

**Researcher’s guide to**

**Massachusetts state education data**

2024

Office of Planning and Research

**Table of Contents**

Contents

[I. Introduction 2](#_Toc159409337)

[II. DESE’s Research Agenda 2](#_Toc159409338)

[III. Available data 2](#_Toc159409339)

[IV. Requesting Access to State Data 7](#_Toc159409340)

[V. Criteria for access to confidential data 11](#_Toc159409341)

[VI. Using SIMS Data 12](#_Toc159409342)

[VII. Using SCS data 20](#_Toc159409343)

[VIII. Using MCAS Data 23](#_Toc159409344)

[IX. Using VOCAL (school climate) data 32](#_Toc159409345)

[X. Using ACCESS for ELLs data 34](#_Toc159409346)

[XI. Using other student-level data: SSDR, AP, SAT, ACT, and GED/HiSET 37](#_Toc159409347)

[XII. Using NSC data 38](#_Toc159409348)

[XIII. Using DUA’s Unemployment Insurance wage data 40](#_Toc159409349)

[XIV. Using educator data: EPIMS and ELAR 42](#_Toc159409350)

[XV. Using financial data 44](#_Toc159409351)

[XVI. Appendices 48](#_Toc159409352)

[Appendix 1: MCAS test data availability by subject and grade 48](#_Toc159409353)

[Appendix 2: Sample SIMS file format 51](#_Toc159409354)

[Appendix 3: Standard template for data-sharing agreements 53](#_Toc159409355)

[Appendix 4: Data Impacts of Coronavirus 60](#_Toc159409356)

[Appendix 5: Additional DESE Policies for Research and Evaluation: 65](#_Toc159409357)

[Policy Regarding the Use of Open Science Practices 65](#_Toc159409358)

# I. Introduction

The Massachusetts Department of Elementary and Secondary Education (DESE) has a wealth of data on students, schools, districts, and educators available for research purposes. The purpose of this guide for researchers and evaluators is to clarify what data is available, how to obtain and interpret it, and, ultimately, to generate better research projects and more accurate results that can help inform decisions.

The Office of Planning and Research (OPR) is the first point of contact for any external researchers wishing to use DESE data, whether on a contract with DESE or for their own research purposes. OPR can direct researchers to the non-confidential data that are available and shepherd researchers through the process of qualifying to receive confidential data, if appropriate. Matt Deninger, Chief Strategy and Research Officer, is the best place to start. He can be reached at [Matthew.J.Deninger@mass.](mailto:Matthew.J.Deninger@mass.)gov.

# II. DESE’s Research Agenda

DESE’s goal is that as a result of their public education in Massachusetts, students will: Attain academic knowledge and skills, understand and value self, understand and value others, and engage with the world so that they can be curious and creative, shape their path, feel connected, and be empowered.

We encourage research teams to consider and promote diversity, equity, and inclusion within all projects.

Research focused on any one or more of these areas within the [Educational Vision](https://www.doe.mass.edu/bese/docs/fy2023/2023-05/item7.1-educational-vision.pdf), whether broad or narrow in scope, would be of potential value to DESE. Please check the Office of Planning and Research’s website for updates to the agency’s research agenda <https://www.doe.mass.edu/research/researchers.html>.

# III. Available data

We have four kinds of data that are generally available for researchers. Those include:

* aggregate data at the school, district, and statewide levels,
* confidential student-level data,
* non-confidential student-level data, and
* individual-level data on educators.

Additional data may be made available for research purposes that are collected by DESE program offices. These data may be at any of the following levels. In what follows, we describe each of these types of data and point researchers to relevant resources.

**Aggregate data**

Most of the aggregate data that DESE collects on students, educators, schools, and districts can be found online at one of the following links. Researchers may also wish to review OPR’s Reports webpage: <http://www.doe.mass.edu/research/reports/>. This page includes links to all of the Department’s major report series as well as to recent program evaluations and legislative reports.

* The **School and District Profiles** website: <http://profiles.doe.mass.edu/>. *Profiles* includes school- and district-level data on enrollment, class size, graduation and college enrollment rates, dropouts, MCAS and SAT results, school and district accountability determinations, plans of high school graduates, per-pupil expenditures, teacher data and salaries, and more. It also provides directories of all the schools in the state by type (public, private, charter, collaborative, etc.). All *Profiles* data are posted in year-by-year tables that are exportable in Excel format.
* The **Statistical Reports** website: <http://www.doe.mass.edu/infoservices/reports/>. This lists all the reports produced by the Education Data Services unit, many of which are in queryable and/or downloadable format. While much of this information (graduation rate, grade retention, mobility, and dropout reports) is also posted on *Profiles*, some provide additional details, breakdowns, and statewide summaries.
* The **school and district finance** website: <http://www.doe.mass.edu/finance/statistics/>. Finance reports on the School Finance and District Support website include per pupil expenditures, teacher average salaries, special education expenditures, and state totals. Other useful reports on charter schools, school choice, and state aid programs are available as well. Greater explanation about the financial data sets can be found in Section XII of this guide. Additional financial data are also available in the DART Detail: Staffing and Finance (see below).
* DESE has created **aggregate school- and district-level files** to make student assessment results, data used for accountability determinations, and higher education enrollment data from the National Student Clearinghouse available in a more accessible format for researchers. MCAS files include student questionnaire data in relevant grade files. Files can be accessed after completing a brief online form available at: <http://www.doe.mass.edu/infoservices/research/>.
* Researchers may also wish to review the Department’s **District Analysis and Review Tools (DARTs)**: <https://www.doe.mass.edu/dart/>. The DARTs offer snapshots of district and school performance, allowing users to easily track select data elements over time and make meaningful comparisons to the state or to comparable organizations. The DARTs for Districts and Schools cover a broad range of district and school interests including demographic, assessment, student support, educator, financial, and achievement gap data. The DARTs for districts and schools are available through the *Profiles* webpage. The Department also has a suite of DART tools that provide in-depth data on targeted topic areas including: DART Detail: Staffing and Finance; DART Detail: English Language Learners; Dart Detail: Success after High School. These, and DART User Guides are available at the same link. A reference listing data points contained across DARTs is posted at <http://www.doe.mass.edu/dart/dart-reference-guide.xlsx>.
* The **Resource Allocation and District Action Reports (RADAR)** provides information on district resources. This data tool allows for the comparison of up to 10 districts, tracking trends over 5 years, and reports on associations between district resources and performance. <https://www.doe.mass.edu/research/radar/>
* **Views of Climate and Learning (VOCAL)** is an annual survey sponsored by the Department, asking students about their views on three dimensions and nine topics related to school climate. Aggregate data may be requested. For more information, see the VOCAL section of this research guide or <https://www.doe.mass.edu/research/vocal/>

* TheMassachusetts **Education-to-Career (E2C) Research and Data Hub** brings together data and research about how children and students progress from early education and care settings, through their K-12 education, and into higher education and the workforce. A variety of data sources for different audiences are available online: <https://educationtocareer.data.mass.gov/>

**Confidential student-level data**

Confidential student-level data are made available only to researchers who have submitted an approved research project proposal and have agreed to the terms of a data-sharing agreement with the agency. Please note, the [Family Educational Rights and Privacy Act of 1974 (FERPA)](https://studentprivacy.ed.gov/node/548/) protects student privacy and permits the sharing of individual-level student data to organizations conducting studies for, or on behalf of, educational agencies or institutions to:

1. Develop, validate, or administer predictive tests; or

2. Administer student aid programs; or

3. Improve instruction; or

4. To audit or evaluate Federal- or State-supported educational programs.

To obtain confidential student-level data, a research project **must** qualify under one of these exemptions for data sharing.

Data that are available for research purposes are described in greater detail in later sections. Links to current data handbooks for most data sets can be found on DESE’s [Data Collection](http://www.doe.mass.edu/infoservices/data/) website and are listed in the About section for each data set in this document.

*Note*: The response to the coronavirus pandemic impacted many of the following datasets during the 2019-2020 and 2020-2021 academic years. Please see [Appendix 5](#_Appendix_5:_Data) for details.

* **Student Information Management System (**[**SIMS**](http://www.doe.mass.edu/infoservices/data/sims/)**)**: Individual student-level records linked by a unique statewide identifier (SASID) and including student demographics, attendance, discipline, program participation, and plans after high school. 2002 to present. Currently collected in October, March, and end of year.
* **Student Course Schedules (**[**SCS**](http://www.doe.mass.edu/infoservices/data/scs/)**).** Individual student-level records on which students are enrolled in which courses, along with course completion and performance. 2011 to present. Currently collected in October and end of year; the October file is generally not shared for research purposes.
* **Massachusetts Comprehensive Assessment System (**[**MCAS**](http://www.doe.mass.edu/mcas/)**).** The Massachusetts Comprehensive Assessment System assesses student performance in mathematics, English language arts, and science and technology/engineering. Student-level records available back to 1998, linked to SASIDs from 2002 to present.
* **MCAS student questionnaire.** As part of the MCAS administration, students in grades 8 and 10 are asked to complete a questionnaire. In 2017, the questionnaire was also administered to grade 5 students; in 2019, grade 4 was added). The questions have changed considerably over the years, 2002 to present. In 2017, this data became focused on student perceptions of school climate and is referred to as the Views of Climate and Learning ([**VOCAL**](http://www.doe.mass.edu/research/vocal)) survey, 2017 to present.
* **English proficiency assessments.** Between 2005 and 2012, the Massachusetts English Proficiency Assessment (MEPA) was administered to all English learners (ELs). Massachusetts switched to the [**ACCESS**](https://www.doe.mass.edu/mcas/access/) assessment in 2012 (to present).
* **School Safety Discipline Report (**[**SSDR)**](http://www.doe.mass.edu/infoservices/data/ssdr/)**.** Individual student-level records of each offense that occurs on school property and/or a student is suspended or expelled. 2002 to present.
* **College admissions test data.** The Department receives SAT and Advanced Placement (AP) examination scores for students who take those assessments from College Board. This data is provided to the department without SASIDs; the Department matches the student names to their corresponding SASID. Graduating classes of 2007 to present.
* **National Student Clearinghouse (**[**NSC**](https://studentclearinghouse.org/)**).** The National Student Clearinghouse (NSC) is a non-profit institution that provides enrollment and degree verification from colleges and universities. NSC data is purchased by the Department each year and used to determine whether its high school graduates went to, persisted in, and/or completed postsecondary study.
* **Department of Unemployment Assistance** **(DUA) wage data**. The Department receives quarterly employment and earnings information for students who have enrolled in Massachusetts public schools and are linked to an individual in DUA’s data system. 2010 to present.
* **Voting Data.** The Department has voting data on Massachusetts residents, from the 2018 election.
* **High school equivalency data.** The Department has data on Massachusetts residents who take the GED or HiSET through adult education programming.

**Non-confidential student-level data**

We make available certain non-identifiable student-level MCAS and SSDR data for research purposes. Files for prior years are available via emailed link after completing a brief online form at <http://www.doe.mass.edu/infoservices/research/>. These data come in two forms:

1. records with individual student demographic information but no school or district identifiers, and

2. records with school and district identifiers but no student demographic information.

The most recent files (i.e., for the previous year) for Assessment data are typically posted in December each year; disciplinary data are typically posted in January or February.

**Data on educators**

Some educator-level data is public record, whereas others are made available only to researchers who have submitted an approved research project proposal and have agreed to the terms of a data-sharing agreement with the agency. These data sets are described in greater detail in later sections.

* **Education Personnel Information Management System (**[**EPIMS**](http://www.doe.mass.edu/infoservices/data/epims/)**)**. Demographic and work assignment information about educators and administrative personnel in schools and districts. 2007 to present.
* **Educator Licensure and Renewal System (**[**ELAR**](http://www.doe.mass.edu/licensure/elar/)**).** Data on all current and formerly licensed educators, plus recent enrollees in and completers of Massachusetts educator preparation programs. Beginning in 2023, DESE takes a snapshot of the database four times per year to facilitate research and data-sharing.

# IV. Requesting Access to State Data

**DESE-initiated Projects**

DESE-initiated projects are listed for public, competitive bidding on COMMBUYS (https://www.commbuys.com/bso/). These projects include contracted research and evaluation work, as well as partnerships with DESE program offices to pursue research grant funding wherein DESE receives any portion of the grant funds (e.g., federal grants, foundation grants, research partnerships).

Research teams that are awarded these contracts may include their winning project proposal in addition to the information requested below.

**Researcher-initiated Projects**

Researcher-initiated projects include projects aimed at generating academic papers, policy reports, letters of support, and/or grant projects where the funding is awarded only to the research team.

The first step in any researcher-initiated data request is to contact Matt Deninger, Chief Strategy and Research Officer, at [matthew.j.deninger@mass.gov](mailto:matthew.j.deninger@mass.gov). Please include a general summary of the idea for the project proposal. This step allows DESE staff to provide high-level feedback on project feasibility, interest, and potential overlap with prior or in-progress research.

**Project Proposal Process**

To receive data from the Department not already made available on our website—including but not limited to confidential student-level data—researchers must go through the following steps:

***1) Review eligibility for DESE data sharing***

OPR will determine whether a project will meet our criteria for data-sharing. The criteria DESE uses to assess whether or not to provide data for a particular research project are listed in [Section V](#_V._Criteria_for).

***2) Submit a project proposal to the Office of Planning and Research (OPR)***

The next step is submitting a project proposal. The proposal (typically 3-5 pages) must be on the letterhead of the sponsoring organization and must contain the following elements:

1. A brief description of the purpose of the study
2. A list of the research question(s) to be answered
3. The proposed methodology or analytical approach for answering those questions
4. The data needed from DESE to answer the research questions, specifying the particular data sets, years, and schools/districts from which data are required. Current data handbooks for most data sets can be found on DESE’s [Data Collection](http://www.doe.mass.edu/infoservices/data/) website and are linked, if available, in the *About* section for each dataset within this document.
   1. Please align this data request to your research questions, at a high level, e.g., SIMS will be used for student demographic characteristics.
   2. Please note whether you specifically need access to student or educator names, dates of birth, and/or local (not state) identification numbers, and if so, why. We only provide these data elements if specifically required for a study design, e.g., for data matching purposes.
   3. If requesting educator data, please state if you need educators’ evaluation ratings. We only give these data out if they are directly tied to a research question for the study, as they are subject to a higher confidentiality requirement than our other data.
   4. Please note if you already have access to the data you are requesting from another approved project (or note which you do and don’t have, if relevant).
5. Which [program office(s)](http://www.doe.mass.edu/contact/phone.aspx?mode=org) your work will support, and a description of any prior contact you may have had with them regarding their interest in the project. (Prior contact is not required; OPR just wants to know if you've already had discussions or not.)
6. The anticipated timeframe and deliverable(s) for the project.
   1. An anticipated end date is required for a data sharing agreement. This timeframe can be months or years depending on the scope of the project.
   2. In the timeframe, please propose appropriate touchpoints for agency program staff to engage with you throughout your project, e.g., early on for framing research questions and providing background information, midway for reviewing methodological decisions and/or draft findings, near the end for helping to interpret results, and so forth.
   3. In most cases the deliverable(s) should include a non-technical version oriented towards policymakers and/or practitioners along with whatever more technical work you may wish to produce for publication. Often this takes the form of a two-page summary in our [EdLines](http://www.doe.mass.edu/research/reports/category.aspx?section=education) series and/or a briefing for agency leadership and program staff.
   4. Where appropriate, deliverables should also include individual reports and/or briefings for any individual districts or schools participating in your project.
7. The source of funding for the project, if applicable
8. Names, titles, email, phone, and organizational affiliations and addresses of the following:
   1. The principal investigator(s) for the project
   2. The person who will serve as the liaison to DESE for the project
   3. The person or people who will sign off on the MOU (usually the PI or another authorized signatory from the research organization; some universities also add on someone from their IRB/sponsored projects office)
   4. The person or people who will require access to our secured drop box to obtain the data

This information will be used to complete the data sharing agreement as well as enter individual users in our secure data delivery portal.

***3) Receive proposal approval, revisions, or denial from DESE***

Approximately once a month OPR reviews proposals to determine whether a proposed research project meets current DESE information needs and data confidentiality requirements. In addition to the eligibility criteria and considerations listed in [Section V](#_V._Criteria_for), OPR discusses the proposals in relation to the following questions:

1. Is it compliant with FERPA regulations for data release?

2. Is the research feasible in light of:

a. the limitations of DESE data,

b. the timeline for data release and the project timeline,

c. proposed analytic approach.

3. What is the value-add of the research to DESE and the field?

a. Are other researchers currently engaged in answering this question?

b. Will the research provide actionable answers? (e.g., for teachers, students, parents, policymakers, staff, or programs)

c. Will the research address questions of interest to the Department in high priority areas for policy development and implementation?

Based on OPR’s review of the proposal, researchers will be sent a notice of approval, a request for revisions, or a denial for the project.

