



Guidelines for Requesting and Accessing Confidential U.S. Energy Information Administration Data

June 15, 2026

The U.S. Energy Information Administration (EIA), the statistical and analytical agency within the U.S. Department of Energy (DOE), prepared this report. By law, our data, analyses, and forecasts are independent of approval by any other officer or employee of the U.S. Government. The views in this report do not represent those of DOE or any other federal agencies.

Table of Contents

About these Guidelines.....	1
General Guidance in Preparing an SAP Application.....	1
Additional Guidance in Applying for Access to EIA Data	1
Data selected	1
Work location	1
Software requirements	1
Research team	2
Research project	2
Project duration	2
Funding	2
Timeline	2
Research question	2
Demonstrated need	3
Project abstract	3
Data linkages	3
User-provided data	4
Methodology	4
List of references	5
Project products	5
Requested output and disclosure review	5
Procedures for Data Access After EIA’s Approval of the Application	6
CIPSEA training and non-disclosure agreement	6
Data access agreement	6
Conducting research at EIA Headquarters	6

About these Guidelines

If you wish to conduct research using confidential U.S. Energy Information Administration (EIA) data, you must submit a research application to EIA through the [Standard Application Process \(SAP\) Portal](#). This document provides general guidance in preparing an SAP application, additional guidance in applying for access to EIA data, and an overview of the procedures that are required to access EIA confidential data after your application has been approved.

General Guidance in Preparing an SAP Application

The National Center for Science and Engineering Statistics (NCSES) provides [general SAP information](#) to guide you in creating an application to request data, to get expert help with your application, and to get answers to questions about the SAP process and what to expect. The SAP Applicant User Guide, which is available in the SAP Portal, will walk you through the process of completing an application and submitting it in the SAP Portal for agency review.

Additional Guidance in Applying for Access to EIA Data

We provide the following additional guidance for completing an SAP application for access to EIA's confidential data. This guidance provides information on how to input the data selected and describe your research team and research project. For questions on completing an application for access to EIA's confidential data, email eiainfocipsea@eia.gov. More information about SAP and access to EIA confidential data is available on the EIA [website](#).

Data selected

The first step to create an application using our data is to select the required datasets for your project by adding them to your basket in the [SAP Portal](#). When selecting EIA datasets, do not select datasets from any other U.S. statistical agency or unit because these datasets will not be available onsite at EIA Headquarters for data linkage. If you select datasets from any other U.S. statistical agency or unit, then EIA will reject your application.

You are encouraged to review all information available for each dataset in the SAP Data Catalog on the SAP Portal and use it when completing the “provisioned by” fields of the application that describe the specific years and geography that you need for your project. If your project is approved, EIA will use these fields of the application to prepare the data for your project. If you need more information about an EIA dataset, email eiainfocipsea@eia.gov. After selecting all datasets needed for the project, the next step is to start a new application in the SAP Portal.

Work location

The SAP lists work locations used by all agencies, but only *EIA Onsite* access modality may be selected for projects that access EIA data.

Software requirements

We have statistical software ready to use in the computing environment, such as SAS, R, Python, etc. You can find available software applications in the SAP Data Catalog. Note that you will not be able to access the internet from the computers you will be using to access confidential data onsite at EIA.

Research team

You must list the names of the principal investigator (PI) and all other members of the research team, including collaborators who do not plan to access the confidential data onsite at EIA Headquarters but who will participate in planning, oversight, or dialogue about the EIA data used in your project prior to disclosure review. All researchers must be U.S. citizens and provide their curriculum vitae (CV) in the research team section of the application. Only two members of the research team will be allowed onsite access to the EIA confidential dataset(s) directly at the same time. Each graduate student will be required to submit a letter of support from their research institution to EIA, in addition to the information provided on the application in the SAP Portal.

Research project

You will complete the pages of the SAP application that describe your research project by focusing on the project details. Each field or section must be written so that it can be understood by a statistician who is not necessarily a specialist in the topic within the field of study. Keep in mind that the audience for an EIA research project proposal is not the same audience that researchers typically address when writing research applications or journal articles.

Project duration

We encourage you to carefully assess the time period for which you request data access and ensure the intended research can be completed in this timeframe. Requests for access extensions beyond the specified end date undergo careful scrutiny, must be justified, and generally are granted only for circumstances beyond the control of the researchers (for example, unexpected illness). The maximum length of time you may access a confidential EIA dataset is 60 months, though a typical project duration is 12 to 36 months. Note that you may discover that it takes longer than anticipated to gain access and become familiar with the data and computing environment.

Funding

List sources of funding for the research project overall. We do not charge researchers for use of the data. We provide a secure environment for researchers to access data onsite while administering confidentiality protections.

