Laboratory for Microdata Analysis
ADELE

User guide

The ADELE Laboratory for Microdata Analysis is a *Research Data Centre* (RDC), i.e. a "safe" and accessible place to researchers and scholars, who can autonomously carry out their statistical analysis on the microdata produced by the National Statistical Institute in compliance with confidentiality rules. Therefore the ADELE Laboratory meets the need of statistical analysis for research use, when microdata are required, though disclosure is not possible due to statistical confidentiality. If compared to other access facilities, in the RDC case security of data and statistical confidentiality are ensured by controlling both the environment and the results carried out by users.

The following guide illustrates the service provided by access facilities to statistical information intended for specialized users, it describes aspects of the access to services and provides clarifications on processing development and on rules for the output release.

For further information: adele@istat.it.

SUMMARY

CHAPTER 1 Services and products provided by Istat for specialised users

The National Statistical Institute is permanently committed to ensure availability of statistical surveys' results in the easiest and most complete way. This commitment is translated in a variety of access facilities to datasets and in the number of services and products offered to the users.

For users in need of specific processing or of surveys microdata for research use, Istat provides microdata files and tailored processing by ADELE Laboratory. Research Protocols, though not a service to the users, can be added as specific research tools and opportunities of access to datasets by external entities.

1.1 Tailored processing

The tailored processing (http://www.istat.it/en/information/users) enables to meet the needs of knowledge which are not considered in the ordinary release plans (i.e. specific unpublished tables) and require microdata processing. In this case users should illustrate their needs to the staff in charge by means of the Cont@ct Centre (https://contact.istat.it/en) and economically support processing costs. The required work to meet tailored processing requests is to be assessed by the bureau in charge. Processing should not be too complex as a rule. The results are available according to costs and timetable to be communicated to the users.

1.2 Microdata files

Microdata files (http://www.istat.it/en/products/microdata-files) are collections of survey elementary data where anonymity of statistical units, though already cleared of direct identification information, is safeguarded by means of various statistical methodologies which make indirect identification of statistical units highly improbable.

Requirements and conditions for file disclosure vary according to entities applying for them and are subject to the subscription of specific agreements for use.

1.3 Research protocols

By means of research protocols, personal data can be communicated to researchers operating on behalf of universities, public institutions and research entities, in the framework of joint projects which include Istat (art. 7 paragraph 3, Annex A.3. of the legislative decree n. 196/2003). Research protocols are undertaken within projects involving Istat, the latter can therefore take advantage from specific knowledge available from outside and capitalize the value added deriving from cooperation with other actors of the scientific research sector.

1.4 ADELE Laboratory

By means of the ADELE Laboratory (http://www.istat.it/en/information/researchers/laboratory-for-elementary-data-a), the Institute makes available validated microdata of surveys owned by Istat. These data are previously cleansed from direct identification elements and from confidential and judicial information, to comply with the requirements provided by the Code of deontology and good practice on processing personal data for statistical purposes and for scientific research in the framework of the National Statistical System (Annex A4, Art. 9. Treatment of sensitive or judicial data; Annex A3 Art. 4. Criteria for the assessment of the identification risk, Art. 4.bis Processing of personal, sensitive and judicial data, within the National Statistical Program, of the legislative decree 196/2003). This is a powerful tool for scientific research to carry out statistical analysis that because of their nature are to be performed on microdata (e.g. statistical models), and therefore cannot be satisfied by means of the other above-mentioned services (e.g. that related to microdata files). Tables as well as ordinary descriptive statistics not published by Istat, when not too complex, are to be applied for through the service of Tailored processing addressing the Cont@ct Centre (https://contact.istat.it/en). Tables

resulting from the General Population Census (years 1971-2001) are an exception; for these census rounds access can be applied for at ADELE Laboratory for table output, complying with the disclosure rules as provided in section 5.2.

In order to access the service, the designated representative that should be either a professor/researcher of a University, of a research institute or entity, or a member of a scientific society, should submit a research project to be authorized by Istat President and should sign a contract of statistical confidentiality undertaking. Once the researcher is authorized, he/she autonomously carries out (see section 3.3) processing on the required microdata using one of the assigned access positions to the Laboratory[*]. When processing ends, the output is assessed with reference to confidentiality safeguard, and if positive, it is released to the researcher. Whatever the case, microdata are never released (see section 5). It is worth emphasizing that the responsibility for all conclusions drawn from the data lies entirely with the author, the above-mentioned conclusions are not part of official statistics and do not involve Istat commitment in anyway.