***4) Sign a memorandum of understanding***

If the project is approved, OPR will draw up a memorandum of understanding (MOU) detailing the nature and terms of the data-sharing agreement. The project and the MOU are then reviewed by DESE’s Legal Office and Education Data Services Office. Together with OPR, these three offices serve as DESE’s IRB, which meets once a month to discuss new agreements. The IRB may raise additional questions or concerns about the proposed project. Proposals are rarely denied at this stage, but researchers may be asked for additional revisions or clarifications.

The standard template for this MOU is included in the appendices to this document ([Appendix 4](#_Appendix_4:_Standard)). All parties must agree to the terms of the MOU before any data-sharing can occur. The language of the MOU is determined by federal guidelines for student and educator confidentiality. The requesting researcher is responsible for managing any approval process at his or her home institution, such as an Office for Sponsored Research or legal department.

A key consideration in the Memorandum of Understanding is how the security of the state’s confidential student data will be maintained. Recipients of state data are required to agree to many conditions, including:

* Storing data using industry-standard encryption and authentication;
* Not using the data for any purpose other than to conduct the specific approved research project;
* Not copying, reproducing, or transmitting the data to anyone else;
* Not reporting study results in a manner that could identify an individual student, including observing a minimum *N* size of 10 for any disaggregation;
* Implementing administrative, physical, and technical safeguards to protect data confidentiality, security, and integrity and to prevent redisclosure;
* Destroying all data once no longer needed for the purpose of the research project.

Researchers who cannot agree to these terms will not receive access to state data.

***5) Receive access to data***

Once the MOU is in place, OPR will request a secure portal be created for data delivery (i.e., a “dropbox”). OPR will also request the data from the Data Analysis and Reporting group within Education Data Services. The primary focus of this group is completing the federal and state reporting requirements. Depending on the complexity of the request and the timing relative to certain annual reporting requirements, it could take several weeks to fulfill a data request for back files.

As new data become available, OPR coordinates with the Data Analysis and Reporting group to deliver files to researchers in line with their MOUs. Researchers will be notified via email when new data files are available in their secure portal and do not need to contact the Department to request each file annually. We provide data in SPSS format only.

***Sample Timeline***

Given the rigor of the approval process and competing priorities of staff, it can take a substantial amount of time for a proposal navigate this process. What follows is a “perfect world” example timeline for approval through data delivery:

January 1 Project submitted.

January 15 OPR Proposal Review Meeting.

January 16-18 OPR sends notice of project approval.

February 15 DESE IRB Meeting to review project and draft MOU.

February 16 OPR sends draft MOU to researcher’s institution for legal review.

March 1 Researcher’s institution approves, signs, and returns the MOU; MOU is

executed by DESE; a secure portal is requested for the project.

March 15 A secure portal is created for the project.

March 30 All data available to date for the project are delivered to the secure portal.

A researcher who receives approval during the OPR review process should plan that their project will be reviewed by the IRB the following month. If there are revisions or further questions from either OPR or the IRB, it can extend this time frame substantially. MOU review and approval at the researcher’s institution can also extend this time frame. Please take this sample timeline and potential delays into account in your project planning.

# V. Criteria for access to confidential data

***Required criteria***

* The researcher must be affiliated with an academic institution or independent, nonpartisan research organization.
* The researcher and any institutions with which s/he is affiliated, including third-party funders of the proposed research project, must not be in an advocacy role with respect to the topic of study or have a stake in the study’s outcome.
* The study design and methodology must support an objective analysis of the proposed research question(s).
* The study must propose to examine a topic relevant for policymaking.
* The requester must be willing and able to comply with the requirements associated with being an authorized representative of the Department for the purpose of conducting the study, as defined in the MOU.
* If the study request is from a junior researcher such as a student or research assistant, the project proposal must be under the direction of a faculty advisor who already has an approved data-sharing agreement with the Department. The proposal must include a letter of support from the senior advisor.

***Additional considerations: Study design and methodology***

* The proposal includes specific, well thought out research questions that relate to a topic relevant for policy decision-making.
* The methodology aligns with the research questions (i.e., the research questions can be answered with the proposed methodology) and is consistent with current best practices in research design.
* The research questions can be answered effectively with data the Department has available.
* Previous work published by the researcher (and/or faculty advisor) suggests that he or she has expertise in the topic to be studied and the methodology to be used.
* Previous work published by the researcher (and/or faculty advisor) suggests that he or she has credibility with other researchers and experts.

***Additional considerations: Benefit and relevance to DESE policy and programming***

* Gathering more information on the proposed topic is a high Department priority.
* The information gained will be relevant statewide or for high priority districts, schools, or students.
* The study is likely to provide information in time to be of use for anticipated decision-making.
* The study does not overlap undesirably with other work already in progress.

Meeting all listed criteria is not a guarantee that the Department will make individually identifiable data available to a researcher; these criteria are necessary but not sufficient. Final discretion to share identifiable data with any researcher rests with the Department.

# VI. Using SIMS Data

**About SIMS**

The Student Information Management System, which launched statewide in 2002, contains individual student records, with a unique statewide identifier (SASID) for each student. SIMS collects data on student demographics, attendance, discipline, program participation, and plans after high school, among other information. Linking SIMS data files across years can allow for longitudinal calculations such as grade retention rates.

SIMS data are generally submitted three times per year school year: in October (SIMSA), March (SIMSC), and end of year (June) (SIMSD).[[1]](#footnote-2) These data usually become available for research purposes in January, June, and September, respectively. Most research projects use the October collection for the basic demographic information about each student, linking to the end-of-year collection for information on student dropout and graduation. The March series is collected primarily as preparation for administration of the spring MCAS and is not generally used for research purposes.

**SIMS Data Handbooks:**

Current: <http://www.doe.mass.edu/infoservices/data/sims/SIMS-DataHandbook.docx> Previous available by request at: <http://www.doe.mass.edu/infoservices/research/>

**Which students are reported in SIMS?**

Massachusetts public school districts must report all students in a district between the ages of 3 and 21 who are:

* Enrolled in the district, regardless of the reason;
* Enrolled in private schools or collaboratives (in state and out of state) and for whom the district is financially responsible;
* Not educated by the district but have active IEPs and receive their special education services from the district (may include home-schooled students or students educated in private schools); or
* Outplaced to the Department of Youth Services (DYS) or a correctional facility.

SIMS does *not* include:

* Home-schooled students (unless they receive special education services from the district);
* Parentally-placed private school students (unless they receive special education services from the district);
* Students in alternative/adult programs working toward their GED; and
* Students traveling abroad for the school year.

**What is a SASID?**

A SASID, or State-Assigned Student Identifier, is required for all students reported in the SIMS data collection. It is a ten-digit number unique to each student, and it stays with the student throughout the student’s public education in Massachusetts, including pre-K and higher education. The SASID is included in all SIMS submissions. SASIDs ensure that all the information associated with a student does not get associated with another student sharing similar identifying information.

**Common questions on SIMS business rules**

***Terminology for SIMS administrations***

The Department has several ways of identifying a particular administration of SIMS. In addition to the date, an administration of SIMS is given a period number representing the consecutive administration of SIMS as well as a year and letter. For example, the October 1, 2004, administration of SIMS is also Period 13, meaning it is the 13th administration of SIMS since SIMS originated, and SIMS 05A, meaning it is the October administration in the 2004–05 school year. The letter B was used for December administrations, which no longer occur. Consequently, most years use only the letters A, C, and D.

***Variable names***

The datasets for SIMS use variable headings that are shortened versions of full variable names. These full variables are linked to DOE data elements: the categories schools use when reporting their data. For example, the full variable name for the variable heading ENSTAT is “enrollment status at time of data collection” and is also DOE12. The handbook for a given period provides information, organized by DOE element, on the full definition of the variable as well as the possible responses.

Several variables in the SIMS datasets are not included in the DOE elements. The first is ORG\_CODE, or organizational code. This is a four-digit code for the school district and is the same as the first four digits of the school code, which is DOE015. The second variable not included in the DOE elements is PERIOD, which is explained in “Terminology for SIMS administrations” above.

***Variable definitions***

The current SIMS Data Handbook provides definitions, codes, and usage information for each of the DOE data elements. Some of the variables have changed over time. **Before including a given variable, be sure to account for any changes in the definition or coding of that data element.** The total number of variables also changes from year to year, as some are discontinued, and others added. Some of the important changes in variables are described below. Additional information on variable changes over time can be found in Appendix 4. Researchers should note that updated handbooks continue to be posted on an annual basis. Hence, all researchers should check the website for any updates that have been added after this version of the Researcher’s Guide was compiled and compare handbooks from the years relevant to their research.

***Changes to students’ characteristics across SIMS administrations***

Students sometimes have different codes for categories such as race, special education status, low-income status, or English Learner status in different SIMS administrations. Income and special education status can change over time; race represents whatever the parent or guardian indicates on school enrollment forms; and English Learner status reflects English proficiency per school determination. Researchers using data from a single administration of SIMS should use the code that is current at the time of that administration. For researchers using data from more than one year, determining which code to use will depend on the research question. The Department can help researchers decide how to address this issue for their particular research focus.

***Multiple SASIDs***

While the Department tries to avoid having multiple SASIDs for a single student, it does happen on occasion. In this case, use the SASID that is linked with an MCAS score when possible. Data made available to researchers will already have most multiple SASID issues resolved.

***Multiple records for the same student***

Once a student has been reported as enrolled in a district in a particular school year, he or she must be included in all district SIMS reports for the remainder of the school year even if he or she transfers out. As a result, students may have more than one record in any given SIMS administration. The enrollment status (ENSTAT) variable indicates whether a student is currently enrolled in a given school or is being reported as a transfer. The *reason for reporting* and *reason for enrollment* variables can also contribute to a student having multiple records. See below for the Department’s process of ordering enrollment status to make a SIMS file unique.

***Minimum N***

DESE uses a minimum N of 6 students for reporting any student demographic information and a minimum N of 10 for reporting student test outcomes. Accountability determinations are made for schools and districts with a minimum N of 20 in the aggregate and for subgroups.

***Graduation and dropout rates***

Researchers conducting analyses that include graduation or dropout rates should read Annual Dropout Rate vs. Cohort Graduation Rate, available online at <http://www.doe.mass.edu/infoservices/reports/gradrates/dropoutvsgrad.html> and the Frequently Asked Questions on graduation rates, available online at <http://www.doe.mass.edu/infoservices/reports/gradrates/gradratesfaq.html>.

In 2006, DESE began to calculate both a four-year and a five-year graduation rate for each cohort. Prior to 2006, graduation rates were estimated from annual dropout data or from grade-level enrollment information. For each cohort since 2006, the standard graduation rate calculation counts students who move within or between districts in whichever school and district they are in when they graduate. However, to understand better how schools and districts are doing with students that started ninth grade with them and didn't transfer out, DESE also publishes an adjusted graduation rate. The adjusted graduation rate includes only those students who were in the original cohort and did not leave the cohort. It does not include the students that transferred into the school or district after October 1st of ninth grade.

Linking SIMS files across time allows for calculations like dropout and graduation rate. The process of attributing student outcomes to particular schools and districts can be complex in certain situations and may rely on the March SIMS collection which is not typically shared with researchers. If it is critical for a research project to be able to precisely identify all dropouts and/or members of a graduation cohort as DESE defines them, a graduate and/or dropout file can be requested.

**Common questions on SIMS data elements**

***Race categories***

Through the 2004–05 school year DESE collected a single data element covering both race and Hispanic ethnicity. Its categories were: American Indian or Alaskan Native; Asian or Pacific Islander; Black; White; and Hispanic.

Beginning in 2005–06, DESE started collecting one data element for race and another for Hispanic ethnicity. It also separated the Asian or Pacific Islander category into two separate categories and allowed respondents to choose to identify with more than one race. The current data elements and possible categories are:

* *Ethnicity:* Not Hispanic or Latino; Hispanic or Latino
* *Race:* White; Black or African American; Asian; American Indian or Alaska Native; Native Hawaiian or other Pacific Islander

Each student now has one designation for ethnicity (either Hispanic or not Hispanic) and at least one designation for race, resulting in 62 possible combinations.

We include any student reporting Hispanic ethnicity in the Hispanic category regardless of race. Students classified as not Hispanic are included in the race category they designated. Non-Hispanic students with more than one race are considered multi-race. For example, a student who is both Black and Hispanic would count only as Hispanic when collapsing categories. A student who is Black, White, and Hispanic would still be included in the Hispanic category. A non-Hispanic student who is both Asian and White would be in the multi-race category.

This results in seven racial/ethnic categories: Hispanic or Latino; White; Black or African American; Asian; American Indian or Alaska Native; Native Hawaiian or other Pacific Islander; and multi-race, non-Hispanic. Additional information about the race/ethnicity categories can be found on the Department’s website at: <http://www.doe.mass.edu/infoservices/data/guides/race-faq.html>

***Bilingual education/English language learners***

For the 2003–04 SIMS administrations, “Bilingual Education Program Status” changed to “English Language Learners Program Status,” and the response codes also changed. The new variable reflected the passage of legislation in Massachusetts changing the default model of second-language education from transitional bilingual education to sheltered English immersion. The variables for the scores of limited English proficient students on a range of standardized tests were also eliminated. Also, between the 2003–04 and 2004–05 administrations, the meaning of response code 04 changed from “receiving English as a second language services (for waivered students only)” to “LEP student whose parent/guardian has consented to opt out of all ELL programs offered in the district.” In 2009, the LEP definition was modified, and ELL Program status codes were clarified. Until 2018, a former EL student was defined as a student not currently an EL but had been at some point in the two previous academic years. Beginning in the 2018–19 school year, former EL status was extended to the prior four academic years.

***Low income/economically disadvantaged status***

For the SIMS 2001–02 and 2002–03 administrations, the Department divided income status into those eligible and not eligible for free or reduced-price lunch. In the 2003–04 administration, the codes were expanded to allow differentiation between eligibility for free lunch and eligibility for reduced-price lunch.

In 2014–15, DESE substantially changed its data collection on low-income status in response to a policy change at the U.S. Department of Agriculture, which sponsors the free- and reduced-price lunch program. Under USDA’s [Community Eligibility Program](https://www.doe.mass.edu/cnp/nprograms/nslp/cep-eligibility.html), many of the state’s largest and poorest districts may offer free lunch to all students rather than having to individually qualify them for the program. These districts could choose to report each student’s free and reduced-price lunch status in SIMS, or they could report the default 00 value for all students. As a result, DESE no longer has systematic, statewide individual-level data on students’ free and reduced-price lunch status, and DOE19 is no longer used by DESE for reporting and statistical purposes.

Beginning in the 2014–15 school year, DESE used a metric called “economically disadvantaged,” which is based on a student’s participation in one or more of the following state administered programs: the Supplemental Nutrition Assistance Program (SNAP), the Transitional Assistance for Families with Dependent Children (TAFDC), the Department of Children and Families’ foster care program, and MassHealth (Medicaid). This variable appears as ECODIS in the data. Due to the change in methodology, the number of economically disadvantaged students in most schools in 2014-15 and beyond is lower than the number of low-income students reported in 2013-14 and prior years. Enrollment and achievement data for economically disadvantaged students cannot be directly compared to low-income data in prior years. For more information, see <http://www.doe.mass.edu/infoservices/data/ed.html>.

In 2021-22, the definition of low-income students in the Massachusetts SIMS data expanded because the state’s 2019 Student Opportunity Act required DESE to modify its definition. In recent years (2015-2021), students were identified as “economically disadvantaged” (ECODIS) based on participation in one or more of the following state-administered programs:

* the Supplemental Nutrition Assistance Program (SNAP);
* the Transitional Assistance for Families with Dependent Children (TAFDC);
* the Department of Children and Families' (DCF) foster care program; and
* certain MassHealth (Medicaid) programs.

Starting in the 2021-22 academic year, the state has returned to the “low-income” descriptor and has expanded the definition to include three additional groups of students:

* MassHealth program participants up to 185% of the federal poverty level;
* students identified by districts as homeless, and
* students that districts have confirmed have met the low-income criteria through a supplemental process, which involves the collection of required supporting documentation (SIMS DOE056).

See [Redefining Low-income Under the Student Opportunity Act (SY 2021-22)](https://www.doe.mass.edu/infoservices/data/sims/redefining-lowincome.html) for more information.

This change in measurement creates meaningful differences in who is included in this group of students, which has strong impacts for data analysis and reporting. Please see the [technical memo](https://www.doe.mass.edu/research/2021-01tech-memo.docx) for details.

As we continue to learn about the implications of this change, DESE has committed to including three variables related to low-income designation in SIMS files available for research purposes:

1. ECODIS\_old a variable that matches the previous definition of economic disadvantage from 2014-15 to 2020-2021 (i.e., up to 135% of the federal poverty level);
2. ECODIS a variable that is folded into the new LOW\_INCOME variable, that includes up to 185% of the federal poverty level;
3. LOW\_INCOME a variable that includes all students identified via federal poverty level as well as students identified by districts as homeless or low-income through the supplemental process.