Timeline

Using the Project Timeline Template that is available for download in the SAP application, upload a table with a proposed timeline for the project, including a column for each year and a row for each task. An example is given in Table 1 below. The timeline must include activities such as completing confidential data access training, establishing EIA onsite access, cleaning data, merging data, and conducting specific analyses, as well as an estimated timeframe for outputs, such as disclosure review of datasets intended for use in publications or presentations.

Research question

Provide an overall description of the research question. This section must describe the purpose of the project including any key concepts that will inform the research methods or models and a general description of the data used in the project.

Demonstrated need

Discuss why the project needs non-public data. Explain in detail why the project’s research question can only be addressed using the requested confidential data. Be specific on why the public use microdata files that EIA publishes are not sufficient for your use. This includes listing which variables in the confidential file will be used and what methodological advantages are perceived by using these variables in the confidential file, instead of the public use microdata file. You will be able to use variables from both the confidential and public use microdata files when conducting your research at EIA.

Table 1. Example of a Proposed Project Timeline

Task	Year		
	1	2	3
Researcher submits request in SAP to access EIA data	x		
Researcher completes CIPSEA training	x		
Researcher establishes onsite access to EIA confidential data	x		
Researcher conducts research	x	x	x
Researcher prepares statistical aggregates for disclosure review			x
Researcher discusses any required changes required to the data table resulting from the disclosure review			x
Researcher requests expert review of data for release			x
Researcher submits and completes paper and submits to journal.			x

Project abstract

Project abstracts must range between 150 and 250 words in length. This section must reference major datasets, key methods, and discuss expected findings. You may include one or two sentences about why the research is important in your field of study. The project abstract should capture the essence of the project application.

Data linkages

If you plan to use non-EIA data in the project (see the next section for guidance on user-provided data), upload a document summarizing how the data used in your project will be linked and any expected overlap, including expected sample sizes (if applicable). State the datasets to be linked, the unit at which the linkages will occur (for example, zip code, building type, etc.), the purpose of the linkage, and information on how the linkages will be performed.

It is important to describe how and at what unit of observation or aggregation datasets will be linked. Linking survey data to other survey data may lead to small sample sizes (and therefore potential disclosure issues) if the survey samples do not significantly overlap. When linking confidential data to externally provided contextual variables, detail the level of geography (for example, county, state, etc.) at which this linkage will be performed. You may use our public use microdata to help evaluate subpopulation sample sizes prior to submitting a research application.

User-provided data

If you plan to use non-EIA data in the project, enter information about each dataset. The list must be comprehensive. Enter the dataset name, a short description of the dataset, the format and approximate size of the file, dataset source, and whether the data are publicly available or proprietary (and can be shared with EIA). If available, include the link where the data can be downloaded.

Be prepared to provide these data to us prior to workspace access so we may upload all user-provided data to the computer that you will use to access EIA confidential datasets. Future requests for user-provided data not listed in the application will require additional justification and approval, delaying project timelines.

Methodology

This section describes the statistical equations, if any, that will be estimated, what key variables are needed and how they will be measured, and how the data from one or more datasets will be used. It is important in this section to describe the methods used to answer the research question. For a complete list of available variables for EIA confidential datasets, download codebooks from the SAP Data Catalog.

The methodology section must clearly justify your access to each requested dataset and state the role of each dataset in the analysis. Data cleaning and crafting is an essential element of research design. The application needs to be as detailed as possible in describing how measures will be constructed and how the data will fit together to construct the datasets used for analysis. Generally, this is accomplished by discussing how variables are constructed and from which datasets they are retrieved.

You must show that you have read the publicly available documentation about the datasets you are requesting and have considered implications of the data in relation to your research approach. Sources of information about datasets that are linked to the SAP Data Catalog include the website link for survey program information. EIA [survey questionnaires](#) and methodology documentation are available on the EIA website.

The units of observation for the analysis and the groups of units for which researchers will carry out analysis are both significant. For each, be specific. Is the unit of observation, for example, a commercial building, or will the unit of observation be at a more aggregate level (for example, census division)? Will the analysis use some combination of units? Will the researcher conduct separate analysis by region, building type, etc.? If the project uses user-provided data, how large are the datasets that will be analyzed? Information that describes the units of analysis and the groups of units/levels of dataset(s) the project will be using are necessary for EIA reviewers to assess both the feasibility of the project and the risk of disclosing confidential information.

Clearly specify all models that will be used for analysis. You do not need to give the exact functional form you plan to use, but you must write out, in as much detail as possible, any equations and relevant variables and discuss how you will fit the data to the models. If EIA confidential data include sample weights, researchers should describe how weights will be used.

List of references

To demonstrate the scientific merit of the research, provide a brief, targeted discussion of the relevant literature, focusing on the key concepts and data to be analyzed. Include a list of references in the application.

