

In this final step, you will review bar graphs of the results for the two main questions. In the accompanying Excel spreadsheet, a template will generate these graphs for you. Review the graphs in the "Q1 Graphs," "Q2 Graphs-Geometry," and "G2 Graphs-Algebra II" sheets of the Excel template. When reviewing graphs, keep two points in mind:

- Subgroups of students are not mutually exclusive groups. For example, multilingual learners are represented in the "All Students" results and can also be represented in other subgroups, such as students facing socioeconomic barriers. The subgroups provide additional descriptive information on how coursetaking and proficiency levels can vary by student demographic information, baseline performance, and graduation diploma type.

- Results and graphs should not be finalized until you meet with content experts. Meet with the mathematics content experts (for example, mathematics coaches, instructional coaches, mathematics supervisors) to review the results and graphs for a "gut check" on the data. Content experts may be able to tell you if the results seem similar to (or far from) what they expected, given their knowledge of placement policies and student outcomes. If results are far from what content experts expected, you may want to double check your analytic sample, coding, and analysis.

Once the results and charts have been verified by the content experts, update the data as needed and finalize the charts. When finalizing the charts, include a "figure note" at the bottom of each figure to state the sample sizes of each group.

> Tip: Insert a figure note at the bottom of each figure. An example figure note can say: There were [n number] of students who completed Algebra I in grade 7 with valid data, represented in the "All students" bar. Among this group, [n] were multilingual learners in grade 5, [n] were students facing socioeconomic barriers in grade 5, [n] scored Advanced proficient in grade 5 mathematics, [n] scored Proficient in grade 5 mathematics, [n] earned a Standard diploma, and [n] earned an Advanced Studies diploma.

Next steps

This memo provides step-by-step instructions to help database managers and analysts answer important questions about the highest-level mathematics courses and the Geometry and Algebra II proficiency levels of students who were in grade 5 in the 2009/10 school year, completed Algebra I in grade 7 (2011/12), and graduated from high school in 2016/17. These analyses present information about one cohort of students who were hyper-accelerated in their mathematics coursetaking pathway, that is, students who completed Algebra I in grade 7. With this information, database managers and analysts can support school division leaders and mathematics content experts to use data to inform and improve decisions about policy and practice related to mathematics coursetaking pathways.

The next steps include implications for policy and practice and implications for additional analyses and further study.

Implications for policy and practice

When sharing these data and graphs with school division leaders and mathematics content experts, database managers and analysts may want to use a two-step process. The first step is to review the data with school division leaders and mathematics content experts. The purpose is to help those stakeholders understand the results and make sense of the data. Database managers and analysts can guide a data-review discussion, answer any questions about the students represented in the analysis, describe data anomalies and missing data, share the results, and help interpret the graphs. Questions to consider when reviewing the data might include:

- Looking at the data and graphs, describe what you see?

- What patterns do you see in the data?

- What questions come to mind as you look at the data and graphs?

During this initial review, school division leaders and mathematics content experts should focus on making observations about the data and results, rather than identifying solutions or changes to policies and practices.

Once stakeholders understand the results, the second step is to reflect on mathematics policies and practices. The purpose is to help school division leaders and mathematics content experts use the data and results to inform and improve policies and practices. Results from the analysis can help school division leaders and mathematics content experts think about inputs and outcomes of mathematics coursetaking pathways. Questions to consider when thinking about implications might include:

- How are students placed into mathematics coursetaking pathways?

- What differentiated supports do students receive during Algebra I?

- How are students supported in advanced-level mathematics courses throughout high school?

School division leaders and mathematics content experts can use this information to reflect on policies and practices related to Algebra I placement and mathematics instruction for these hyper-accelerated students.

School division leaders and mathematics content experts can also think about what other types of information would help them better understand student experiences in different mathematics coursetaking pathways. The additional analyses below highlight how database managers and analysts can adapt the instructions in this memo to conduct additional analyses of interest.

Implications for additional analyses

Database managers and analysts can adapt the steps presented in the analysis memo to conduct additional analyses. Below are examples of ways to adapt the steps in the memo to answer other questions. The options below would follow the same process to create the analytic dataset and conduct your analyses. The same Excel template can be used to generate graphs.

Different cohort of students. If you would like to conduct the same analyses but with students who graduated recently, say in the 2020/21 school year, in step I.1 you would identify students who were in grade 5 in the 2013/14 school year, and then continue with all the subsequent steps in the memo (accessing data from when these students were in grade 5 in the

2013/14 school year and identifying the students who completed Algebra I in grade 7 in the 2015/16 school year and graduated high school in 2021). Replicating these analyses with different cohorts of students to monitor changes over time is important as leaders consider implications for policy and practice in algebra readiness and success.