These variables will be included for at least 3 years, to provide continuity in measurement across projects. DESE is partnered with researchers to identify and develop best practices for navigating this change and will provide additional information and recommendations in the future.

***Special education***

In the 2003–04 SIMS administrations, the Department added six new special education categories. Some of these were discontinued in subsequent years. In 2007–08, the variable for special education status was divided by age, with one for children between ages three and five, and another for students between ages six and 21.

***High needs***

DESE calculates a “high needs” student group based on the student’s characteristics across SIMS cycles. A student is high needs if he or she is designated as low income (prior to the 2014–15 school year) economically disadvantaged (starting in the 2014–15 school year), or low income (starting in the 2021-2022 school year), an English learner or former English learner (note change in definition of former EL described above), or a student with disabilities. Because of the change from low income to economically disadvantaged in the 2014–15 school year and from economic disadvantage to the new low income in the 2021-2022 school year, be cautious in interpreting or reporting changes in the high needs group over time.

***Days of membership/days of attendance: 555 code***

Starting in the 2006–07 school year, if a student was reported as enrolled in the previous school year’s end of year report, the student must be submitted again by the district in October of the following school year. If a student’s status change occurred over the summer, districts use the 555 response code for the days in membership and days in attendance variables when reporting summer transfers, graduates, and dropouts. In prior years, the Department relied on a Missing Students Report for students who were enrolled in the end of year report but did not appear again in the following school year’s October submission.

***Discontinued Variables***

* **Suspension.** Beginning in 2012–13, the in- and out-of-school suspension variable in the SIMS data collection was discontinued, with the default code for those variables as 0.
* **AP courses.** Beginning in 2012–13, the AP course variables in the SIMS data collection was discontinued, with the default code for those variables as 500. Students taking AP courses can be identified through the Student Course Schedule data sets.

Please see [Section XI](#_XI._Using_other) for further information on data related to student discipline and advanced course taking.

**Making a SIMS file unique**

Students may be attributed to multiple schools within a SIMS data collection for a variety of reasons. For example, a student may transfer from public School A to public School B, where they are currently enrolled. In this case the student will appear in the SIMS file as a transfer in state (enstat=20) at School A and enrolled (enstat=01) at School B.

When a unique file is needed to merge data sets or attribute a student to one school for a given time period, the following enrollment hierarchy process is performed on the SIMS file.

|  |  |  |
| --- | --- | --- |
| Enstat | Enrollment Hierarchy | Explanation |
| 01 | 1 | Enrolled |
| 06 | 2 | Deceased |
| 04 | 3 | Graduated |
| 10 | 4 | Certificate of attainment |
| 11 | 4 | Completed grade 12 and district approved program |
| 09 | 5 | Reached maximum age, did not graduate or receive a CA |
| 05 | 6 | Permanent expulsion (expulsion) |
| 31, 32…36 | 7 | Dropout |
| 20 | 8 | Transferred in state public |
| 21 | 9 | Transferred - in state private |
| 22 | 9 | Transferred out of state (public or private) |
| 23 | 9 | Home school |
| 24 | 9 | Adult diploma program leading to MA diploma |
| 40 | 10 | Not enrolled but receiving Special Education services only |
| 41 | 10 | Transferred, no longer receiving special education services only |

The file is then sorted by the SASID, enrollment hierarchy and days of membership (after removing the 555 days of membership coding used for summer transfers).

For example, SPSS coding would be:

RECODE ATTEND MEMBER (555=SYSMIS).

RECODE ENSTAT (‘01’=1) (‘06’=2) (‘04’=3) (‘10’=4) (‘11’=4) (‘09’=5) (‘05’=6) (‘31’=7) (‘31’=7) (‘32’=7) (‘33’=7) (‘34’=7) (‘35’=7) (‘36’=7) (‘20’=8) (‘21’=9) (‘22’=9) (‘23’=9) (‘24’=9)(‘40’=10) (‘41’=11) into ensort.

Sort cases by SASID (A) ensort (A) MEMBER (A).

Select if (sasid ~=lag(sasid)).

Note that SIMS DOE 12 expanded in 2007. Prior to 2007 enrollment codes were:

01 Enrolled

02 Transferred

03 Dropped Out

04 Graduate with Competency Determination

05 Permanent Exclusion

06 Deceased

09 Reached maximum age, did not graduate or receive Certificate of Attainment

10 Certificate of Attainment

**What other resources are available to help researchers use SIMS data?**

* SIMS website at <http://www.doe.mass.edu/infoservices/data/sims/>
* List of Massachusetts District and School Codes: available on the Department’s website at <http://profiles.doe.mass.edu/search/search.aspx?leftNavId>. From the pull-down menu, select Public School or Public School District and click on the blue Get Results button. The list of either schools or districts will come up and to the right is an export button that will export the information displayed into an Excel file.
* SIMS Explanation of the SIMS Summary Reports, available on the Department’s website at <https://www.doe.mass.edu/infoservices/data/sims/sumreports.html> provides information on the reports provided to districts and clarifies how certain calculations are done with the SIMS data. For example, the school enrollment report notes what students are included in calculating enrollment numbers based on certain SIMS variables.

# VII. Using SCS data

**About SCS**

The Student Course Schedule data set collects data on the courses students are enrolled in and their performance (marks) in those courses. By linking course assignments across SIMS, EPIMS, and SCS, we can identify which teachers serve which students. SCS is collected in October and at the end of the school year, though the October data is generally not shared for research purposes. The end-of-year data includes all the course data from October and has course completion and performance information (if applicable). End-of-year data is collected in mid-August and files are typically available to researchers around December.

**SCS Data Handbook:** <http://www.doe.mass.edu/infoservices/data/scs/SCS-DataHandbook.docx>

**Common questions about SCS data**

***Terminology of SCS administrations***

Similar to SIMS, each administration of SCS is given a period number as well as year and letter. For example the collection for the end of the 2014 school year is titled SCS14D and is also SC\_Period 13, meaning it is the 13th administration of the SCS.[[2]](#footnote-3) SCS is collected twice each school year, once in October, which is denoted by the letter A, and once at the end of the school year, which is denoted by the letter D. Typically only the end-of-year SCS data collection is shared with researchers.

***Variable names***

Similar to SIMS, the SCS dataset uses variable headings that are a shortened version of the full variable name.

|  |  |  |
| --- | --- | --- |
| Element | Name | File Heading |
| SCS01 | LASID | LASID |
| SCS02 | SASID | SASID |
| SCS03 | School/Program ID Number | CS\_LOC |
| SCS04 | Local Course Code | LCC |
| SCS05 | Subject Area-Course Code | COURSE |
| SCS06 | Class Section | SECTION |
| SCS07 | Course Term | TERM |
| SCS08 | Course Enrollment Status | CS\_STAT |
| SCS09 | Course Level | CS\_LEVEL |
| SCS10 | Course Credit Available | CS\_CAVAIL |
| SCS11 | Course Credit Earned | CS\_EARN |
| SCS12 | Course Letter Mark | CS\_LMARK |
| SCS13 | Course Numeric Mark | CS\_NMARK |

The SCS data set also includes the ORG\_CODE or organization code representing the school district, as well as the period which is explained in “Terminology of SCS administrations” above.

***Minimum N***

DESE uses a minimum N of 6 students for reporting any student demographic information and a minimum N of 10 for reporting student test outcomes. Accountability determinations are made for schools and districts with a minimum N of 20 in the aggregate and for subgroups.

***Variable definitions***

The current SCS handbook provides definitions, codes, and usage information for each SCS data element. Additional information about some of the data elements are below.

***Course types***

Districts provide information on what courses the students are enrolled in through two data points, local course code (SCS04) and state-defined subject code (SCS05). For state analysis, Course Subject Area Course Code (SCS05) is used to determine the subject area.

***Student performance and course marks***

Information on student performance can be provided a number of ways. Students may receive a letter mark, a numeric mark, or both. As outlined in the SCS data handbook, letter marks include standard alphabetical marks (e.g., A-), Pass/Fail, or categorical assessment (e.g., Minimally Acceptable). Numeric marks are on 100-point scale and can go to two decimal places. Students may also be noted as “mark not required” or “ungraded course.” Mark not required and ungraded courses appear frequently in the elementary grades, and to a lesser extent in middle school grades, particularly 6th graders in K–6 schools. For example, in the 2011 SCS data set, nearly 20 percent of grade 6 students had all their core courses marked as ungraded or mark not required. In high school grades, mark not required and ungraded courses for core subject courses are most common in proficiency-based programs. When doing analysis that focuses on course performance, be aware of the prevalence of ungraded coursework in your sample and make sure you are appropriately accounting for them.

When students have both letter marks and numeric marks, the state uses the letter mark in determining passing and failing. Courses are deemed as passing for state calculations if the student earned a letter grade of D- or better, a rating of minimally acceptable or better, “Pass,” numerical mark of 59.5 or higher, or full credit awarded for a credit-bearing course that is ungraded/mark not required.

Students only receive grades for completed courses (based on SCS08) in SCS. Courses in progress are noted as 88 or 88888.

***Course level (SCS09)***

It is not clear that this data element appropriately captures full, comparable information on the level of difficulty of all courses. This is not the variable the Department uses for calculating Advanced Placement or other advanced coursework. The state uses the NCES Course Subject Area Course Code (SCS05) to determine if it is an Advanced Placement course. A list of advanced courses used for accountability purposes can be found at <https://www.doe.mass.edu/accountability/lists-tools/>.

***Matching SCS to other files***

In many cases, the end-of-year SCS file is merged with the end-of-year SIMS file by SASID or by SASID, Org Code, and School Code, depending on the analysis. To merge SCS with EPIMS, match files on Org Code, School Code, Course, Section and Term.

***Course term***

The SCS data handbook outlines the array of options that may be used in referencing the length of a course. In aggregating the amount of course work taken in a given subject area, DESE assumes a semester course equals 0.5 of a full-year course, a trimester equals 0.33 of a full-year course, a quarter equals 0.25, a quinmester equals 0.2, etc. Multiple trimester and multiple non-consecutive trimesters are assumed equal to 0.667 of a full-year course, multiple quarters/multiple non-consecutive quarters are equal to 0.5 of a full-year course, and multiple quinmesters/multiple non-consecutive quinmesters are equal 0.4 of a full-year course.

***Other data anomalies***

* For a small number of high schools prior to 2016, all or most courses for graduating 12th graders were erroneously marked as withdrawn (and letter and numeric mark reflecting “Withdrawn”) in the SCS submission. DESE put validations in place beginning in the 2016 school year but be aware of the anomaly if looking at 12th grade course completion or performance rates.
* A few districts have missing SCS data in one or more collections, as follows:

|  |  |  |
| --- | --- | --- |
| District | Data missing | Data reported but EPIMS cross-validation waived |
| Boston |  | SCS16A |
| MA Virtual Academy |  | SCS17D |
| Greenfield | SCS11D |  |
| Lowell | SCS11D |  |
| Medford | SCS16D | SCS16A, SCS17A, SCS17D |
| Sharon |  | SCS17D |
| South Hadley | SCS16A |  |
| Watertown |  | SCS16A |

**What other resources are available to help researchers use SCS data?**

* SCS website at <http://www.doe.mass.edu/infoservices/data/scs/>.

# VIII. Using MCAS Data

**Who takes the MCAS?**

The 1993 Education Reform Law mandates that all students in the tested grades who are educated with Massachusetts public funds participate in MCAS, including students:

* Enrolled in public schools
* Enrolled in charter schools
* Enrolled in educational collaboratives
* Enrolled at public expense in approved and unapproved private special education schools within and outside Massachusetts
* Receiving educational services in institutional settings at public expense
* In the custody of either the Department of Social Services or the Department of Youth Services
* With disabilities (see note below)
* With limited English proficiency (see note below)

Some students with disabilities may receive certain accommodations to facilitate their participation, such as changes in the test timing, setting, presentation, or how the student responds to questions.

The small numbers of students with significant disabilities who cannot participate in MCAS even with accommodations are assessed using the MCAS Alternate Assessment (<http://www.doe.mass.edu/mcas/alt/>).

English learners (EL) are exempted from the MCAS ELA test in their first year in the U.S., though they must take any mathematics and science assessments offered at their grade level. Their ELA results, however, are not included in performance aggregations for any subject area during their first year. Since 2006, ELs have been considered first-year students if they first enrolled in a U.S. school after March 1 of the prior school year.

See the [MCAS website](http://www.doe.mass.edu/mcas/) for more information on the testing program in general. The [Test Administration Resources](http://www.doe.mass.edu/mcas/admin.html), [Test Questions](http://www.doe.mass.edu/mcas/testitems.html), [Results](http://www.doe.mass.edu/mcas/results.html), [Technical Reports](http://www.doe.mass.edu/mcas/tech/), and [Accessibility and Accommodations](http://www.doe.mass.edu/mcas/accessibility/) links may be particularly helpful.

**About MCAS Administration**

State assessment data includes data from the English language arts and mathematics MCAS assessments in grades 3 to 8 and 10; the science and technology/engineering (STE) assessments in grades 5, 8, and high school; the MCAS student questionnaire data in grades 8 and 10 (and grade 5, beginning in 2017); and the MEPA/ACCESS data on English learners’ proficiency in reading, writing, listening, and speaking English. When students take MCAS, they also have the opportunity to take the Views of Climate and Learning Survey (VOCAL) and the MCAS Student Questionnaire, which provide additional information about students’ experiences in school (detailed below).

**MCAS – Implementation Timeline.** Student-level MCAS data are available back to 1998 but can only be linked with SIMS data starting with the 2002 MCAS test administration. Prior to that time, some student demographic data were available from MCAS, but in the form of student self-reports rather than linked by unique student identifiers. The quality of the match between SIMS and MCAS improved substantially by the 2004 administration; for this reason, we recommend that researchers use student data from 2004 and beyond where possible. MCAS tests from before 2015 are referred to as Legacy MCAS tests.

From 2014-2019, Massachusetts transitioned to administering the Next-Generation (Next-Gen) MCAS: computer-based assessments with innovative item types aligned to the new curriculum frameworks. The transition to this test occurred over five years (2014 to 2019) and created some data availability and comparability issues that are important for researchers to understand.

In 2014, Massachusetts began a field test of the Partnership for Assessment of Readiness for College and Careers (PARCC) test. In spring 2014, approximately 700 randomly selected schools participated in the Performance Based Assessment field test of math or ELA tests developed by PARCC. In most cases, PARCC field testing was done in one subject (math or ELA) in only some classes at some grade levels in the participating schools. No performance level information was provided for students who took the PARCC field tests.

In 2015 and 2016, the PARCC was implemented in some districts. The majority of these participating schools chose to have the students who took the PARCC field test also take the MCAS; these students have complete student assessment data for 2014. About 10 percent of schools who took part in the field test did not double-test and have “PRC” noted as the performance information for field tested students in one subject and an MCAS score for the other subject area in the field-tested grades and classes. Thirteen schools only took PARCC tests in both subjects and all grades. This results in just under 10,000 students in grades 3 through 8 who do not have MCAS performance information in one or both subjects for 2014, representing 2.3 percent of students tested in those grades.

In 2015 and 2016, Massachusetts school districts had the option of administering MCAS or PARCC to their students in grades 3 to 8 to fulfill their state testing requirement in English language arts and mathematics. Districts that selected PARCC also had the choice to administer that test on paper or on a computer and could make different test mode choices by school. Districts that had selected PARCC in 2015 could not switch back to MCAS in 2016, but those who selected MCAS in 2015 could switch to PARCC in 2016. See the section below, “Working with 2015 and 2016 Massachusetts assessment data,” for details.

In fall 2015, the Department and Board of Elementary and Secondary Education decided to pursue development of the Next-Generation MCAS assessments, building upon the best aspects of MCAS and PARCC, combining innovative items developed by PARCC with new items specifically designed to assess the Massachusetts learning standards.

In 2017, the state administered the first Next-Generation MCAS in ELA and math to students in grades 3 through 8. Although the new assessments are intended to be administered on computer, transition to primary online testing in grades 3 to 8 was introduced two grades at a time (grades 4 and 8 in 2017, grades 5 and 7 in 2018, and all 3 to 8 grades in 2019). The Next-Generation MCAS ELA and math tests in grade 10 began in spring 2019, as did Next-Generation science for grades 5 and 8. For more information on the transition to next generation STE tests, see: <https://www.doe.mass.edu/mcas/tdd/sci.html?section=transition>. For a summary of MCAS test administrations by grade, subject, and year, please see [Appendix 1: MCAS test data availability by subject and grade](#_Appendix_1:_MCAS).

In 2020, all mandatory MCAS testing was cancelled due to the pandemic. The 2019-2020 10th grade ELA and math MCAS was not taken. The competency determination was waived for the class of 2020, which resulted in fewer retests and increased the graduation rates for 2020. 12th graders in the 2019-2020 academic year were given the option to meet their competency determination for graduation based on coursework. Because the 2020 MCAS was cancelled, there are no 2020 MCAS data available for research purposes.