The ADELE Laboratory experience is not isolated within the European context: other national statistical institutes, besides Eurostat, provide similar services at their Research Data Centre (RDC). Terms of use for laboratories, access facilities and rules for statistical disclosure control are shared in the main guide lines of the different countries, and they are included in a process of harmonization at an international level.

---

[*] The service is provided at Istat headquarters in Rome via Cesare Balbo 16 as well as at all Istat territorial branches (except the Provincial statistical institute of the Autonomous Province of Bolzano).

CHAPTER 2 Preparing the application for access to the Laboratory

The user who intends to use the provided service   submits to the Laboratory and application addressed to the Istat President (Annex A 'Application form for access to the Laboratorio di Analisi di Dati ELEmentari').

For submitting the application form a computer-aided system is available on the Istat web-site at http://www.istat.it/adele/MenuAdele.html.

Some information on the application's preparation and submission follow.

2.1 How to fill-in the form

1)   Duly designated representative of the research project

The duly designated representative of the research project subscribes the application form and undertakes the responsibility before Istat as for the therein exposed obligations. Access to the ADELE Laboratory can be exclusively granted for research projects where the designated representative is either a professor/researcher of a University, of a research institute or entity, or a member of a scientific society (i.e. belongs to an organization under the "Code of deontology and good practice on processing personal data for statistical purposes and for scientific research"). This   excludes some types of individuals and entities (for example private citizens or societies performing market researches). It is worthy specifying that the ADELE Laboratory services address users not belonging to Sistan. Information in the form should refer to the position of the designated representative at University, research institute or entity. PhD students are not accepted as designated representative, but can be included and carry out processing when a professor is the designated representative. In case the organization of affiliation is not a University, and in order to carry out the assessment of the entity profile   (i.e. on the applicant eligibility), a copy of the entity statute as well as of other official documents should be submitted to enable assessment of aims and scientific activities.

Access to the Laboratory is exclusively allowed for scientific research purposes, the latter are to be clearly understandable from the submitted research project. (item 6).

2)   Possible designated person in charge of processing data at Istat.

The designated representative of the research project can designate one or more persons for practically carry out data processing at the ADELE Laboratory. If this is the case personal data of the designated person are to be provided, specifying at the same time the position occupied within the entity of affiliation; access is never permitted to graduating students.

3)   Subject of the application

The exact name of the survey to be analyzed should be clearly indicated as well as the year and reference time of the survey data (e.g. 'Labour Force Survey: all quarters from 2005 to 2007').

As a support to user's search of the survey they are interested in, the National Statistical Institute made the following resources available to users:

  a) directly on-line:
  - List of surveys (www.istat.it/adele/ListaRilevazioni);

  b) request for further information:
  - Contact Centre (https://contact.istat.it/en);
  - Centers for Statistical Information (www.istat.it/it/informazioni/per-gli-utenti/sportelli-sul-territorio).

The list of integrated databases accessible at the ADELE Laboratory is also available online (www.istat.it/it/archivio/microdati-integrati).

In case other microdata files of the Institute are available for the required surveys (File standard, File for Research or mIcro.STAT), their inappropriateness is to be specified in relation to the research purposes.

4) Operational details

The applicant is invited to specify the software to be used for processing, among those made available in the Laboratory (see section 3.7). In addition to that the applicant should specify the preferred access facility, because the ADELE laboratory service is also supplied by Istat territorial bureaus and not only by the headquarters in Rome.

5) Result dissemination

The planned outputs should be indicated (e.g.: scientific publication, Phd. thesis, etc.).

6) Description of the research project

The name of the research project and the pursued scientific purpose should be reported. A "short" description of the research project is enough, as far as it meets the standards set by the scientific community (i.e. a detailed description of the legitimate research purpose, the theoretical background and the methods of analysis, if possible including references).

The project description is aimed at focusing the scope of the user's and the data sets to be made available, and therefore the results produced at the Laboratory. Istat does not carry out any assessment on the quality of the research conducted by the applicant, but in accordance with the law provides data for a purpose which should be legitimate and declared. The use of data for a purpose other than the declared one is not permitted.

7) Type of processing, and output to be collected

As for the type of processing, the Institute should assess that there are no other and more suitable forms of dissemination/communication to get the same results. The description of processing and output to be obtained should be very much detailed, because, before disclosure, the research project output should be assessed according to the rules illustrated in section 5.2; in case the required outputs do not comply with the provided rules, the Laboratory would not give the output back, and researchers would lose their work. A detailed description of the expected results to be released, enables to carry out a better assessment in advance on the feasibility of the research and to deal with possible impediments before processing are carried out, and therefore reduces the possibility of having output which cannot be disseminated.