Project products

Describe the deliverables that you expect to produce, such as journal articles, conference presentations, technical reports, data tables, and other expected products. Include the names of journals targeted for publication.

Requested output and disclosure review

Describe the output you expect to release, and include table shells. This description is important in assessing both the substance of the application and the risk of disclosing confidential information. Output from your research using confidential data must focus on aggregate values (for example, totals, averages, regression coefficients, and relative standard errors [RSEs]) and be consistent with the approved project objectives. Types of output expected to be released include model parameters and aggregate statistics. You should share what type of publication standards you plan to apply. For example, we do not publish commercial building estimates where the RSE (coefficient of variation) exceeds 50% or the count of buildings used for the calculation is less than 20. Prior to submitting an application, you may use the public use microdata file to evaluate subpopulation sample sizes.

To protect confidential data, any products created from internal data are confidential and limited to onsite access at EIA until they have undergone disclosure review and have been approved for release by EIA.

EIA looks for several things in evaluating applications for disclosure risk. The project application must clearly state the proposed methods and output. Reviewers understand that, because the research team is conducting research, it is unlikely that they know all the details in advance. Nevertheless, EIA reviewers need enough detail to assess whether the proposed project can succeed without posing disclosure risk.

Disclosure risk is a complex topic, and researchers will need to discuss their proposed output with their research team as they develop their application. Researchers may also contact eiainfocipsea@eia.gov with disclosure-related questions specific to the requested confidential data.

When you are ready to release your output, you will need to submit a confidentiality package describing how the output fits within the scope of your application, underlying data, and variables. Plan for six weeks after the output is produced to receive your cleared results. This period allows for initial review by EIA and possible requests to revise your output or supporting documentation prior to final review by EIA. Short, simple requests take less time to prepare and review, while requests that require more detailed review due to higher disclosure risk will take longer. Output that includes only signs and statistical significance is reviewed more quickly.

Procedures for Data Access After EIA's Approval of the Application

All information EIA requires to apply for access to confidential data is included in the SAP application. Once a research proposal is approved, however, EIA requires additional documentation for researchers to become temporary, unpaid agents of EIA to gain onsite access to confidential data at EIA. This section details the procedures for obtaining data access after a research proposal has been approved.

CIPSEA training and non-disclosure agreement

The Confidential Information Protection and Statistical Efficiency Act (CIPSEA) is a federal law (Title 5 of the E-government Act of 2002, reauthorized in 2018) that provides confidentiality protection for some data collected by federal agencies for statistical purposes. All persons authorized under CIPSEA to access EIA confidential data are required to take annual CIPSEA training. The training materials explain, with examples, the confidentiality constraints and the consequences of violating CIPSEA. After researchers complete the training, they must submit a certificate of completion to EIA personnel. Each researcher accessing confidential EIA data must also submit a signed non-disclosure agreement (NDA). The CIPSEA training and NDA must be completed before a Data Access Agreement can be signed between EIA and your research institution. Also, your research institution must be a U.S. educational institution or nonprofit entity, and you may only use the data for statistical purposes.

Data access agreement

Before data access can be granted, a data access agreement (DAA) must be signed between EIA and your research institution. EIA will prepare a draft DAA and submit it to the principal investigator for the project. The draft will include a description of the data to be shared, a list of researchers who will access the data (names provided in SAP), responsibilities of agents of EIA, and conditions for access. The DAA may be prepared for a term not to exceed five years. If the research has not been completed at the end of the term, a new DAA will be required to continue work on the project.

If your research institution agrees to sign the draft DAA, a concurrence process takes place in which officials from EIA and your research institution sign the DAA. If your institution requests modifications to the draft, your institution must negotiate the changes with EIA. Several drafts with various modifications may be exchanged. The concurrence process moves forward for signature only after the negotiations produce an agreement that is acceptable to both parties.

Conducting research at EIA Headquarters

After the DAA has been signed by both parties, researchers may make an appointment to enter EIA Headquarters (1000 Independence Ave., S.W., Washington, DC 20585) as visitors and perform research using the confidential EIA data. Visits must take place Monday through Friday between the hours of 8:00 a.m. and 5:00 p.m. No access to EIA Headquarters will be available on federal holidays or days when the building is closed. No more than two researchers from one team (one research proposal) may enter EIA Headquarters at a time.

All researchers must have a government-issued form of identification (for example, driver's license, passport) to enter EIA Headquarters. EIA will provide laptops, other electronic equipment, and software as arranged in advance. Researchers will not have access to the internet, printers, copy machines, cell

phones, or personal devices. The designated area for conducting the research must always remain unlocked and accessible by EIA staff. After completing the project, researchers will submit aggregate data files to EIA for disclosure review, as described above. After disclosure review, EIA will transfer the approved tabulations to the principal investigator in the agreed-upon file format.