In 2021, MCAS was administered in the spring. Grades 3-8 had 94% participation, with 17% of tests given remotely. Results for the grade 8 STE MCAS remote administration were unable to be scored due to a contractor error, and therefore scores were not provided for those students. Grade 10 ELA and Math had 91% participation, with no remote testing option. Please see [Appendix 5](#_Appendix_5:_Data) for more information on the impacts of the Coronavirus pandemic on MCAS data.

In 2022, MCAS administration returned to pre-pandemic schedule and Next Generation High School Introductory Physics and Biology tests were introduced.

A table documenting the changes to the MCAS over time can be found in [Appendix](#_Appendix_1:_MCAS) 1. These changes in test administration have made using MCAS data across these years challenging, but not impossible. Additional information on using the data from these assessments is below in the “Working with MCAS Data” section.

**Student questionnaire.** As part of the MCAS administration, students in grades 8 and 10 are asked to complete a questionnaire. In 2017 and 2019, the questionnaire was also administered to grade 5 and grade 4 students, respectively. Presently, grade 4 ,5, 8, and 10 students take the questionnaire.

The questions have varied over the years and have included topics such as post-high school plans, computer use, after-school activities, college preparation activities, hours spent on homework, self-efficacy on college and career ready skills, and other items. Some questions have been asked consistently for years. Use the search tool on the Department’s website to find the questionnaires for each year; suggested search term: “MCAS student questionnaire [year].”

Starting in 2015, many of the items on the questionnaire were developed to be reported with a Rasch index. The 2015 and 2016 questionnaires related to student perceptions of self-efficacy on the expectations of college- and career-ready standards in mathematics (2015) and ELA (2016).

From 2017 onwards, this data became focused on student perceptions of school climate and is referred to as the Views of Climate and Learning ([VOCAL](http://www.doe.mass.edu/research/vocal)) survey. The VOCAL measures student perception of three dimensions of school climate: engagement, safety, and environment. See [Section IX: Using VOCAL data](#_IX._Using_VOCAL) for more information.

**MEPA/ACCESS for English learners.** Between 2005 and 2012, the Massachusetts English Proficiency Assessment (MEPA) was administered to all English learners (ELs): those whose native language is not English and who are not able to perform ordinary classroom work in English. MEPA scores are available from 2005 to 2012.

Starting with the 2012–13 school year, ELs have participated in the ACCESS test instead. The results for 2012–13 include a crosswalk that translates ACCESS and MEPA scores. More information about ACCESS for ELs is available at <http://www.doe.mass.edu/mcas/access/>.

Our policy is to release MEPA and/or ACCESS data only for projects specifically looking at the English learner population.

**The MCAS and graduation requirements**

Students must pass the grade 10 English Language Arts (ELA) and mathematics and (for the class of 2010 and beyond) high school science/technology assessments in order to graduate from high school; this is called the “competency determination.” Students are permitted to retake each assessment as many times as they need to in order to pass; therefore, students may have multiple assessment records for these tests. Data regarding the competency determination rate can be found at <http://www.doe.mass.edu/mcas/results.html>.

The Coronavirus pandemic impacted graduation policies beginning for the class of 2020. Specific information on the MCAS graduation requirements for the graduating classes of 2020 through 2031 and beyond can be found at: <http://www.doe.mass.edu/mcas/graduation.html>

**General MCAS Data Information**

***Merging MCAS and SIMS databases***

DESE’ databases use a student’s SASID (unique student identifying code) to track student data. The SASID is, therefore, the best way to merge any databases. MCAS files from 2005 to the present have been checked for missing or incorrect SASIDs. MCAS files prior to 2005 may have a small number of duplicate or missing SASIDs. Duplicate SASIDs will not affect aggregation at the school or district levels. However, researchers tracking individual students should omit these cases from the data.

***Accounting for students with MCAS scores but no SASID***

Despite the Department’s best efforts, each year there are some students with MCAS scores but no SASIDs. These students’ scores are counted in state, district, and school aggregate results. From 2004 on, students without SASIDs are *not* counted in any subgroup totals. Prior to 2004, these students were counted in subgroups based on information provided on MCAS test booklets.

***Additional variables in the MCAS data sets***

Each year, the MCAS team merges in some variables from SIMS into the MCAS data sets to facilitate reporting. These often include calculated variables that can be useful to researchers, such as urban district designation; official school type (elementary, elementary/middle, middle, middle/high school, high school); number of years student has attended MA schools; number of years of continuous enrollment; whether the student was ever designated an English learner; and so forth. These variables vary from year to year depending on reporting needs.

***Race codes***

Before SIMS became available in 2004, information on race was reported directly on the MCAS booklets. Because of this, there may be some discrepancies on race prior to 2004. Additionally, the race/ethnicity variables for MCAS were not updated in the same year as for SIMS. In the 2006–07 school year, the previous year’s SIMS race/ethnicity codes for the student were used.

***Minimum N***

DESE uses a minimum N of 6 students for reporting student demographic information and a minimum N of 10 for reporting student test outcomes. Accountability determinations are made for schools and districts with a minimum N of 20 in the aggregate and for subgroups.

***Attribution of scores***

When a student attends a public school out of his or her district, that student’s MCAS scores are included with the school and district the student attends, even if a different district pays for the student. However, if a student attends a private school and is paid for by a public school district, as is sometimes the case for special education students, that student’s results are included in those of the sending district. The SPRP\_DIS and SPRP\_SCHvariables represent the district and school to which a student’s MCAS scores are attributed. The coding is the same as for the ORG\_CODE and SCHOOL variables in SIMS, although the actual school and district may be different in SIMS than in the MCAS database for a given student.

***Accounting for students with raw scores only***

In some cases, students have raw scores but no scaled scores. This can occur for several reasons. English learners who have attended school in the United States for less than one year are required to take the math and science MCAS tests, but their scores do not count and they do not receive scaled scores. Students who take only part of an MCAS exam—for example, if they are absent for one of the testing days—also receive only raw scores.

***Accounting for student absences during testing***

Before 2006, absent students were assigned scaled scores and performance levels based on the raw score earned (usually 200, or warning/failing). Since 2006, absent students receive a performance level of ABS, which counts in school and district participation rates, but not in aggregate performance levels.

***Accounting for students who change schools after October 1***

Beginning in 2006, the scores of students who enrolled in a school after October 1 of that year are counted in the district’s scores only, not in individual school scores. The exception to this is single-school districts, such as charter or regional vocational-technical schools, for which all enrolled students’ scores count. The OCTENR variable in the MCAS dataset identifies which students were and were not enrolled after October 1 each year.

**Working with MCAS data**

***Using scaled and raw scores from the Legacy MCAS***

Beginning in 2007, the Department modified the statistical technique that it uses to ensure the stability of the measurement of MCAS performance standards in grades 3-8. Details of the change are described in the [2007 MCAS Technical Report](http://www.mcasservicecenter.com/documents/MA/Technical%20Report/TechReport_2007.htm). Beginning in 2014, DESE modified the equating technique for the grade 10 ELA and mathematics tests. Details are available on page 6 of the [2014 MCAS technical report](http://www.mcasservicecenter.com/documents/MA/Technical%20Report/TechReport_2014.htm).

Researchers comparing scores across years on the Legacy MCAS (that is, data from prior to 2015) need to be aware of several important caveats. Raw scores represent the number of raw score points earned on the test. Scaled scores rescale the raw scores onto a common metric from 200 to 280 on two-point intervals, adjusting for test difficulty. Scaled scores represent different performance levels:

* 200 to 218 represent a performance level of Warning/Failing;
* 220 through 238, Needs Improvement;
* 240 through 258, Proficient; and
* 260 to 280, Advanced.

These scaled scores are common across all tests, but the raw scores associated with each scaled score vary across tests, years, grades, and subjects. To determine which raw scores translate to which scaled scores, simply crosstabulate the data for the particular test and year in question. Note that from 2001 through 2008 MCAS administrations contain performance levels, but no Advanced level, and no scaled scores for Grade 3 students.

We strongly recommend that researchers *not* use scaled scores for descriptive or multivariate analysis on Legacy MCAS data, in order to avoid estimating incorrect standard errors. Each performance level essentially has its own scale. The number of raw score points that translate into a single scaled score can vary considerably depending on the performance level, with error greater in the tails. As a result, the scale is not linear, and the scaled scores are heteroskedastic (different variances at different parts of the scales) so are not appropriately used for calculating means or conducting significance testing.

Instead, we recommend that researchers either calculate statistics on raw scores and convert them to the associated scaled score when reporting or convert the raw scores to standardized units (mean 0, variance 1) and report results as effect sizes. More information on this issue is available in the MCAS technical reports, available online at <http://www.mcasservicecenter.com/documents/MA/Technical%20Report/TechReport_2007.htm>.

**Working with 2015 and 2016 Massachusetts assessment data**

In 2015 and 2016, Massachusetts school districts had the option of administering MCAS or PARCC to their students in grades 3 to 8 to fulfill their state testing requirement. This section provides details on the testing process for the 2015 and 2016 data and how researchers should handle the resulting complexities of data analysis to ensure comparability and interpretability of findings.

***Test administration details***

Among the state’s school districts that serve grades 3 to 8, 46 percent administered MCAS and 54 percent PARCC in 2015; see [details](http://www.doe.mass.edu/news/news.aspx?id=13541). In 2016, 28 percent of districts administered MCAS, and 72 percent administered PARCC; see [details](http://www.doe.mass.edu/news/news.aspx?id=21379). Almost all districts were required to choose a single test district-wide; only Boston, Worcester, and Springfield were permitted to choose school by school.

In addition, PARCC districts had the option to administer the test online or on paper and could choose test mode school by school. In 2015, 31 percent of districts administered the test entirely on paper, 50 percent entirely online, and 19 percent a mix of the two. In 2016, 39 percent administered entirely on paper, 44 percent administered entirely online, and 17 percent a mix of the two.

All grade 10 students continued to take the Legacy MCAS, as it was currently the test used to meet the state’s high school graduation requirement. In 2015, a small sample of districts chose to administer PARCC in grades 9 and 11, and some grade 8 PARCC-takers took an Algebra I test rather than a grade 8 mathematics test.

***Representative samples***

Districts had the option to choose which test to administer, so the districts that selected MCAS may have been systematically different than those that selected PARCC. Differences in 2015 were small but DESE took additional steps to ensure that comparisons of data were fair and accurate.

Analysts selected representative samples of 2015 MCAS and PARCC test-takers that mirrored the test-taking population in 2014 and used those students’ results to identify which MCAS score was equivalent to which PARCC score in terms of proficiency. DESE then applied this information to the entire sample to generate a statewide percent proficient. DESE’s Office of Student Assessment published a [white paper](http://www.mcasservicecenter.com/documents/MA/Technical%20Report/2015/Appendices/Appendix%20A%20-%20Representative%20Samples%20and%20PARCC%20to%20MCAS%20Concordance%20Studies.pdf) detailing the methodology and also includes concordance tables that equate PARCC and MCAS scaled scores by subject and grade level. Student-level data provided to researchers for 2015 includes a flag for whether the student was in the representative sample. Researchers who want to calculate statewide statistics should select on this flag and run analyses just on this group of students.

DESE reported MCAS equivalent scores for students taking PARCC in 2016 using refined concordance tables. DESE did not report state results, however, because selection bias issues were too large to be overcome through this method. As a result, DESE did not report 2016 state-level achievement and growth results in grades 3 to 8 ELA and Mathematics. School- and district-level results were reported specific to whichever assessment was administered locally.

***Mode effects***

In online test administrations, DESE observed a small but meaningful difference in PARCC test results by test mode, with students who took the test on paper scoring somewhat higher than those who took it online. This is a common issue in the first year an assessment is administered online, particularly among students with limited access to computers at home and/or school. The Center for Analysis of Longitudinal Data in Education Research has published a [working paper](https://caldercenter.org/publications/pen-mightier-keyboard-effect-online-testing-measured-student-achievement) on this issue.

In 2017, the state used the representative sample and concordance work from 2015 to 2016 to render mode adjustments on PARCC thetas from those years. These mode-adjusted thetas are available for all grade 3 to 8 assessments in those two years and were used to calculate student growth percentiles from 2017 onward. (SGPs calculated in 2015 and 2016 did not adjust for mode.) Analyses conducted with growth percentiles indicated that the majority of mode effects were addressed through the method. Another mode adjustment was made during the 2019 administration for two districts that used the paper accommodated form for all students because they were unable to test online.

**Working with Next-Generation MCAS data**

In 2017, Massachusetts began testing students in grades 3 through 8 on the Next-Generation MCAS assessment in English language arts and mathematics. In 2017, students in grades 4 and 8 were required to take the test online, and 95% of students did so (a few got waivers). Districts had the choice of online or paper administration for the other grades. In 2018, students in grades 5 and 7 also tested online. In 2019, all students in grades 3 to 8 tested online unless they had an accommodation to take a paper form.

The next-generation ELA and math MCAS in grade 10 began in 2019, along with next-generation science tests for grades 5 and 8. The policies for future graduating cohorts and for the transition to next-generation science was planned to be set by the state Board of Elementary and Secondary Education in 2020 but was postponed to 2022 due to the Coronavirus pandemic. The next-generation science tests for Biology and Introductory Physics will be next-generation, computer-based tests starting in 2023. The Chemistry and Technology/Engineering test will be a legacy, paper-based test in 2023 and will no longer be offered beginning in 2024. Specific information can be found at: <https://www.doe.mass.edu/mcas/tdd/sci.html?section=transition>.

Unlike the PARCC scaled scores, the Next-Generation MCAS scaled scores already account for potential mode effects. Mode adjustment studies were conducted in each year in which schools were allowed to select the mode of administration. For grades in which online testing was the expectation, psychometric procedures were used to link the paper results to the online scale.

The scaled scores on the next-generation MCAS have also been designed to address the prior issues around non-comparability across grades and non-linearity of scales across performance levels. Studies using data just from the Next-Generation MCAS therefore should be able to work directly with the scaled scores, rather than using thetas or converting raw or scaled scores to z-scores.

Nonetheless we recommend that researchers do the usual tests of model assumptions before reporting findings. Thetas are provided to researchers starting in 2015 and continuing through the present. For more details, see the [MCAS technical reports](http://www.doe.mass.edu/mcas/tech/?section=techreports).

**File availability**

MCAS data from the previous year is certified in the fall of the following year. Research files are *typically* available as follows:

* MCAS data: December.
* Retest data: June.
* Competency Determination data: December and June.

**Combining assessment data across systems**

Researchers may wish to combine data from legacy MCAS, PARCC, and Next-Generation MCAS. This can be done but, particularly when significance testing will be used, should be done carefully to account for the psychometric properties of the tests and the sample selection issues during 2015 and 2016. See important notes above regarding using raw scores for analysis of legacy MCAS and adjusting for mode in PARCC data.

To combine data across assessments, we recommend the following:

* Up to 2014: Convert legacy MCAS raw scores to z-scores.
* 2015 and 2016: Use mode-adjusted thetas.
* 2017 forward: Use thetas provided via lookup table.

Researchers should consider the appropriate population from which to compute z-scores given their research question.

**What other resources are available to help researchers use MCAS data?**

* School and district-level legacy MCAS data from 2003 to 2018 are posted on the Department’s website at <http://profiles.doe.mass.edu/statereport/mcas.aspx>; Next-Generation MCAS data from 2017 onward are available at <http://profiles.doe.mass.edu/statereport/nextgenmcas.aspx>.
* MCAS Technical documents, including technical reports, validity studies, and information about test item development are available at <http://www.doe.mass.edu/mcas/tech/>
* The MCAS District File Layout provides variable names and descriptions for MCAS data for each year. To locate this, complete the Download Infoservices Files Form located at: <http://www.doe.mass.edu/InfoServices/research/download_form.aspx>. After selecting either the MCAS Student-Level Files or the MCAS Aggregated Files, scroll down to the hyperlink for the Excel Workbook with an Excel icon placed just to the right of the hyperlink. This file provides a list of variables, their meanings, and the possible response codes in effect for that year.

# IX. Using VOCAL (school climate) data

**What is VOCAL?**

Views of Climate and Learning (VOCAL) is an annual survey sponsored by the Department, asking students about their views on three dimensions and nine topics related to school climate, based on the U.S. Department of Education’s School Climate Surveys, detailed in the table below.

|  |  |  |
| --- | --- | --- |
| **Engagement** | **Safety** | **Environment** |
| Cultural competence | Emotional safety | Instructional environment |
| Relationships | Physical safety | Mental health environment |
| Participation | Bullying/cyber-bullying | Discipline environment |

After a pilot year in 2017, the survey has been administered in grades 5, 8, and 10 since 2018 and in grade 4 since 2019. It is taken as part of the Massachusetts Comprehensive Administration System (MCAS) and is optional for districts, schools, and students. In 2018, 84 percent of eligible fifth grade students, 89 percent of eighth grade students, and 60 percent of tenth grade students statewide participated in the survey, respectively. Additional information about the survey, including the survey instruments and summaries of findings, is available at <http://www.doe.mass.edu/research/vocal/>.