For each statistics and processing the following items should be specified: type, reference (sub)population and variables or size for the phenomena to be studied. As for variables there is no need of providing a specific and complete list, rather of circumscribing the features to be analyzed in relation to the surveyed phenomena.

Special care should be devoted to detail the possible descriptive statistics and/or tables that represent for the ADELE Laboratory an exceptional output, they should therefore be limited to those supported by statistical models. Tables deriving from the general censuses of the population for 1971- 2001 are an exception to this rule; access to the ADELE Laboratory can be applied for with reference to these years of census following the rules for disclosure provided in section 5.2.

2.2 Submitting the application

The applicant is invited to previouslysend by email to adele@istat.it the "Application form for the access to the laboratory of microdata analysis" (Annex A). The assisted compilation is available at: www.istat.it/adele/MenuAdele.html, for a preliminary assessment. Following this assessment the persons in charge of the Laboratory may require possible changes or integrations.

After that the applicant can send the application form approved by the ADELE Laboratory staff using one of the following means:
  - fax n. 06.46678004;
  - certified email to adele@postacert.istat.it;
  - email to adele@istat.it. In this case the original copy of the request should be kept and delivered to the Laboratory staff at the first appointment.

The form, when sent , should be accompanied by a copy of the identity document of the person in charge of the research project.

In this way the formal authorization process starts and the user will be informed via email by the staff on the results.

CHAPTER 3 Service access and use

3.1 Service access

In case the authorization is released, a workstation associated to the research project is created. In addition to the person in charge, one or more persons in charge of processing data can be associated. Credentials will be prepared for the first designated person indicated in the application form, and in in case other persons indicated in the application form (or the same person in charge of the project) need to autonomously access (i.e. while the first designated person is absent), an email request should be sent in order to receive personal credentials.

Credentials for the access point are emailed by the Laboratory staff to the recipient alone and they are strictly personal. Therefore the user should not communicate credentials to any third party, including the Istat staff.

The service is provided at Istat headquarters in Rome - via Cesare Balbo 16, and at Istat territorial bureaus (http://www.istat.it/it/istituto-nazionale-di-statistica/organizzazione/sedi-territoriali) except of the Provincial statistical institute of the autonomous province of Bolzano.

It should be considered that whatever the chosen access facility, an appointment should be agreed with the chosen access facility at least three working days in advance. Contacts for the chosen access facilities will be communicated to the user after credentials were sent.

The Laboratory staff (adele@istat.it), should be, anyway kept informed on any communication between the user and the chosen access facility.

3.2 Prerequisites for processing data

The ADELE Laboratory research is addressed to autonomous researchers for the following operations:

- specifying statistical surveys and their necessary variables for the research project aims;

- use of hardware and software tools made available at the access point;

- interpretation of data and of the carried out processing.

At www.istat.it/adele/ListaRilevazioni the list of all surveys and dataset accessible through the ADELE Laboratory, including their metadata, by means of which the user can access the list of all present variables and related modes is available. It has to be noticed that for each of the indicated surveys, the variables shown are the only ones available at the Laboratory. The latter, in compliance with the law requirements, does not make available to the users variables that enable the identification of statistical units, confidential or judicial information.

3.3 Metadata and microdata files

The microdata file available at the ADELE Laboratory are coded by means of ASCII and are provided in a separate format by the tabulator, with the variable names reported in the first line. If users prefer, they can request at adele@istat.it, files to be made available in a fixed-size fields format.

Data related to the various research projects are downloaded by the Laboratory staff directly within an area dedicated to the research project. Even in the case that various project are carried out by the same members, the file copies cannot be shifted from an area to another.

The metadata are supplied with the microdata, the former report information on variables, on the corresponding modes and possible further information on the data themselves. Metadata are provided

in the HTML format and can be downloaded also from the page for assisted compilation of the access request. The page is available at http://www.istat.it/adele/MenuAdele.html.

The tool for assisted compilation also enables to download examples of the solicited microdata files. If necessary the example files enable to settle the code to be sent to the Laboratory according to the modes specified in section 3.6.

3.4 Working environment

The user by means of the provided work environment accesses the server of the Institute where data to be processed are kept.

Access to the personal working environment is allowed through any of the access points to the Laboratory and researchers of the same project (i.e. included in the same and authorized access application) can also simultaneously access. Each user will be required to change the password at the first access. The password should be at least 8 digits and include at least three of the following groups of digits: capital letters, lowercase letters, figures, special characters. The user will be requested to change password also in the case of no accesses during an over three-month period.