Different Algebra I completion. If you would like to conduct the same analyses but with students who completed Algebra I in a different grade, say grade 8, in step I.2 you would identify students who completed Algebra I in grade 8, and then continue with all the subsequent steps in the memo. Comparing the inputs and outcomes of students who completed Algebra I in different grades is important as leaders consider implications for policy and practice in placement in mathematics coursetaking pathways and differentiated supports.

Different subgroups of students. If you would like to conduct the same analyses but with different subgroups of students, say students with disabilities, in step I.1 you would merge in the variable on student disability, and then continue with all the subsequent steps in the memo. This is true for student race/ethnicity and other student demographic variables. Conducting analyses of different subgroups of students is important for accountability purposes as leaders consider implications for policy and practice to ensure that all students are learning and succeeding in mathematics.

References

Allensworth, E. M., & Easton, J. Q. (2005). *The on-track indicator as a predictor of high school graduation.* The University of Chicago, Consortium on Chicago School Research.

Allensworth, E. M., & Easton, J. Q. (2007). *What matters for staying on-track and graduating in Chicago Public Schools : A close look at course grades, failures, and attendance in the freshman year.* The University of Chicago, Consortium on Chicago School Research.

Burkham, D.T., & Lee, V.E. (2003). *Mathematics, foreign language, and science coursetaking and the NELS:88 transcript data (National Center for Education Statistics Working Paper No. 2003-01).* U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics.

Galanti, T. M. (2019). *Hyper acceleration of Algebra I: Narrating opportunity to learn from a situative perspective*. George Mason University.

Holian, L., & Mokher, C. (2011). *Estimating college enrollment rates for Virginia public high school graduates* (Issues & Answers Report, REL 2011-104). U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Appalachia.

Jonas, D., Dougherty, C., Herrera, A., LaTurner, J., Garland, M. W., & Ware, A. (2012). High school predictors of college readiness: Determinants of high school graduates' enrollment and successful completion of first-year mathematics and English college courses in Virginia. Virginia Department of Education.

Jonas, D., Garland, M. W., & Yamaguchi, R. (2014). Following Virginia's career and technical education completers out of high school and into college: A study of high school graduates' college enrollment, persistence, and completion. Virginia Department of Education.

LaFave, A. (2019). *Data point: Algebra I coursetaking and postsecondary enrollment* (NCES No. 2019-154). National Center for Education Statistics Working Paper. U.S. Department of Education.

Matthews, M. S., & Farmer, J. L. (2008). Factors affecting the Algebra I achievement of academically talented learners. *Journal of Advanced Academics, 19*(3), 472–501.

Stein, M. K., Kaufman, J. H., Sherman, M., & Hillen, A. F. (2011). Algebra: A challenge at the crossroads of policy and practice. *Review of Educational Research, 81*(4), 453–492.

Virginia Department of Education. (2020). *VLDS data dictionary*. Retrieved June 2, 2020, from https://vlds.virginia.gov/insights

Virginia Department of Education. (n.d. b). *Advanced studies diploma: Minimum course and credit requirements*. Retrieved June 2, 2020, from http://www.doe.virginia.gov/instruction/graduation/advanced_studies.shtml#11

Virginia Department of Education. (n.d. c). *Standard diploma: Minimum course and credit requirements*. Retrieved June 2, 2020, from https://www.doe.virginia.gov/instruction/graduation/standard.shtml

Virginia Department of Education. (n.d. a). *SOL test scoring and performance reports*. Retrieved June 2, 2020, from http://www.pen.k12.va.us/testing/scoring/index.shtml

Virginia Department of Education Virginia Mathematics Pathways Iniviative (n.d.). *What is the Virginia Mathematics Pathways Initiative?* Retrieved May 17, 2021 from https://www.doe.virginia.gov/instruction/mathematics/vmpi/index.shtml

Watson, J., Carlivati, J., & McCarroll, J. (2010). *Eighth grade algebra: Findings from the eighth-grade round of the Early Childhood Longitudinal Study, Kindergarten Class of 1998–1999* (ECLS-K) (NCES No. 2010-016). National Center on Education Statistics Working Paper. U.S. Department of Education, Institute of Education Sciences.

Yamaguchi, R., Jonas, D., Lynne, Schmidt, R. A., Sieber, M., Buffington, P., Neumayer DePiper, J., & Araoz, C. (2020). *Algebra I and college preparatory diploma outcomes Among Virginia students who completed Algebra I in grades 7–9* (REL 2021-038). U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.