**About VOCAL data**

***Survey administration details***

The MCAS test is administered annually to students and the VOCAL surveys are taken after the last testing session in relevant grades; schools are responsible for the MCAS and survey administration. Survey administration was as follows:

* In 2017, a paper survey was piloted in grades 5, 8, and 10. In grades 5 and 8, the surveys were administered as part of the MCAS Science and Technology/Engineering (STE) achievement test. Grade 10 surveys were administered during the mathematics MCAS test.
* In 2018, grades 5 and 8 were administered the survey as part of the STE test via computer. Grade 10 survey was administered via paper during the mathematics MCAS test.
* In 2019, all surveys were computer based. In addition to grades 5 (STE), 8 (STE), and 10 (mathematics), the survey was also administered to grade 4 students via the mathematics MCAS test.
* In 2020, VOCAL and MCAS were not administered due to the Coronavirus pandemic; as a result, the VOCAL survey data are not available for that year. For more information, please see [Appendix 5](#_Appendix_5:_Data).
* In 2021 through the present, all surveys were computer based. The survey was administered to Grades 4 (mathematics), 5 (STE), 8 (STE), and 10 (mathematics).
* See [Appendix 1: MCAS test data availability by subject and grade](#_Appendix_1:_MCAS).

For paper surveys, students marked their responses in their MCAS student answer booklets. The computer-based surveys were designed to provide one item per screen; students provided their response and then advanced to the next screen and item. Each item/screen was prefaced with the words, “Think of the last 30 days of school.” More details of the survey administration protocols can be found [here](http://www.doe.mass.edu/research/vocal/2018/validity-study.docx). Beginning in 2019, each grade-level form was composed of 40 items and computer-based.

***Types of scores***

Two types of scores are available: student-level item response data and Rasch-based index scores.

* A Likert scale with four response options was used to rate students’ perceptions of school climate; coding for all items dictated that a response of “0” (*untrue*) indicated the lowest level of school climate, with a “3” (*always* *true*) denoting the most positive school climate. Response scoring categories “1” and “2” corresponded to *mostly untrue and mostly true*, respectively. Note, sixteen items were reverse-scored: eight bullying behavior items, five physical safety items, one emotional safety item, and one mental health environment item were reverse-scored. A higher item score, irrespective of whether the item is positively or negatively valanced, is associated with a more positive school climate.
* Four Rasch-based index scores are provided to schools: an overall VOCAL score, an engagement score, a safety score, and an environment score. Each student receives a index score for the four measures. These scores are transformed to make them more interpretable and provided in aggregate to schools and districts. In addition, districts receive a bullying index score, based on a subset of safety items. In 2018 (baseline year), student-level index scaled-scores were centered at 50 with a standard deviation of 20. To accurately measure trends, student responses from 2019 were placed on the baseline scale using common anchoring items. Before aggregation to the school-level, student scores were truncated to range from 1 to 99. A similar process was used for each dimension or bullying score.

**For more technical information about the VOCAL measure, see:**<https://www.doe.mass.edu/research/vocal/>

***Minimum N and reliability***

For schools and districts to receive VOCAL-related index scores, DESE uses a minimum N of 10 students for reporting student survey outcomes and a minimum index score reliability of 0.7 (Rasch-based person separation reliability measure). Similarly, for schools and districts to receive VOCAL-related item-level frequency response data, DESE uses a minimum N of 10 students.

***Access to VOCAL data***

Aggregate VOCAL results are public and are currently available for use. Researchers may request school and district level VOCAL data to answer related to questions related to school climate and culture.

**What other resources are available to help researchers use VOCAL data?**

* The VOCAL website <https://www.doe.mass.edu/research/vocal/> includes copies of the surveys, frequently asked questions, and other tools as well as research briefs, studies, and webinars on VOCAL use.

# X. Using ACCESS for ELLs data

**What is ACCESS for ELLs?**

Federal and state laws require that English learner (EL) students be assessed annually to measure their proficiency in reading, writing, listening, and speaking English, as well as the progress they are making in learning English. In fulfillment of these laws, EL students are required to participate in ACCESS for ELLs (Assessing Comprehension and Communication in English State-to-State for English Language Learners) tests, which replaced MEPA tests beginning in the 2012-2013 school year.

ACCESS is an assessment of English language proficiency administered through the WIDA consortium to students who have been reported to the Massachusetts Department of Elementary and Secondary Education as ELs. The ACCESS assesses student achievement of the WIDA English Language Development Standards in reading, writing, listening, and speaking and is administered to all publicly funded ELs in grades K-12. In addition, the Alternate ACCESS for ELLs (Alternate ACCESS) is administered to EL students in grades 1-12 who have significant cognitive disabilities and are unable to take the regular ACCESS tests. The ACCESS is available in paper-based and computer-based versions for grades 1-12 (kindergarten tests are available on paper only, as is the same for the Alternate ACCESS).

Massachusetts has administered the ACCESS for ELLs tests since 2013 to ELs who attend Massachusetts schools and whose education is publicly funded. The ACCESS assessments are administered during January and February each year to all EL students in grades K–12, as required by ESSA and Massachusetts Chapter 71A, including students who have “opted out” of receiving English language services.

**About ACCESS data**

ACCESS for ELLs is administered once annually in January and/or February.

ACCESS tests measure how well EL students have achieved the English Language Development (ELD) Standards developed by the WIDA consortium. Massachusetts adopted the WIDA standards because they describe English language development in four areas, or “domains” (listening, reading, speaking, and writing); academic literacy in four subjects, plus social and instructional language. The WIDA standards are aligned with the Massachusetts Curriculum Frameworks. Only students who participate in all four domains receive overall scores. Students participate in tests designed for specific grade-level clusters as follows:

ACCESS Tests

| Format | Grade-Level Cluster Tests |
| --- | --- |
| Online | 1, 2-3, 4-5, 6-8, 9-12 |
| Paper | K, 1, 2, 3, 4-5, 6-8, 9-12; and Alternate ACCESS in grades 1-12 |

In 2016, WIDA introduced a new, more rigorous computer-based test. Schools in Massachusetts were given a choice to administer either computer- or paper-based tests, with the expectation that the state would transition virtually all EL students to computer-based testing for the 2019 test administration. Paper-based tests remain available for use as an accommodation due to a disability, or for a first-year EL who lacks familiarity with or the ability to use a computer.

For paper-based testing, students in grades 1-12 take either Tier A or B/C of the designated grade-cluster test, depending on their level of English language proficiency. For computer-based testing, the listening and reading tests are adaptive, with students directed to questions of equal, lesser, or greater difficulty depending on how they responded to a previous cluster of questions. Based on listening and reading results, students are routed to the appropriately tiered speaking and writing tests.

Students taking the ACCESS test receive scaled scores between 100-600 overall, in each domain, and several composite scores. Proficiency level scores are reported as a whole number followed by a decimal (e.g., Level 3.4), indicating the student’s overall language proficiency level, and the relative position within the proficiency level of the student’s scaled score, rounded to the nearest tenth. English proficiency is reported in one of six English language proficiency levels: 1-Entering, 2-Emerging, 3-Developing, 4-Expanding, 5-Bridging, and 6-Reaching.

Assessment subscores are also reported using the same scaled scores and proficiency levels in the four domains of listening, speaking, reading, and writing; and reported as a composite score in the following combined areas:

Overall composite score combining all four domain scores;

Oral Language composite score, combining scores from listening and speaking;

Literacy composite score, combining scaled scores from reading and writing;

Comprehension composite score, combining scaled scores from listening and reading.

**Alternate ACCESS**

The Alternate ACCESS was administered in Massachusetts to EL students with significant cognitive disabilities. The Alternate ACCESS is given in grades 1–12 to students whose disabilities prevent participation in the ACCESS general assessment, and is recommended for students who participate in the state’s academic alternate assessment, the MCAS-Alt. It is administered individually by trained and certified educators in four grade-level clusters: Grades 1–2, 3–5, 6–8, and 9–12. Students are assessed in the four domains of reading, writing, listening, and speaking. Results are reported at six proficiency levels (Levels A1, A2, A3, P1, P2, and P3; and on a numerical scale from 900 to 960).

**Participation**

Participation means that a student participated in all four sections of the ACCESS or Alternate ACCESS test―reading, writing, listening, and speaking.

**Progress Indicators**

In 2018, a method was established to indicate progress in learning English that provided the following information for each EL student:

1. **Future Progress Target** representing the minimum ACCESS proficiency level score needed on the following year’s ACCESS test to remain on track to reach English proficiency (i.e., attain a score of Level 4.2) within a total of six years in a Massachusetts school. Future progress targets are reported as an ACCESS proficiency level, ranging from Levels 1.0 to 4.2. Individual student targets are provided in the summer of each year.
2. **Difficulty Index** showing an estimate of how difficult it will be to reach next year’s target, relative to the student’s current proficiency level. The difficulty index ranges from 1–99, with one being the lowest difficulty (i.e., least difficult to achieve the target) and 99 the highest (i.e., most difficult to achieve the target). Students with difficulty indices greater than 60 are considered to have a high difficulty index. Educators are directed to use the difficulty index as an indicator of which students may need additional instructional assistance in order to succeed.
3. **Student Growth Percentile for ACCESS** **(SGPA)** indicating the amount of progress made by a student on the ACCESS test from one year to the next, relative to other EL students in the same grade who earned similar ACCESS scores the prior year. In general, SGPA indicates whether a student has made low, moderate, or high gains in their ACCESS test scores from one year to the next compared with other ELs in the same grade level who took the ACCESS test in two consecutive years. Like the difficulty index, the SGPA is a number from 1-99, with 40-59 representing *average* growth. Higher SGPA numbers (i.e., above 60) indicate more growth, and lower SGPA numbers (i.e., below 40) show less growth.
4. **Progress Indicator**, either “Yes” or “No,” showing whether the student has met last year’s future progress target in the current year. Students who meet or exceed their future target for that year will be designated as having “made progress.” Students who do not meet their targets will not make progress for that year. Students taking the Alternate ACCESS will make progress if they increase the Alternate ACCESS achievement level in any subdomain by one level.

**Attaining proficiency**

Students who obtained an overall ACCESS score of Level 4.2 and a literacy composite score of Level 3.9 will be designated as having attained English language proficiency. Students who were absent were also counted in the overall attaining proficiency results. For more information on proficiency, please visit: <http://www.doe.mass.edu/news/news.aspx?id=24444>

**What other resources are available to help researchers use ACCESS data?**

For an extensive list of ACCESS materials, please visit: <http://www.doe.mass.edu/mcas/access/>

Understanding student scores: <https://wida.wisc.edu/assess/access/scores-reports>

# XI. Using other student-level data: SSDR, AP, SAT, ACT, and GED/HiSET

**School safety and discipline data (SSDR)**

**Handbook for SSDR data:** <http://www.doe.mass.edu/infoservices/data/ssdr/>

The SSDR report tracks each time an offense occurs on school property. Prior to 2012–13, this report only collected information on drug, violent, or criminal offenses for all students, along with all offenses by students with disabilities that resulted in suspensions or expulsions.

Beginning in 2012–13, the report expanded to also include any other suspensions or expulsions for non-drug, non-violent or non-criminal related offenses, irrespective of type of student, and the resulting disciplinary action. (Note: The in- and out-of-school suspension variable in the SIMS data collection was discontinued in 2012–13. SSDR is now the sole source of suspension information.) See <http://www.doe.mass.edu/infoservices/data/ssdr/> for more information on SSDR.

The trend line in the state’s discipline data was affected by the implementation of Chapter 222 of the Acts of 2012, a policy change regarding suspensions which took place on July 1, 2014. See DESE’s [highlights of changes to the state law](http://www.doe.mass.edu/sfs/) and the [advisory to districts](http://www.doe.mass.edu/lawsregs/advisory/discipline/studentdiscipline.html) for more details.

**College admissions tests (AP, SAT, and ACT)**

DESE receives student-level data on SAT and Advanced Placement (AP) examination scores for students who take those assessments from College Board. This data is provided to the department without SASIDs, and the Department matches the student names to the corresponding SASID. These are available for the graduating classes of 2007 to present.

Before 2017, DESE received SAT data on all students in each graduating cohort. In 2017, the data pulls changed to include all students taking the test that year. Currently we are missing data for students in the 2017 cohort who took the test before 2017, e.g., in 2016 as juniors.

DESE also received student-level data from the ACT program through 2015. Very few students in Massachusetts take the ACT and the data are not linked to SASIDs prior to 2015, so they are not commonly used in research projects. After 2015, DESE no longer collects this data.

**High school equivalency (GED/HiSET)**

Data on high school equivalency credentials, including the GED and HiSET, is available from the Adult and Community Learning Services (ACLS) office at DESE. Data on all individuals who complete a HiSET or GED through adult education programming are available for research purposes. These data are not matched to SASID and require individually identifiable data (e.g., name and DOB) to be matched to other datasets.

# XII. Using NSC data

**What is the National Student Clearinghouse?**

The National Student Clearinghouse (NSC) is a non-profit institution that provides enrollment and degree verification from colleges and universities.[[3]](#footnote-4) The Department has a contract with NSC to obtain these data and links them to other state data using the SASID.

**Handbook for NSC data:** <https://www.studentclearinghouse.org/educational-organizations/studenttracker-for-educational-organizations/>

**How the Department requests data from NSC**

Before 2012, the Department submitted batches of recent student graduates. For example, a file titled “2007 Graduates” contained all the students who were identified as graduates in 2007.[[4]](#footnote-5) These files provide a snapshot of student enrollment in postsecondary prior to the date they were created by NSC; they are not ideal for tracking students’ persistence for longitudinal studies. In spring 2012, the Department submitted a file to NSC containing all Massachusetts graduates from 2003 to 2010. We maintain these data for historical purposes.

Starting in 2012, the Department began submitting an updated file of graduates to NSC in late winter/early spring each year; these data are typically available to researchers in late spring/early summer. As new Massachusetts graduates are added to the file, earlier graduating classes are left off. The Department does not generally submit data on a graduating class until 16 months after graduation (i.e., information on the graduating class of 2015 will be available spring/summer 2017). This timeline aligns with our federal reporting requirements and provides the most complete picture of graduates’ enrollment in the year after high school graduation.

In 2013, 2015, 2017, 2019, and 2022 the Department submitted a file of students who were ever enrolled in the prior 10 school years in grades 8 through 12 or special education beyond 12th grade, but who were not currently enrolled or a graduate from a Massachusetts public school, to NSC for the purpose of determining the enrollment status and educational achievement at postsecondary institutions of these students who are not part of the graduate file. DESE refers to these files as “non-graduate” files. Because of NSC contract and budget constraints, we do not have a firm schedule for future submissions of the non-graduate file, though we do intend to continue to request these data occasionally.

SEARCH DATE is a variable in the NSC files which DESE includes in their submission to NSC. It is the date at which point NSC looks for enrollment in a postsecondary institution, and the format is YYYYMMDD. This variable has changed over time.

* Prior to 2016, the SEARCHDATE was January 1 of the year the student graduated from high school. For example, if a student graduated from high school in 2012, the SEARCHDATE was 20120101.
* Beginning in 2016, the SEARCHDATE is three years prior to the year the student graduated from high school. DESE changed what SEARCH DATE was included to capture a student’s college enrollment during their high school years (e.g., dual enrollment). For example, if a student graduated from high school in 2012, the SEARCHDATE is 20090101. Beginning in 2021, this variable returned to representing SEARCHDATE as January 1 of the year of graduation.

In summary, prior to 2016, SEARCHDATE is January 1 of the graduation year, from 2016-2020 it represents 3 years prior, and from 2021-present it returns to January of the gradation year.

**How NSC matches DESE to data**

NSC uses a matching algorithm to determine if a student enrolled in a postsecondary institution based on the first name, last name, middle initial and student’s date of birth submitted by DESE with the information submitted by postsecondary institutions. The Record Found indicator states if a match was made: Y means postsecondary enrollment information exists for the student, and enrollment begin and end dates indicates when the student was enrolled in postsecondary; N means the student is not found in NSC postsecondary records. This is not a perfect matching algorithm; both false positives (students who did not continue on to postsecondary are mistakenly matched up with students who did go to college) and false negatives (students reported as N who did enroll at college) can occur.

**What other resources are available to help researchers use NSC data?**

* Additional information on the National Center for Education Statistics Classification of Instructional Program (NCES CIP) codes used to classify majors is available at <http://nces.ed.gov/ipeds/cipcode>.
* For additional information on the strengths and weaknesses of the NSC data set and lessons learned about using this data in research, please refer to [The Missing Manual: Using National Student Clearinghouse Data to Track Postsecondary Outcomes](http://www-personal.umich.edu/~jmhyman/dynarski_hemelt_hyman_missing_manual.pdf) and other resources on the NSC’s [Notes from the Field](http://nscresearchcenter.org/workingwithourdata/) section on the Working with Our Data tab.