The working environment is based on Microsoft Windows platform with some peculiarities such as limitations to alteration or execution of commands that are potentially dangerous for system security and isolation from the Internet. No any input/output action is allowed (e.g. use of printers or USB entries) exception made for keyboard, mouse or screen. Some statistical software are made available for each user Id (SAS, STATA, SPSS and R) in addition to the Microsoft Office package and the RStudio interface.

The working space is an area shared by all the researchers of the project and it is in the disk **H**. In the disk you find a folder named **Work**, all the files of the project **SHOULD** be saved in that. If files are saved in other positions they will be missed when ending the working session. Within the directory **H:/Work** the users can freely create any new subfolder they consider useful to develop the project.

Two subfolders are contained in the folder **H:/Work,** they have specific functions and they should neither be deleted nor renamed:

- **Work/Dati**: it contains microdata files and the related metadata required for the research project and uploaded by the Laboratory staff;

- **Work/Output**: it is set to contain the output of the research project the user is going to submit for statistical disclosure control. The user will find in it the form for the output description (Annex C), it should be filled in to apply for output assessment. In this folder no other files are to be saved except those containing the output to be submitted and the filled-in form. For example in progress and partial processing results should be excluded, while files related to the code used for final processing can be included in compliance with information provided in section 3.6.

The user is required to delete all files which are no more necessary for processing data.

Every day the user SHALL close the working session choosing "ESCI" from the "START" menu (and NOT "DISCONNETTI").

3.5 Installation of additional packages

In case the user needs to use either STATA or R packages which are not included in the basic installation, a specific request should be submitted to adele@istat.it with at least seven working days in advance. Required time may vary depending on the number of requested packages.

The request should contain the explicit name of the packages to be installed and the software they refer to (STATA or R). The user will be informed when installation is completed.

However it should be noticed that whatever package installation is limited by its compatibility with the operating system of the Laboratory.

3.6 Use of external source codes

The user can organize personal source codes for the available statistical software and request the uploading of them in his/her personal working area, e.g. making use of example metadata and files available on the Istat web page.

Files shall necessarily be in text format and sent with at least three working days in advance to adele@istat.it. Uploading time may vary according to the number and volume of files.

When the used source codes are saved in the Output folder they can be requested together with the output results at the end of the research project. It is however required that these files are limited to those necessary to reproduce the final project output, and that no any type of data or information is contained in these files.

3.7 Uploading external files of data

In case of real need the uploading of external files of data in the personal working area can be requested. However some constraints on data themselves should be taken into account, data should be:

- Aggregated data (no any microdata can be requested for uploading);

- Open data;

- Not to be authorized or with any use limit.

In order to request uploading, the user should send a specific email to adele@istat.it therewith attaching:

- Data files (either in text format or Microsoft Excel only);

- The duly filled-in form for external data (Annex D).

The Laboratory staff, in case of need, may request further information to the user.

The request should be submitted with at least five working days of advance to the day of data use. However assessment time may vary according to the volume and complexity of the data. The user will be informed either on the positive outcome of the request or on the reason of the rejection.

CHAPTER 4 Data safeguard at the ADELE Laboratory

Safeguarding access to statistical information, as a community patrimony, is one of the Istat institutional duties. On the other hand the right to statistical information is confronted by the right to *privacy* of the subjects information is referred to. The delicate task of NSIs in this field is therefore safeguarding privacy of single persons while ensuring statistical information to the community.

Access to micro data for research purposes is therefore allowed in compliance with national rules on statistical confidentiality and personal data protection. This access should be performed according to conditions and guarantees provided by the art.7 of the "Codice di deontologia e di buona condotta per i trattamenti di dati personali per scopi statistici e di ricerca scientifica effettuati nell'ambito del Sistema statistico nazionale" (Allegato A.3 del d.lgs. n. 196 del 2003)[Code of ethics and good conduct for the processing of personal data for statistical and scientific research purposes carried out within the National Statistical System "(Annex A.3 of Legislative Decree No. 196 of 2003) TN].

Data confidentiality at ADELE Laboratory is safeguarded by means of three actions:
1. physical protection of the Laboratory;
2. personal responsibility of the accessing user;
3. output disclosure control (see chapter 5).

4.1 Physical protection of the Laboratory

Working stations for access to ADELE Laboratory are located within Istat offices. The access to working stations is monitored and permitted to authorized users only, and their identity is verified by the Laboratory staff.

Working stations are disconnected from the Institute Intranet and from the Internet. This means that the user can access neither the Internet nor his/her email account. In addition to that downloading or uploading data is physically prevented and formally forbidden: all forms of data entry and exit (exception made for the ordinary use of mouse, key board and screen) are prevented to the user and the laboratory staff only is enabled to, after the content has been checked (see chapter 3).

4.2 Personal responsibility in access

The authorized user when accessing the ADELE Laboratory undertakes (Annex B):
- to use data made available for research purposes only as declared in the application;
- not to attempt in any way or for any reason to identify statistical units referred to the made available data;
- not to attempt in any way or for any reason to copy, even partially, the data made available or the results of their data process.;
- to comply with the provisions on statistical confidentiality and personal data protection;
- to cooperate with the Laboratory experts and provide all the necessary information for a correct interpretation of the data process performed and the required output (see chapter 5).

Whoever is responsible for violating the commitments undertaken with Istat will be liable to report to the competent judicial authority, in addition to exposing themselves to the following penalties:
1) service stop;
2) denial to the right of access to the Laboratory in case of possible future requests;
3) reporting of facts to the institution of affiliation for adopting the subsequent disciplinary measures.

CHAPTER 5 Rules for the release of the processed data and output description

5.1 Conclusion of data processing, request for output release and description

When all data processing related to the research project are completed, the user sends the request for the output release by email to adele@istat.it.

The output to be released has to be assessed in relation to statistical confidentiality by the ADELE Laboratory staff. In order to do that, the output has to be saved in the folder "Work\Output" together with the "Form for output description" (Annex C) completely filled-in. In the form all the following information are to be provided: purpose and the modes of analysis, output file name and content, the processes performed on the original data set and the possible (sub)populations analyzed, the meaning of each variable (including the definition of the derived variables), and any other information considered useful to a correct interpretation of the output files. The output description should be sufficient to understand it (reference to other sources is not permitted, i.e. the used syntax files), and can be provided also after data processing, and by email: in this case the assessment, and the possible release, have to be delayed.

The "Output" folder should exclusively contain the file to be released, and their description.

Documents are to be correct and comprehensive, drafted as indicated in order to enable the Laboratory to organize the useful information to carry out the necessary assessment prior release: in case of insufficient documentation, the necessary documents should be provided, and further appointments can be agreed; anyhow ADELE Laboratory has the right to require further explanations on the produced output.

Not any release of midway output is permitted (only final output of the project can be released).

5.2 Rules for output release at the ADELE Laboratory

The ADELE Laboratory staff is in charge of assessing the output according to the following guidelines:

- The output should be clearly and completely documented (see section 5.1);
- The output should be organized in order to be released as it is, with no need of changes by the Laboratory staff during assessment; in case of no releasable output some further appointments should be agreed in order to make it releasable.;
- Descriptive tables and statistics should be limited to those that support statistical models. exception made for tables deriving from the General Population Census (years 1971-2001) for the latter access can be required to the ADELE Laboratory, complying with the following rules;
- In order to enable assessment by the Laboratory Staff, the output is preferred in a text file, either Word or Excel, but not in the format of the owner of the statistical applications used; in case of descriptive statistics or tables the excel format should be used;
- Producing results without weight normalization is strongly discouraged; however, for assessment purposes, the users are invited to communicate if standardized (normalized) weights are used and how (whether it is a normalization to the total population or to a specific sub-population);
- The output volume may be itself a reason for refusal to disclosure: the output to be released should be minimal and correspond to the output included in the work to be disseminated; as an indication, a maximum of 30 pages is suggested (~ 60Kb in ASCII text format);
- **Disclosure of intermediate output (**i.e. not concluding the work**) is not permitted**.

Some specific rules for the most frequent output disclosure are listed below:

1) **Descriptive statistics and tables to support statistical models**
   Each value reported in tables or descriptive statistics in general must refer to at least 10 statistical units.

In details:

- descriptive statistics that report precise data on the individual units (for example maximum and minimum for continuous variables) cannot be released;
- mode, minimum and maximum can be released if the modes that identify them are assumed by at least 10 units;
- quantiles: the median is considered releasable if referring to a distribution of at least 50 units; the other quantiles are not released except for special cases to be agreed;
- means, ratios and indicators: these outputs must be presented in their disaggregated form (for example for means and ratios separate numerator and denominator should be presentes; for means of dichotomous variables also the complement should be presented, etc.); each element must be accompanied by the number of units (at least 10) that contribute to determine its value; this also applies to any additions, also to be explicitly presented (for example, if an indicator shows the value of 95%, it should be possible to verify that even 5% corresponds to at least 10 units; dichotomous variables, etc.);
- intensity tables: users must specify the number of units (at least 10) that contribute to determine the value of each cell (match the relative frequency table);
- frequency tables: tables with a cell number of less than 10 unweighted units are not released in any case.