# XIII. Using DUA’s Unemployment Insurance wage data

**About the data**

Through a partnership with the Massachusetts Department of Unemployment Assistance (DUA), DESE receives quarterly employment and earnings information for students who have enrolled in Massachusetts public schools and are linked to an individual in DUA’s data system. Wage records are available from 2010 to present (typically with a reporting lag of three to four months) for individuals working in Massachusetts in employment covered under the unemployment compensation program. This does NOT include federal/military employment, self-employment, and employment outside of Massachusetts. DESE and DUA are able to link about 85% to 90% of students in each high school graduation cohort to an individual in the wage record system.

The two main elements of the wage data are quarterly wages and industry code of the employer. Employers report individual employee wages to DUA on a quarterly basis. Records are unique by person, wage year, wage quarter, and employer, so an individual working for two employers in the same quarter would result in two records. Each employer is assigned a North American Industry Classification System (NAICS) code based on the goods/services they produce. This is provided by DUA at the 4-digit level.

**Record matching**

To initiate the wage record matching process, the Executive Office of Education Information Technology team (EOE IT) uses the state’s Edwin Analytics data warehouse to create an input file with a record for all students of working age (14+) reported in SIMS from 2001 to present. This file contains a unique, non-SASID identifier, name fields, date of birth, and social security number (SSN) if applicable. Because DESE does not collect SSNs from K-12 students, SSNs are only available for Massachusetts K-12 students who later attended a Massachusetts public higher education institution; this provides the best linkage to DUA wage records.

EOE IT also creates a copy of the database containing the unique identifier, but masking all other identifiers including:

* First, middle, and last name
* SASID
* Locally Assigned Student Identifier (LASID)
* SSN
* Date of birth (exact date is changed to one that is +/- 30 days from actual DOB)

Following the file creation and the Edwin duplication/masking process, DUA links wage records using SSN if available and, if not, a partial matching procedure using name and date of birth. DUA then removes the identifiers (except for the unique identifier) and sends EOE a return file with the wage data. EOE then joins the wage data to the masked Edwin environment using the unique identifier. This process is repeated four times per year.

**What does this mean for researchers?**

As a condition of the data sharing agreement with DUA, EOE IT created the masked Edwin Analytics data warehouse to reinforce student privacy protections. Consequently, individuals in the wage dataset cannot be linked by SASID to student records in the standard files used by researchers (SIMS, SCS, NSC, etc.). Instead, the identifiers in the masked environment link to tables in the data warehouse which contain the same data as the standard files (except for masked fields), but in a format that is optimized for database storage.

To work with wage data, researchers must use extracts from the masked data warehouse which contain historical data for all years requested, not individual files for each collection period. Some column names match those found in the standard files, but some are as they appear in the warehouse. Researchers using these data will be sent a data dictionary to use as a guide for these new columns. These data are connected similarly to the state’s standard data sets and allow for teacher-student linking through district, school, course, and section identifiers.

Some research projects require connection of data collected by researchers (e.g., waitlists, assessment scores not collected by DESE) to the DUA data. In limited cases DESE can integrate these records into the Edwin Analytics data warehouse prior to a forthcoming DUA data match and Edwin masking procedure. Through that process, the SASIDs can be masked and assigned the same unique identifiers that correspond with wage records. After DUA matching, the records can be returned to the researchers with wage data attached. DESE needs to receive the researcher’s data for matching at least three to four months ahead of the targeted date for data release. This is resource-intensive for our staff, so DESE only does so for the highest priority research questions, at the discretion of the research director.

# XIV. Using educator data: EPIMS and ELAR

**What data are available about educators?**

DESE has two sources of information about educators: the Education Personnel Information Management System (EPIMS) and the Educator Licensure and Renewal data (ELAR). EPIMS collects demographic and work assignment information on individual public school educators twice per year (October 1 and end of year). ELAR collects educator preparation and licensure information on licensed educators and educators currently enrolled in preparation programs in the state, whether they are practicing in a public school or not. ELAR is a constantly updated transactional database; DESE takes a snapshot of it four times per year to generate data for analytical purposes. Please note that because ELAR and EPIMS are pulled multiple times per year, at different times, there may not be perfect alignment between the two datasets (e.g., a teacher may change positions or earn a new license at different points in the year).

**About EPIMS data**

**EPIMS Data Handbook:** <http://www.doe.mass.edu/infoservices/data/epims/epims-DataHandbook.docx>

***EPIMS data structure***

EPIMS collects two separate sets of records about each educator, linked by MEPID, the unique educator identifier. The staff roster collects one record per educator per district of employment, with demographic information such as age, gender, race/ethnicity, and employment status at the time of data collection. The work assignment data collects one record per educator per work assignment (i.e., a course assignment), with information about the specific course and section taught, grade level, course term, and full-time equivalency.

***Linking to the Student Course Schedule data set***

To merge SCS with EPIMS, match files on Org Code, School Code, Course, Section and Term. The student-teacher link is of higher quality in the 2011–12 years and forward, once school districts had fully adopted the SCED codes at the elementary and middle school levels.

***Data quality issues***

In some years, the beginning teacher field in EPIMS was not consistently reported across districts. It is best to identify new teachers as those that have not previously appeared in EPIMS. This will correctly capture all teachers new to teaching in Massachusetts public schools since 2007, but will incorrectly classify experienced teachers whose prior experience was in private schools or out of state.

Also, hire date appears to be reported by most districts as the date that employment started, rather than the date an offer of employment was made.

A few districts have missing data for some EPIMS collections, as follows:

* Boston EPIMS16A (except the Dever and UP Holland schools)
* MA Virtual Academy EPIMS17D
* Medford EPIMS16A, EPIMS16D (data accepted but cross-validation waived), EPIMS17A
* Sharon EPIMS17D
* South Hadley EPIMS16A
* Watertown EPMIS16A (data accepted but cross-validation waived)

**About ELAR data**

***Sharing ELAR data***

DESE generally shares ELAR data only for research projects specifically looking at issues related to educator preparation or licensure. Researchers should specify which of the five ELAR files (listed below) they need for their project.

***General data elements***

ELAR collects first/middle/last name, self-reported race, and self-reported gender on all records. MEPIDs are attached to the records through EPIMS identification maintenance. Some older records, particularly for non-active educators, may not have MEPIDs attached.

ELAR data comes in five files:

***1. Preparation program data elements***

ELAR collects preparation program organization code and name, general program area (e.g., Academic: Administrator), program description (e.g., elementary), grade level the program covers (e.g., 5 to 8), the credential being earned (initial/professional, baccalaureate/post-baccalaureate, etc.), the dates of program enrollment and completion, and the program completion status for each candidate enrolled in a preparation program. DESE began collecting preparation program data in 2009–10, but initially on completers rather than all enrollees.

***2. Educator license data***

ELAR collects educator license number, license credential field (e.g., biology, elementary), license grade level (e.g., 1 to 6, 5 to 8), license stage level (e.g., initial, professional), license classification (e.g., Academic: Teacher), date the credential was issued, date the credential expires or expired (where available), and status of the credential (e.g,. expired, licensed) for each license held by a Massachusetts educator.

***3. MTEL data***

ELAR collects MTEL test code, name, and date; test result (e.g, P=pass, F=fail), reading test result, writing test result, and numerical test scores for all tests taken by Massachusetts educators.

***4. Self-reported degree data***

Educators may choose to self-report their college major and degree level.

***5. Waiver data***

ELAR contains data on educators who received an emergency license waiver and those who received waivers to teach out of their content area.

**What other resources are available to help researchers use educator data?**

* EPIMS website at <http://www.doe.mass.edu/infoservices/data/epims/>

# XV. Using financial data

**What financial data are available?**

Sources of expenditure and revenue finance data include the End of Year Financial Report (EOYR) for municipal and regional districts and the Charter School End of Year Report (CSEOYR). Additional information on state aid and reimbursements are reported by the School Finance office, including Chapter 70 aid calculations and allocations, calculations of tuitions to charter schools and school choice, state reimbursement for special education extraordinary expenses (“circuit breaker”), and state transportation reimbursement.

Expenditures are typically reported by expenditure function and funding source. Funding sources include general appropriations (district and municipal expenditures for education), grants (federal, state, and local), and local revolving funds (e.g. athletics fees, food service revenues). Expenditures are reported out as annual dollar and per pupil amounts from the EOYR used by municipal and regional districts and CSEOYR by charter schools. However, the two reports do not have the same chart of accounts and cannot be merged. The EOYR chart of accounts can be found at <http://www.doe.mass.edu/charter/finance/revexp/>; the CSEOYR chart of accounts is embedded in the CSEOYR workbooks. Data from these different sources should only be compared at highly aggregated levels that include approximately the same expenditure categories.

**End of Year Report (EOYR) for municipal and regional districts**

The End of Year Financial Report (EOYR) is a comprehensive report of revenues and expenditures by fiscal year. Information on End of Year Financial Reports are available at <http://www.doe.mass.edu/finance/accounting/eoy/>. Data from this report allows for several spending comparison reports available as downloadable Excel files at <http://www.doe.mass.edu/finance/statistics/> and in the DARTs.

These reports include instructional and operational expenditures; capital, debt, community activities, and regional assessments are not included. These reports typically are aggregated at the function level and by functional categories, and by funding sources (general appropriations and grants/revolving funds). The Chart of Accounts-Criteria for Financial Reporting used in EOYR is available at: <http://www.doe.mass.edu/finance/accounting/eoy/ChartOfAccounts.docx>

These reports distinguish between expenditures for in- and out-of-district students. In-district students either reside in the district or are tuitioned in from another district. Instructional and operational costs are divided by this number to get per-in-district pupil expenditures. Out-of-district expenditures include tuitions and transportation for students who attend charter schools, students who attend other districts through interdistrict school choice or other tuition arrangements, and students with disabilities served through out-of-district placements. An example of other tuition arrangements would be tuitions that a town might pay to send its students to high school in a nearby district because the town has no high school and doesn’t choose to be a member of a regional high school district. The out-of-district total cannot be properly reported as a per-pupil expenditure because the cost of tuitions varies greatly depending on the reason for going out of district. School choice tuitions are approximately $5,000 by state regulation, while charter school tuitions are approximately the same as a district’s cost per pupil, and a special education placement for a student with complex needs could cost more than $100,000. Two districts with one student each in the same special education placement would have very different costs “per out-of-district pupil” if one of them also sent 100 students to other districts at $5,000 each. School choice and charter opportunities are not equally available to every district’s students and those tuitions vary a great deal by district.

Reports include:

* [**Per Pupil Expenditures, All Funds**](http://www.doe.mass.edu/finance/statistics/per-pupil-exp.html) In addition to showing the overall cost per pupil, this report provides detail about how much districts spend in specific functional areas such as administration, teaching, and maintenance.
* [**Special Education Direct Expenditures Trends**](http://www.doe.mass.edu/finance/statistics/) This report provides details on expenditures for students with disabilities, both in and out of district. Expenditures from general appropriations are reported by placement category, and expenditures from grants are reported as a lump sum.
* [**State Total Spending Trends**](http://www.doe.mass.edu/finance/statistics/) This report provides statewide spending from all funds by functions and functional categories. It also offers calculations of year-over-year change and some charts of the data.
* [**Teacher Average Salaries**](http://www.doe.mass.edu/finance/statistics/) This report includes total spending on teacher salaries divided by full-time-equivalent (FTE) teachers to calculate average teacher salary by district. The report explains which salary data from the EOYR, and which teachers, are included.
* [**District Analysis and Review Tools (DARTs)**](https://www.doe.mass.edu/dart/dart-finance-staff.xlsx) Other finance reports using EOYR data can be found in the DART Detail: Staffing and Finance. This tool provides a wide array of reports: district staffing, teachers, SPED staffing, finance summary, expenditure by source, per-pupil expenditure, out of district expenditures, district revenues, and revolving fund revenue and expenditure. See the [DART Finance User Guide](https://www.doe.mass.edu/dart/dart-finance-user-guide.docx) for descriptions of the data and guidance on potential uses, as well as staffing and finance definitions.

**Charter School End of Year Reports (CSEOYR)**

Charter schools are required to keep an accurate account of all of their financial activities and report their revenues and expenditures. This data is reported slightly differently than the EOYR for municipal and regional districts. Aggregated annual data of the Charter School End of Year Financial Report is available at <http://www.doe.mass.edu/charter/finance/revexp/>. Reports show dollar and per pupil expenditures, as well as revenue. Once the workbook is open, researchers can access worksheets with the raw data. Note: Charter schools do not enroll out-of-district pupils. Additional information about charter school finance available at <http://www.doe.mass.edu/charter/finance/>.

**Additional revenue data, including state aid**

***Chapter 70***

Chapter 70 state education aid is allocated to districts each year largely based on their foundation enrollment and their property and income wealth. The aid ensures that all districts can spend at or above their foundation budgets (the state’s estimate of the minimum needed in each district to provide an adequate educational program). The funds are deposited into local general funds and are appropriated to and budgeted by school committees without any distinction between local and state aid dollars. Information about how Chapter 70 is calculated, links to legislation, state aid, and spending requirements and the following data resources are available at <http://www.doe.mass.edu/finance/chapter70/>:

* + *Chapter 70 district profiles*: For each school district, yearly spending and state aid totals in comparison to the foundation budget, are available going back to FY93.
  + *Chapter 70 Trends in Aid and Local Contribution*: The trend tool shows key factors influencing Chapter 70 required local contributions and state aid calculations going back to FY2007.

***School choice and charter tuition***

Some districts have tuition expenditures and/or revenues from interdistrict school choice. Details of these tuitions, by sending and receiving districts, are available at: <http://www.doe.mass.edu/finance/schoolchoice/>. A significant part of charter school revenues are tuitions paid by municipal districts; details about tuitions by sending district and receiving charter district can be found at <http://www.doe.mass.edu/charter/finance/tuition/>.

***Circuit Breaker (reimbursements for extraordinary special education expenditures)***

The Special Education Circuit Breaker program reimburses districts for per student costs that exceed four times the state average foundation budget per pupil as calculated under the Chapter 70 program, with the goal of the state paying up to 75 percent of the costs above that threshold, subject to appropriation. It also includes a provision that allows districts to claim for extraordinary relief when claimable Special Education costs exceed 125% of the previous year's claimed costs. Worksheets available at <http://www.doe.mass.edu/finance/circuitbreaker/> report Circuit Breaker funding by year.

***Transportation***

Regional districts are reimbursed for a percentage of their transportation expenses set by the legislature in each year’s budget appropriations. Transportation reimbursements data are available at <http://www.doe.mass.edu/finance/transportation/> by fiscal year.

***Grant allocation and awards***

Reports on state and federal entitlement and allocation grants, as well as competitive awards, are available at <http://www.doe.mass.edu/grants/awards.html> and posted in workbooks by fiscal year at <http://www.doe.mass.edu/finance/accounting/eoy/>. Expenditures from grants are reported in aggregated categories, rather than by individual grants, in the EOYR and CEOYR reports.

**Common questions about finance data**

***Student enrollment calculation***

Student enrollment has specific definitions in the finance area that differ from standard reporting. For end-of-year per pupil expenditure reports, FTE pupil numbers are calculated based on days in membership over the whole year. For the state’s Chapter 70 aid program, foundation enrollment is all students enrolled in a district on October 1st of the prior year for which the district is financially responsible because they reside in the district. Student enrollment reported in Profiles is the number of students enrolled in a district (regardless of residence or tuition) as of October 1st each year.

***Teacher FTEs***

Teacher FTEs also have specific definitions in the finance area that differ from standard reporting. The FTE of teachers for the teacher average salary report is calculated from EPIMS (staffing data) but includes teachers who are on salary but not actually teaching for part or all of the year, and teachers that teach post-secondary, post-graduate vocational programs and are not collected through EPIMS. This differs from teacher FTEs reported in Profiles reports, which use data reported as of October 1 each year.

***Fiscal year***

Fiscal years refer to the same time period as school years, July 1 to the following June 30. FY16 or 2016 refers to school year 2015–2016.

***Caution in comparing across data sources***

Finance data analysis requires clarity about what is and is not included across and within the various data reports. For example, municipal/regional districts and charter districts report in separate templates and charts of accounts. Charter data includes capital and debt, whereas municipal/regional expenditure per-pupil reports do not include debt, capital, community activities, and regional assessments. Please contact the Office of Planning and Research for questions about using these data for research purposes.