2) Charts on variables

Charts on non-continuous variables must be accompanied by the corresponding table of the values they represent; this will be evaluated as specified in the previous point. Charts on continuous variables must be saved as images and deprived of the abscissa values.

3) Regressions

The following output can be released:

a) (p-1) *estimated parameters*, where p is the number of regressors, when all the conditions specified below are verified:

- ✓ the total number of observations must exceed the number of explanatory variables of at least 100 units;
- ✓ among the explanatory variables the presence of at least one variable for which the operations of sum, difference, product and quotient make sense is necessary;
- ✓ the observations on all data must refer to at least 100 different units of analysis.

b) *Diagrams on the correct model specification*:
1. the histogram of residues, deprived of the abscissa values;
2. the diagram of the density of residues, deprived of the abscissa values;
3. the Q-Q plot of residues, deprived of the abscissa and ordinate values;
4. the P-P plot of residues;
5. the diagram of the ranks of the residues vs. the ranks of the predicted values of the variable to be explained;
6. the diagram of the ranks of the residues vs. the ranks of an explaining variable;

c) *Statistics on the adaptation and correct specification of the mode*:
1. The statistics expressed by a scalar;
2. The statistics expressed by a vector, having a size not greater than the number of parameters estimated, that is (p-1). Only the conventional level of significance is released from the obscured regressor (0.005, 0.01, 0.025, 0.05, 0.1).

However are excluded from the release:
1. residuals of regression;
2. "aforesaid" values of the variable to be explained.

4) Factorial analysis and models with structural equations

The following output can be released:
1. Model parameters,
2. the (possible) correlation matrix among factors,
3. standard errors and statistics on the significance of the model parameters,
4. communality and specificity for each variable,
5. factorial scores referring to units of analysis that are not individuals, families or enterprises,
6. statistics on the goodness of the model, expressed by a scalar,
7. the scree plots relating to the eigenvalues of the covariance / correlation matrices observed,
8. diagrams of relational models between manifest and latent variables.

5) Analysis by principal components

The output listed below can be released:
1. eigenvalues,
2. the following statistics:
   a) variance explained by the factor axes,
   b) matrix (p x k) of the **relative contributions (squares of the cosines) of the points-variable**,
   c) matrix (p x k) of the absolute contributions of the points-variable,
   d) matrix (p x k) of the coordinates of the points-variable,

where p is the number of variables and k is the number of eigenvalues which, organized in non-decreasing order, accumulate a fraction of the total variability not exceeding 85%,
3. scree plot of the eigenvalues,
4. diagrams of the projection of the variable points on the factorial plans.

6) Analysis of correspondences

The output listed below can be released:
1. eigenvalues,
2. the following statistics:
   a) inertia explained by the factorial axes,
   b) matrix (p x k) of the relative contributions (cosine squares) of the mode-points (column and/or row),
   c) matrix (p x k) of the absolute contributions of the mode-points (column and / or row),
   d) matrix (p x k) of the coordinates of the mode-points (column and / or row), where p does not exceed the total number of modes and k is the number of eigenvalues which, ordered in non-decreasing order, accumulate a fraction of the total inertia not exceeding 85%,
   e) test values, expressed by scalars, on the significance of each additional mode (in the analysis of multiple correspondences),
3. scree plot of the eigenvalues,
4. diagrams of the projection of the row and / or column point-modes on the factorial plans.

With regard to the units of analysis, for any type of processing, the values observed and the statistics that do not comply with the rules on "Descriptive statistics and tables" remain excluded from the release.

CHAPTER 6 Modification of the research project

During the execution of a research project at the Laboratory the initial conditions or aims of the project may change. In this case, the user is obliged to immediately communicate these changes, report them in the application form and re-submit (as described in chapter 2) an application form similar to the one previously authorized one, adding or modifying only the necessary information .

The new request should be submitted in the following cases:

• addition of microdata related to further surveys compared to the previously authorized application;

• addition of persons in charge of carrying out the analyzes at the Laboratory;

• change in the membership and / or qualification of one or more members of the research project and forfeiture or extension of the employment contract;

• change of the aims of the approved research project.

Following receipt of the new application form by the Staff, the user will receive subsequent communication of the outcome of the request in relation to access to the Laboratory.

CHAPTER 7 Project conclusion

7.1 Project conclusion

A research project is concluded when:

- more than six months have passed with no direct or carbon copy communication at the address adele@istat.it by the researchers participating in it.

- The produced output has been released to the user. In this case the working area will be maintained for a six-month period after the output was released. The user can request at adele@istat.it, before the six-month term, a time extension for maintenance specifying why (for example, the use of the same area for further and forthcoming research projects). It is worthy reminding that together with the final output also the code used for processing data can be required according to section 3.6.

- 24 months have passed from the beginning of the project. In case the user needs to continue the project beyond this term, they can send a request for a time extension to adele@istat.it, before the deadline and specify why.

For the last two cases the Laboratory Staff will acknowledge receipt to the user by e-mail of the request and the duration of the time extension granted.

7.2 Form for the Laboratory assessment

At the end of each project, users are asked to optionally answer a short questionnaire (Annex E). The purpose is to assess the service provided from the user perspective. The form does not include personal data. The information collected are exclusively used to produce reports on the service quality and are not in any case associated to users' personal data for dissemination.

7.3 Sending of scientific works containing the released output

The user who carried out processing at the ADELE Laboratory should send via email to adele@istat.it an electronic copy of the work where the output released by the Laboratory appears, or any other form of dissemination of the results achieved by means of the processing carried out at the ADELE Laboratory (i.e. article on a scientific journal, book or chapter of a book, graduation or Phd, research report, presentation in a conference, etc.).

In addition to that the researchers should quote the source (e.g. Istat, the Survey the data contained in the files refer to) in texts, in publications and in any other form of dissemination of studies, analysis and processing of any type for which they used all or part of the results of the processing conducted at the ADELE Laboratory. They should specify that the processing was conducted at the Istat Laboratory for the microdata analysis, ensuring in any case compliance with the legislation on the protection of statistical confidentiality and of personal data, and also indicating that the results of the analysis carried out at the Laboratory are the sole responsibility of the author, do not constitute official statistics and do not in any way implicate Istat. Eventually, in case analysis were conducted without using weight normalization, this fact must be clearly specified in the disclosure of the results.

The following text is supplied as a model:

The data used in this work were supplied by Istat and referred to the Survey ….

Processing were carried out at the Istat Laboratory for the Microdata analysis and in compliance with the norms on the protection of statistical confidentiality and of personal data. The results and opinions expressed here are the sole responsibility of the author, they are not official statistics and do not implicate Istat in anyway.
It is worthy specifying that the analysis were conducted without using weight normalization.

Al Presidente dell'Istat
c/o Segreteria del Laboratorio ADELE
Via Cesare Balbo, 16
00184 Roma

APPLICATION FORM FOR THE ACCESS TO THE LABORATORY OF MICRODATA ANALYSIS

Sec. 1) PERSON IN CHARGE OF THE RESEARCH PROJECT[2]

Full name …….

email…………………………………………………………………………Tel. …………………………………………………………………

Organization[3] …….

Position and function……..

Organization address: ……………………………………………………………………………………………

Municipality/city ……………………………………………………………………………………………...

Country …………………………………………. ……………………………………….

Sec. 2) DESIGNATED PERSON IN CHARGE OF CARRYING OUT PROCESSING AT ISTAT[4]

Full name …….

email …………………………………………………………………………Tel. ……………………………………………………………..

Organization ………………………………………………………………….…………………………………………………

Position and function ……..

Sec. 3) SUBJECT OF THE REQUEST

Microdata of the survery/s (please list reference periods of each survey, for instance "Inserimento professionale dei laureati years 2007 and 2011")

………

………

………

Sec. 3a) INADEQUACY OF STANDARD FILES, FILES FOR RESEARCH OR MICRO.STAT

In case other Istat microdata files are available for the submitted surveys (File standard, research files or mIcro.STAT), explain the reasons why they are not suitable for your research purposes:

………..

………..

---

[2] The person in charge of the research project, who has the responsibility of compliance with the contract rules. Entrance to the laboratory is permitted to: University professors (full, associate and contract professors), researchers of Universities, research bodies or institutes, and members of scientific societies out of SISTAN to which the Code of deontology and good conduct for the processing of personal data for statistical and scientific research purposes applies (annex A.4 of d.lgs. n. 196/2003).

[3] In case the applicant is not a University its statistical or scientific research purposes should be specified.

[4] This part is to be filled-in only in the case access is required for another person different from the "Person in charge" of the research project. It is worthy specifying that also the designated person has the responsibility of complying with the contractual rules. In case of more "persons" replicate the section.

Sec. 4) OPERATIONAL DETAILS

Software to be used (tick the box):

SAS ☐ STATA ☐ SPSS ☐ R ☐

Location of the first access[5]………

Sec. 5) DATA DISSEMINATION

Purposes and scope of dissemination of the results of the research project:

☐ Phd Thesis

☐ Scientific work (for a publication on a journal or presentation at a conference)

☐ Writing reports for public or private bodies (specify):
…………………………………………………………

☐ other (specify): ………………………………………………………………………………………

Sec. 6) DESCRIPTION OF THE RESEARCH PROJECT

Title of the project: …….