# 

# XVI. Appendices

# Appendix 1: MCAS test data availability by subject and grade

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table 1. *Summary of MCAS test data availability and mode of delivery by subject and grade.* | | | | | | | | |  |  |  |
| **Grade** | **2007 to**  **2013** | **2014** | **2015** | **2016** | **2017** | **2018** | **2019** | **2020** | **2021** | **2022** | **2023** |
| ***ELA/Mathematics*** | | | | | | | | | | |  |
| 3 | Leg-PB | Leg-PB | Leg-PB | Leg-PB | NG-mix | NG-mix | NG-CB | N/A | NG-CB | NG-CB | NG-CB |
| 4 | Leg-PB | Leg-PB | Leg-PB | Leg-PB | NG-CB | NG-CB | NG-CB | N/A | NG-CB | NG-CB | NG-CB |
| 5 | Leg-PB | Leg-PB | Leg-PB | Leg-PB | NG-mix | NG-CB | NG-CB | N/A | NG-CB | NG-CB | NG-CB |
| 6 | Leg-PB | Leg-PB | Leg-PB | Leg-PB | NG-mix | NG-mix | NG-CB | N/A | NG-CB | NG-CB | NG-CB |
| 7 | Leg-PB | Leg-PB | Leg-PB | Leg-PB | NG-mix | NG-CB | NG-CB | N/A | NG-CB | NG-CB | NG-CB |
| 8 | Leg-PB | Leg-PB | Leg-PB | Leg-PB | NG-CB | NG-CB | NG-CB | N/A | NG-CB | NG-CB | NG-CB |
| 10 | Leg-PB | Leg-PB | Leg-PB | Leg-PB | Leg-PB | Leg-PB | NG-CB | N/A | NG-CB | NG-CB | NG-CB |
| ***Science/Technology*** | | | | | | | | | | |  |
| 5 | Leg-PB | Leg-PB | Leg-PB | Leg-PB | Leg-PB | Leg-CB | NG-CB | N/A | NG-CB | NG-CB | NG-CB |
| 8 | Leg-PB | Leg-PB | Leg-PB | Leg-PB | Leg-PB | Leg-CB | NG-CB | N/A | NG-CB | NG-CB | NG-CB |
| ***Biology*** | | | | | | | | | | |  |
| 9 | Leg-PB | Leg-PB | Leg-PB | Leg-PB | Leg-PB | Leg-PB | Leg-PB | N/A | Leg-PB | NG-CB | NG-CB |
| 10 | Leg-PB | Leg-PB | Leg-PB | Leg-PB | Leg-PB | Leg-PB | Leg-PB | N/A | Leg-PB | NG-CB | NG-CB |
| 11 & 12 | Leg-PB | Leg-PB | Leg-PB | Leg-PB | Leg-PB | Leg-PB | Leg-PB | N/A | Leg-PB | Leg-PB | NG-CB |
| ***Chemistry*** | | | | | | | | | | |  |
| 9 | Leg-PB | Leg-PB | Leg-PB | Leg-PB | Leg-PB | Leg-PB | Leg-PB | N/A | Leg-PB | Leg-PB | Leg-PB |
| 10 | Leg-PB | Leg-PB | Leg-PB | Leg-PB | Leg-PB | Leg-PB | Leg-PB | N/A | Leg-PB | Leg-PB | Leg-PB |
| 11 & 12 | Leg-PB | Leg-PB | Leg-PB | Leg-PB | Leg-PB | Leg-PB | Leg-PB | N/A | Leg-PB | Leg-PB | Leg-PB |
| ***Physics*** | | | | | | | | | | |  |
| 9 | Leg-PB | Leg-PB | Leg-PB | Leg-PB | Leg-PB | Leg-PB | Leg-PB | N/A | Leg-PB | NG-CB | NG-CB |
| 10 | Leg-PB | Leg-PB | Leg-PB | Leg-PB | Leg-PB | Leg-PB | Leg-PB | N/A | Leg-PB | NG-CB | NG-CB |
| 11 & 12 | Leg-PB | Leg-PB | Leg-PB | Leg-PB | Leg-PB | Leg-PB | Leg-PB | N/A | Leg-PB | Leg-PB | NG-CB |
| ***Technology*/*Engineering*** | | | | | | | | | | |  |
| 9 | Leg-PB | Leg-PB | Leg-PB | Leg-PB | Leg-PB | Leg-PB | Leg-PB | N/A | Leg-PB | Leg-PB | Leg-PB |
| 10 | Leg-PB | Leg-PB | Leg-PB | Leg-PB | Leg-PB | Leg-PB | Leg-PB | N/A | Leg-PB | Leg-PB | Leg-PB |
| 11 & 12 | Leg-PB | Leg-PB | Leg-PB | Leg-PB | Leg-PB | Leg-PB | Leg-PB | N/A | Leg-PB | Leg-PB | Leg-PB |
| ***Student Questionnaire*** | | | | | | | | | | |  |
| 4 | -- | -- | -- | -- | -- | -- | VOCAL CB-Math | N/A | VOCAL CB-Math | VOCAL CB-Math | VOCAL CB-Math |
| 5 | -- | -- | -- | -- | VOCAL  PB-STE | VOCAL CB-STE | VOCAL CB-STE | N/A | VOCAL CB-STE | VOCAL CB-STE | VOCAL CB-STE |
| 8 | PB | PB | PB | PB | VOCAL PB-STE | VOCAL CB-STE | VOCAL CB-STE | N/A | VOCAL CB-STE | VOCAL CB-STE | VOCAL CB-STE |
| 10 | PB | PB | PB | PB | VOCAL PB-Math | VOCAL PB-Math | VOCAL CB-Math | N/A | VOCAL CB-Math | VOCAL CB-Math | VOCAL CB-Math |
| ***Retests*** | | | | | | | | | | |  |
| ELA grade 10 | Leg-PB | Leg-PB | Leg-PB | Leg-PB | Leg-PB | Leg-PB | Leg-PB | N/A | Either-PB | Either-PB | Either-PB |
| Math grade 10 | Leg-PB | Leg-PB | Leg-PB | Leg-PB | Leg-PB | Leg-PB | Leg-PB | N/A | Either-PB | Either-PB | Either-PB |
| ***PARCC – ELA/Mathematics*** | | | | | | | | | | |  |
| 3-8 | -- | Field Test | Opt In  PB/CB by school | Opt In  PB/CB by school | -- | -- | -- | -- | -- | -- | -- |
| 9 & 11 | -- | -- | Opt In  PB/CB by school | -- | -- | -- | -- | -- | -- | -- | -- |
| *Note*. For years in which state assessment data can be linked to SIMS; Leg = Legacy MCAS; NG = Next Generation MCAS; PB = Paper-Based; CB = Computer-Based; Mix = schools can choose computer or paper based. In 2017 about 40% students did computer-based, and in 2018 over 60% did computer-based in grades with both paper and computer options; TBD = To Be Determined. N/A = Not Applicable, in 2020 MCAS was cancelled due to the Coronavirus pandemic. Across all years, when tests are computer-based, there may be up to 2,000 students who take paper-based assessment accommodations. The mode of test administration is always present in data files. | | | | | | | | | | |  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Table 2. *Pre-2007 MCAS information on testing subjects by grade* | | | | | | |
| **Discipline** | **2002** | **2003** | **2004** | **2005** | **2006** | **2007 to**  **Present\*** |
| *ELA* |  |  |  |  |  |  |
| 3 | ü | ü | ü | ü | ü | ü |
| 4 | ü | ü | ü | ü | ü | ü |
| 5 |  |  |  |  | ü | ü |
| 6 |  |  |  |  | ü | ü |
| 7 | ü | ü | ü | ü | ü | ü |
| 8 |  |  |  |  | ü | ü |
| 10 | ü | ü | ü | ü | ü | ü |
| *Composition* |  |  |  |  |  |  |
| 4 | ü | ü | ü | ü | ü | ü |
| 7 | ü | ü | ü | ü | ü | ü |
| 10 | ü | ü | ü | ü | ü | ü |
| *Mathematics* |  |  |  |  |  |  |
| 3 |  |  |  |  | ü | ü |
| 4 | ü | ü | ü | ü | ü | ü |
| 5 |  |  |  |  | ü | ü |
| 6 | ü | ü | ü | ü | ü | ü |
| 7 |  |  |  |  | ü | ü |
| 8 | ü | ü | ü | ü | ü | ü |
| 10 | ü | ü | ü | ü | ü | ü |
| *Science/technology* |  |  |  |  |  |  |
| 5 |  | ü | ü | ü | ü | ü |
| 8 |  | ü | ü | ü | ü | ü |
| Biology |  |  |  |  |  | ü |
| Chemistry |  |  |  |  |  | ü |
| Physics |  |  |  |  |  | ü |
| Tech./Eng. |  |  |  |  |  | ü |
| *Retests* |  |  |  |  |  |  |
| ELA grade 10 | ü | ü | ü | ü | ü | ü |
| Math grade 10 | ü | ü | ü | ü | ü | ü |
| Composition | ü | ü | ü | ü | ü | ü |
| \*There are no available MCAS data in 2020 due to the Coronavirus pandemic | | | | | | |

# Appendix 2: Sample SIMS file format

Below are the variable names and corresponding data elements from SIMS23A. The exact order of the variables in the and variable names may slightly vary over time. Updated handbooks and materials related to SIMS data collection changes are posted to the SIMS website and the Information Services research files web page as they are available. Please check these sites for any updates that have been added after this version of the Researcher Guide was compiled.

| **File Heading** | **Element** | **Name** |
| --- | --- | --- |
| PERIOD | -- | Period |
| SASID | DOE002 | State Assigned Student Identifier (SASID) |
| FNAME | DOE003 | First Name |
| MNAME | DOE004 | Middle Name |
| LNAME | DOE005 | Last Name |
| CITY\_OB | DOE008 | City/Town of Birth |
| DOB | DOE006 | Date of Birth |
| LASID | DOE001 | Locally Assigned Student Identifier (LASID) |
| SCHOOL | DOE015 | School Identification Number |
| ATTEND | DOE017 | Days in Attendance |
| MEMBER | DOE018 | Days in Membership |
| TOWN | DOE014 | City/Town of Residence — Student |
| RACE | DOE010 | Race/Ethnicity |
| RFR | DOE011 | Reason for Reporting |
| ENSTAT | DOE012 | Enrollment Status at Time of Data Collection |
| RFE | DOE013 | Reason for Enrollment |
| GRADE | DOE016 | Grade Level |
| LEP | DOE025 | English Learner |
| PLAN | DOE033 | High School Completer Plans |
| IMMIG | DOE022 | Immigration Status |
| FLANG | DOE024 | First (Native) Language |
| CTE | DOE035 | Career/Vocational Technical Education (CVTE) - Type of Program |
| CTE\_COMP | DOE031 | CVTE- Competency Attainment |
| COUNTRY | DOE023 | Country of Origin |
| ORG\_CODE | -- | District Code |
| SPED\_DIS | DOE036 | Special Education - Nature of Primary Disability |
| LVL\_NEED | DOE038 | Special Education - Level of Need |
| FPS\_504\_CODE | DOE039 | 504 Plan Status |
| C74\_SPOP | DOE042 | CVTE - Special Populations |
| C74\_PROG | DOE043 | CVTE - Chapter 74-Approved Vocational Technical Education Program Participation |
| N74\_PRG | DOE044 | CVTE - Non-Chapter 74 Career and Technical Education Program Participation |
| IND\_RECOGNISD\_CREDS\_1\_CODE | DOE047 | Industry Recognized Credential |
| IND\_RECOGNISD\_CREDS\_2\_CODE | DOE048 | Industry Recognized Credential |
| IND\_RECOGNISD\_CREDS\_3\_CODE | DOE049 | Industry Recognized Credential |
| EARLY\_CHILDHOOD\_EXP\_CODE | DOE050 | Early Childhood Experience |
| TITLE1 | DOE020 | Title I Participation |
| ELP\_PROG | DOE026 | English Language Learners Program Status |
| HQCPP\_TYPE\_CODE | DOE045 | HQCP Program Type |
| HQCPP\_PARTICIPATION\_CODE | DOE046 | HQCP Program Participation |
| TRUANT | DOE052 | Unexcused Student Absences |
| GENDER | DOE009 | Gender |
| SPEDEVAL | DOE040 | Special Education - Evaluation Results |
| GRAD\_CORE | DOE037 | Graduate, Completed Massachusetts Core Curr. |
| ALT\_ED | DOE027 | Program Code |
| SPED\_3TO5 | DOE032 | Special Education Placement, ages 3-5 |
| SPED\_6TO21 | DOE034 | Special Education Placement, ages 6-21 |
| SEAL\_OF\_BILITERACY\_CODE | DOE028 | Seal of Biliteracy |
| LEP\_ARRIVE | DOE021 | EL Students in their First Year in U.S. Schools |
| PERIOD\_NAME | -- | Description of Period |
| ORG\_ID | -- | First four of School ID (DOE0015) - Numeric |
| MD\_MILITARY\_FAMILY | DOE029 | Member of Military Family |
| ZIP\_CODE | DOE051 | ZIP Code |
| TITLE\_I\_NON\_INST | DOE030 | Non-Instructional Title I Targeted Assistance Services |
| SLIFE\_CODE | DOE041 | English Learner Student with Limited or Interrupted Formal Education (SLIFE) |
| CIVICS\_PROJ\_COMP\_CODE | DOE053 | Civics Project Completer |
| DAYS\_ATTENDED\_REMOTE | DOE054 | Days of Attendance Remote |
| DAYS\_ABSENT\_REMOTE | DOE055 | Days Absent Remote |
| SUPP\_LOW\_INCOME\_IND | DOE056 | Supplemental Low-Income Indicator |
| LOW\_INCOME | DOE019 | Low-Income Status |

# Appendix 3: Standard template for data-sharing agreements

**MEMORANDUM OF AGREEMENT BETWEEN**

**THE MASSACHUSETTS DEPARTMENT OF**

**ELEMENTARY AND SECONDARY EDUCATION**

**AND**

**ORGANIZATION**

1. **Parties**
   1. This Memorandum of Agreement (“the Agreement”) is entered into by and between the Massachusetts Department of Elementary and Secondary Education (“the Department”) and ORGANIZATION (“ORG SHORT NAME”).
   2. The Massachusetts Department of Elementary and Secondary Education is a state educational agency, authorized to collect and maintain student educational records and to receive information from local educational agencies (LEAs) consistent with applicable state and federal laws and subject to the federal Family Educational Rights and Privacy Act (FERPA), as authorized by 20 U.S.C. § 1232g (b) and 34 CFR Part 99, and the Fair Information Practices Act (FIPA), M.G.L. c. 66A.
   3. The Department designates Matthew Deninger (Matthew.J.Deninger@mass.gov, 781-338-3117) as its liaison for all communications with ORGANIZATION regarding this project and the Agreement as it relates to the project.
   4. [BRIEF DESCRIPTION OF ORGANIZATION]
   5. The ORGANIZATION designates LIAISON (EMAIL, PHONE) as its liaison for all communications with the Department regarding this project and the Agreement as it relates to the project.
2. **Purpose**
   1. The purpose of the Agreement is:
      1. to document the terms under which the Department is authorized to release to ORGANIZATION personally identifiable student information for the purpose of [describe the project] (hereinafter, “the project”); and
      2. to designate ORGANIZATION as the temporary custodian of the Department’s personally identifiable student information consistent with applicable federal and state laws concerning access to and confidentiality of student record information, including FERPA and FIPA.
   2. As described herein, ORGANIZATION, as authorized by the Department, may have temporary access to data in the custody of the Department for use in the project identified in this Agreement and any addenda to it. A XX-page description of the project is attached hereto.
3. **Authority**
   1. Consistent with FERPA and 34 CFR § 99.31(a(6)(i), the Department is authorized to release to ORGANIZATION personally identifiable student information for the purpose of the project and in connection with the Department’s conducting studies to [develop, validate, or administer predictive tests OR administer student aid programs OR improve instruction].
   2. This release is also consistent with the Fair Information Practices Act, M.G.L. c. 66A, § 2(c) (FIPA). [For any agreement releasing only non-student records, reference only FIPA authorization in this section]
4. **Data shared** 
   1. The Department shall provide ORGANIZATION with the following data for the purpose(s) described in section II:

FOR EXAMPLE

* + 1. All SIMS data elements from school years X to Y, excluding name and date of birth
    2. All SCS data elements from school years X to Y, excluding student name and date of birth
    3. All SSDR data elements from school years X to y, excluding student name and date of birth
    4. All MCAS and MCAS-Alt data elements from school years X to Y, excluding student name and date of birth
    5. All SAT and AP data elements from school years X to Y, excluding student name and date of birth
    6. All National Student Clearinghouse data elements for graduates and non-graduates from the years X to Y, excluding student name and date of birth
  1. This Agreement does not convey ownership of data to ORGANIZATION.
  2. All copies of data of any type, including any modifications or additions to data from any source that contains information regarding individual students, are subject to the provisions of the Agreement and in the same manner as the original data disclosed by the Department to ORGANIZATION.
  3. Nothing in the Agreement shall be construed to authorize ORGANIZATION to have access to additional data from the Department that is not included in the scope of the Agreement, or to govern access to the data by entities other than the Parties.

1. **designation of temporary data custodian**
   1. For purposes of this project and for ensuring ORGANIZATION’s compliance with the terms of this Agreement and all applicable state and federal laws, ORGANIZATION designates CUSTODIAN (EMAIL, PHONE) as the temporary custodian of the Department’s data. The Department will release all data and information for this project to the named temporary custodian.
   2. CUSTODIAN shall be responsible for transmitting all data requests and maintaining a log or other record of all data requested and received pursuant to the Agreement, including confirmation of the completion of the project and the return or destruction of data as described below.
   3. The Department or its agents may upon request review the records required to be kept by ORGANIZATION under this Agreement.
2. **Terms and Conditions**

To effect the transfer of data and information that is subject to state and federal confidentiality laws and to ensure that the required confidentiality of personally identifiable information shall always be maintained, ORGANIZATION agrees to the following terms and conditions.

* 1. **Legal Compliance**
     1. ORGANIZATION agrees to comply in all respects with the provisions of FERPA. For the purposes of the Agreement, FERPA includes all requirements of 34 CFR Part 99 and 20 U.S.C. § 1232g, as well as any amendments or other relevant provisions of federal law.
     2. ORGANIZATION agrees to comply in all respects with the provisions of FIPA, as ORGANIZATION is a “holder” of personal data for the purposes of FIPA, a copy of which is available at <http://www.malegislature.gov/Laws/GeneralLaws/PartI/TitleX/Chapter66A>.
     3. ORGANIZATION agrees to require all employees, contractors, and agents of any kind to comply with the Agreement and all applicable provisions of FERPA, FIPA, and other federal and state laws with respect to the data and information shared under the Agreement.
     4. ORGANIZATION agrees to require of and maintain an appropriate confidentiality agreement from each employee, contractor, or agency with access to data pursuant to the Agreement.
     5. Nothing in this Agreement may be construed to allow either party to maintain, use, disclose, or share student record information in a manner not allowed under federal or state laws or regulations.
  2. **Security Protocols**
     1. ORGANIZATION agrees to implement administrative, physical and technical safeguards that reasonably and appropriately protect the confidentiality, security, and integrity of data obtained pursuant to the Agreement and that prevent use and disclosure of such data other than as permitted under the Agreement.
     2. ORGANIZATION agrees to store data using industry-standard encryption and authentication to ensure that only authorized agents of ORGANIZATION have access to the data.
     3. ORGANIZATION agrees not to access, take, maintain, or store data outside the United States.
     4. ORGANIZATION agrees to immediately inform the Department when ORGANIZATION either (1) knows, or has reason to know, of a breach of security with respect to the data provided under this Agreement, or (2) knows, or has reason to know, that any of the data described in Section 4 was acquired or used by an unauthorized person or used for an unauthorized purpose.
        1. Notice shall be given by telephone and by email to Matthew Deninger (Matthew.J.Deninger@mass.gov, 781-338-3117), and to the Department Assistant General Counsel Lucy Wall, [lucy.a.wall@mass.](mailto:lucy.a.wall@mass.)gov, (781) 338-3400.
        2. In addition to informing the Department as provided herein, ORGANIZATION shall cooperate with the Department and such other state and federal entities as necessary in the investigation and further reporting of such breach.
  3. **Permissible Use of Data**
     1. ORGANIZATION agrees to use data shared under the Agreement for no purpose other than the research project described in this Agreement, and as authorized under 34 CFR §§ 99.31(a)(6)(i).
     2. ORGANIZATION agrees to not amend or alter the scope, design, format, or description of a project or report generated by ORGANIZATION for this project, except as consistent with the Agreement, without prior written approval of the Department.
  4. **Access to Data**
     1. ORGANIZATION agrees not to share data received under the Agreement with any other entity without prior written approval from the Department.
     2. The ability to access or maintain data under the Agreement shall not under any circumstances transfer from ORGANIZATION to any other individual, institution, or entity.
     3. ORGANIZATION agrees not to copy, reproduce, or transmit data obtained pursuant to the Agreement except to ORGANIZATION’s own employees, contractors, or agents acting for or on behalf of the Department and as necessary to fulfill the purpose of the project described herein.
  5. **Preserving Confidentiality**
     1. ORGANIZATION agrees not to disclose data contained under the Agreement in it in any manner that could identify any individual student, except as authorized by FERPA, to any entity other than the Department, or authorized employees, contractors, and agents of ORGANIZATION working on the project approved by the Department consistent with this Agreement.
     2. ORGANIZATION shall neither disclose or otherwise release data and reports relating to an individual student, nor disclose information relating to a group or category of students without ensuring the confidentiality of individual students in that group.
     3. Publications and written reports of this data and information related to it, including preliminary project descriptions and draft reports, shall involve only aggregate data and no personally identifiable information or other information that could lead to the identification of any student.
     4. No report of aggregate data based on an identifiable group of students fewer than ten (10) in number shall be released to anyone other than the Department. ORGANIZATION shall require that all employees, contractors, and agents working on this project abide by that statistical cell size.
     5. The Department shall provide ORGANIZATION with reasonable notification of any changes in Department policies regarding limits on the use of confidential data. As an agent of the Department, ORGANIZATION shall either affirmatively agree to uphold these policies or relinquish access to the data.
  6. **Publication**
     1. In order to protect the confidentiality of personally identifiable information disclosed to ORGANIZATION, ORGANIZATION agrees to provide the Department any proposed publications or presentations which are to make public any findings, data, or results of the research produced under this Agreement for the Department’s review at least thirty (30) days prior to submission of a manuscript or abstract for publication, or at least thirty (30) days prior to the date of the presentation.
     2. ORGANIZATION agrees to provide the Department with an electronic copy of the final versions of all reports and other documents associated with the project and, at the Department’s request, with an electronic copy of a nontechnical summary of findings for potential public distribution. The Department reserves the right to distribute and otherwise use the final report and associated documents as it wishes, in sum or in part.
     3. ORGANIZATION has the right to publish, present, or use the study results it has gained in the course of the research under this Agreement. ORGANIZATION or its agents working on this project, retain the right to publish findings in other publications, provided it first shares a substantively similar report with the Department.
  7. **Destruction of Data**
     1. ORGANIZATION agrees to destroy all data obtained under the Agreement and addenda to it in a manner consistent with industry standards when no longer needed for the purpose for which the Department disclosed the data.
     2. Nothing in this Agreement authorizes ORGANIZATION to maintain data beyond the time period reasonably needed to complete the project and respond to inquiries from other researchers, which shall be DATE.
     3. ORGANIZATION agrees to require all employees, contractors, or agents working on the project to comply with this provision. No other individual or entity is authorized to continue research using the data obtained under the Agreement upon the termination of the Agreement and project described herein.

1. **related parties**
   1. ORGANIZATION represents that it is authorized to bind to the terms of the Agreement, including security, confidentiality, maintenance, publication, and destruction of data disclosed by the Department to it, all related or associated institutions, individuals, employees, contractors, or other agents who may have access to the data or may own, lease, or control equipment or facilities of any kind where the data is stored, maintained or used in any way.
2. **Term**
   1. This Agreement takes effect upon the signature by the representative of each party named below and shall remain in effect until the data access time period defined in paragraph 10 of section V has ended, or until canceled by either party upon 30 days written notice, whichever occurs first. The Agreement is renewable upon written approval by the representative of each party.
3. **Entire Agreement**
   1. This Agreement expresses the entire agreement of the parties and shall not be modified or altered except in writing executed by the representatives of the Department and ORGANIZATION authorized to bind the parties, and in a manner consistent with applicable state and federal laws.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Matthew Deninger Date

Chief Strategy and Research Officer

Massachusetts Department of Elementary and Secondary Education

The signatory below on behalf of ORGANIZATION is authorized to bind ORGANIZATION to the terms of the Agreement.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

SIGNATORY NAME Date

SIGNATORY TITLE

ORGANIZATION

Attachment: Project description

# Appendix 4: Data Impacts of Coronavirus

|  |  |
| --- | --- |
| **Dataset** | **Changes** |
| **SIMS** | SY2020:   * Attendance for the 2019-2020 school year should end with the SIMS-C collection. * The initiation of remote learning during spring 2020 will impact all data in the SIMS-D collection. For example, attendance will reflect differences in district and school closures in response to the pandemic, as well as differences in district policies around remote learning attendance. |
| SY2021:   * There are two new variables: attendance during remote learning and absences during remote learning. |
| SY2022:   * Remote learning attendance collected. |
|  | |
| **SCS** | SY2020:   * SCS was collected as normal for the 2019-2020 academic year. * 9th grade passing rates in comparison to previous years can be found online. |
| SY2021:   * No known impacts. |
| SY2022:   * No known impacts. |
|  | |
| **MCAS** | SY2020-2023:   * The science testing requirement has been waived for the graduating classes of 2021, 2022, and 2023. This is because not all students will have equal access to taking the science exam at the end of their respective time in the corresponding course. Specifically, 2018-2019 9th graders took science, 2019-2020 9th graders did not, 2020-2021 9th graders will take it. |
| SY2020:   * All mandatory MCAS testing in SY2020 was cancelled. * The 2019-2020 10th grade ELA and math MCAS was not taken. * The Competency determination was waived for the class of 2020, which resulted in fewer retests and increased the graduation rates for 2020.   + 12th graders in the 2019-2020 academic year were given the option to meet their competency determination (CD) for graduation based on coursework. * **Because the 2020 MCAS was cancelled, there are no 2020 MCAS data available for research purposes.** |
| SY2021:   * The 2020-2021 10th grade ELA and math MCAS was administered in spring. * MCAS in 2021 had 94% participation in Grades 3-8.   + 17% of tests were given remotely.   + Results for the grade 8 STE MCAS remote administration were unable to be scored, and therefore scores were not provided for those students. * In grade 10 there was no remote testing option for ELA and Math.   + Participation was 91% * MCAS testing window for ELA did not start until April and was allowed to end later than in previous years. * There were differences in testing sessions for Grades 3-8. ([Link](https://www.doe.mass.edu/mcas/2021/results/faq.docx)) * Student Growth Percentiles (SGP) were calculated differently than in prior years. ([Link](https://www.doe.mass.edu/mcas/2021/results/sgp.docx)) |
| SY2022:   * MCAS administration returned to pre-pandemic schedule and SGP was calculated as it had previously. |
|  | |
| **VOCAL** | SY2020:   * VOCAL not administered. |
| SY2021:   * VOCAL was administered in grades 4, 5, 8, &10 with MCAS (grades 4 and 10 with mathematics; grades 5 and 8 with science). * Participation across all grades was lower compared to SY2019 (G4: 82% to 65%; G5: 85% to 67%; G8: 85% to 63%; G10: 76% to 51%) * Students participated across different modes of learning (in-person, hybrid, remote).   + In grades 4 and 5, approximately two-thirds of students reported they were in in-personal learning, 19% remote, and 13% hybrid.   + In grade 8, 8% of student reported they were remote, 46% in-person, 23% hybrid, and 23% transitioned from one mode of learning to another in the previous 30 days.   + In grade 10, 9% of student reported they were remote, 30% in-person, 32% hybrid, and 29% transitioned from one mode of learning. |
| SY2022:   * VOCAL was administered in grades 4, 5, 8, & 10 with MCAS (grades 4 and 10 with mathematics; grades 5 and 8 with science). |
|  | |
| **NSC** | SY2020:   * No known impacts. |
| SY2021:   * No known impacts. |
| SY2022:   * No known impacts. |
|  | |
| **WAGE** | SY2020:   * No known impacts. |
| SY2021:   * No known impacts. |
| SY2022:   * No known impacts. |
|  | |
| **SSDR** | SY2020:   * Collection was the same. * Comparisons can be determined from data published online. * Lower overall compared to SY2019, with very few disciplines from March 2020 onward. |
| SY2021:   * Many fewer disciplines compared to SY2020. * Data published online for comparisons. |
| SY2022:   * Number of students disciplined SLIGHTLY below pre-pandemic levels. |
|  | |
| **AP, SAT, and ACT** | SY2020:   * ~6% decrease in AP test takers as 2019. * SAT had about half as many testers as 2019. |
| SY2021:   * AP – 10% fewer tests taken vs. 2019 * SAT - ~25 point increase in both reading/writing and math vs. 2019. Anecdotal evidence suggests that many colleges dropped SAT as an admissions requirement, but high-achieving students opted to test regardless which drove up scores. Also, number of tests taken was ~60% of 2019 count. |
| SY2022:   * AP: Slightly below pre-pandemic number of tests taken * SAT: Number of tests taken was 70% of 2019 count. Average scores decreasing from 2021, but still a few points higher in both subjects vs. 2019. |
|  | |
| **ELAR** | SY2020: License expirations were extended through state of emergency. |
| SY2021:   * Addition of “BA Verification” * Addition of Emergency license. ([Link](https://www.doe.mass.edu/licensure/)) * CAP data includes a new option for indicating completion via an alternative assessment due to the state of emergency. |
| SY2021:   * Emergency licenses were available for teachers and administrators through December 12, 2021 ([Link](https://www.doe.mass.edu/licensure/)) |
| SY2022:   * Emergency licenses were available for teachers and administrators from May 26, 2022 through September 22, 2022. ([Link](https://www.doe.mass.edu/licensure/)) * ELAR will include information about candidates who passed an alternative assessment as part of our [pilot](https://www.doe.mass.edu/news/news.aspx?id=26384). |
|  | SY2023:   * Emergency licenses were extended through June 30, 2024. ([Link](https://www.doe.mass.edu/news/news.aspx?id=26944)) |
| **EPIMS** | SY2020:   * Educator Evaluation ratings were impacted – with many fewer evaluations statewide than previous year. * Comparisons to previous years can be found online. |
| SY2021:   * Addition of Emergency license. * Out-of-Field calculation may be impacted (TBD). * Changes to work assignments for SY2021 may be impacted (staffing numbers and capacity will impact assignments; TBD). * Some teachers were assigned to virtual district school codes with IDs ending in “VK12”. This assignment may not match up with the school codes that students are assigned to. |
| SY2022:   * TBD |
|  | |
| **Financial** | SY2020:   * No known impacts. |
| SY2021:   * No known impacts. |
| SY2022:   * No known impacts. |
|  | |
| **ACCESS** | SY2020:   * No known impacts. * Collected and scored as normal. |
| SY2021:   * Given in January of 2020, participation was below 90%. |
| SY2022:   * No known impacts. |
|  | |

# Appendix 5: Additional DESE Policies for Research and Evaluation:

****

Policy Regarding the Use of Open Science Practices – June 2023

The following policy regarding the Use of Open Science Practices offers guidance to researchers who want to engage in open science practices in any Massachusetts Department of Elementary and Secondary Education (Department or DESE) sponsored research activities. The policy covers all research activities, whether conducted by employees of the Department or by external research partners.

**Definitions**

* *Open Science:* Open science is the movement to make scientific research and its dissemination accessible to all levels of society. Open science encourages transparency, reproducibility, and accessible knowledge that is shared and developed through collaborative networks. Research is considered open science if access to various research materials, including publications, data, protocols, and software are provided.

**Background**

The Department regularly engages in research activities to evaluate education-related programs and policies. Research activities are conducted by DESE staff and via publicly procured research contractors and may involve public employees as participants (e.g., teachers, administrators, paraprofessionals).

Open science practices among researchers are growing in an effort to encourage transparent, replicable research across many fields of study. These practices inform many aspects of the research process including documentation of analytic code, pre-registration of analyses, open access publications and pre-prints, and documentation of protocols and decisions within the scientific process.

**The following criteria govern the acceptable provision of open science practices when conducting research with DESE confidential data.**

*Code Files*

All code files used to generate analyses based in full or in part on DESE confidential data may be published under the following conditions:

* All personally identifiable information, such as names and dates of birth, must be redacted from any public facing code file.
  + This policy pertains to any Department-associated work, including data scraped from the internet or other public materials.

*Data Files*

No confidential data files may be shared. When needed, researchers may link to publicly reported data presented on the DESE website that may be used to replicate findings.

*Other Study Materials*

Other study materials including but not limited to survey, focus group, or interview protocols may be shared publicly.

*Preprints*

After you have received data security and accuracy review approval from the Department, researchers are welcome to share preprints of their academic publications, in line with the terms set in their data sharing agreement.

*Pre-registration*

After their proposals have been approved by the Department, researchers are welcome to pre-register their analyses. We encourage researchers to use the proposal review process to sharpen their pre-registered analyses in light of the institutional, data, and policy expertise held by Department employees.

1. For several years in the mid 2000s, an additional administration of SIMS that only included special education students occurred on December 1, but this has been discontinued. [↑](#footnote-ref-2)
2. This includes two collections that were done during a pilot year. The first full year of SCS is for the 2010–11 school year. [↑](#footnote-ref-3)
3. The 3,600 colleges and universities participating in NSC enroll over 98% of all students attending public and private postsecondary institutions nationally. [↑](#footnote-ref-4)
4. The 2007 graduate file would include all students identified as a graduate in 2007, not just those who were part of a 4- or 5-year graduation cohort. It does not contain information about graduates from previous years. [↑](#footnote-ref-5)