Abstract (describe the main goals and expected results of the project, max 20 lines)…………………………………….

……

……

Sec. 7) TYPE OF PROCESSING AND OUTPUT TO BE COLLECTED[6]

For each of the following points describe the type of processing and indicate (one for each line) the output to be collected:

☐ Regressions:

………..

………..

………..

☐ Factor analysis and structural equation models:

………..

………

………

---

[5] The service is provided at Istat headquarters in Rome, via Cesare Balbo 16, as well as at Istat territorial branches (exception made for the Statistical bureau of the Autonomous Province of Bolzano).

[6] It is worthy specifying that the description provided here is used for the application assessment and is not in anyway a commitment by Istat to release the output, in compliance with point 7 of the "General conditions for the use of ADELE Laboratory".

☐   Analysis of the main components:

……

……

……

☐   Correspondence analysis:

……

……

……

☐   Descriptive statistics supporting models (means, tables, indicators, etc.) and graphs on variables:

……

……….

The undersigned (surname, name) _____

In his capacity of person in charge of the above-described research project,

certifies that:

- Information contained in this form are complete and true and that any variation will be immediately communicated to Istat;
- the research activity carried out falls within the scope of the *Code of Deontology and Good Conduct for the processing of personal data for statistical and scientific purposes* (Annex A.4 of Legislative Decree 30 June 2003, No. 196 - *Code for the protection of personal data*);
- with regard to any access to data that can be or is referred to enterprises, neither the undersigned nor any person in charge of carrying out the processing at Istat are in one of the situations of incompatibility or of current or potential conflict of interest envisaged by current legislation on the protection of competition, as well as on the protection of industrial and business confidentiality;
- also declares to have read, understood and approved the "General Conditions of Use of the ADELE Laboratory" reported below and to undertake and adopt all the measures and behaviors necessary for their correct and complete compliance.

The undersigned is aware that, pursuant to art. 76 of the d.p.r. of 28 December 2000, n. 445, false declarations, false documents and the use of false deeds are punished according to the penal code and special laws.

Place_____Date _____   SIGNATURE[7]

                                           (*person in charge of the research project*)

Information pursuant to art. 13 of the legislative decree 30 June 2003, n. 196

Personal data collected by means of this form will only be used to perform the necessary operations for the provision of the requested service. The same data will also be kept for monitoring the management system.

Istat is the data controller of the personal data; the Director of the Central Directorate for the development of information and statistical literacy (dcsi@istat.it) is the responsible for the treatment, it is also possible to apply to the DCSI Director for the exercise of the rights referred to in art. 5-ter of the legislative decree 25 May 2016, n. 97.

---

[7] A photocopy of an identity document of the person in charge of the research project should be enclosed with the form (art. 38, D.P.R. n. 445 of 28/12/2000).

GENERAL CONDITIONS OF USE FOR THE ADELE LABORATORY

The following are the "General conditions of use for the ADELE Laboratory"; further details about the provision of the service are included in the "User Guide" available online on the Institute's website, at:

https://www.istat.it/en/information-and-services/researchers/laboratory-for-elementary-data-analysis

1. For exclusive statistical and scientific research purposes, Istat allows specialized external users, based on a specific request and prior presentation of a research project, to analyze the basic data of the surveys available at the ADELE Laboratory. The data communicated are the result of findings are owned by the Institute, they are free from direct identifying references and sensitive and judicial data, in compliance with the Code regarding the protection of personal data.

2. Access to the Laboratory is subject to verification of the existence of the conditions and requirements established by current legislation (Article 7, paragraph 2, of the Code of Ethics for the processing of personal data carried out within the National Statistical System) and to the assessment of the submitted research project. In order to allow this assessment, the description of the project must highlight both the characteristics of the analysis that make the use of microdata necessary and limit the scope of use of the results.

3. The Laboratory is not authorized under any circumstances to release microdata. The objective of the Laboratory is to enable statistical analyzes to be carried out when, due to their nature, they require access to microdata (for example, statistical models). Tables and common descriptive statistics, when not published by Istat and if not excessively complex, should be requested with the personalized processing service by contacting the Cont@ct Center (https://contact.istat.it//index.php) The tables deriving from the general population census (years 1971-2001) are an exception and access can be requested at the ADELE Laboratory, according to the rules for release provided in the User Guide (section 5.2).

4. The Laboratory makes available to authorized users, in a special room with access controlled by a specialized staff, a workstation for processing. Each user finds in the dedicated station the required microdata (ASCII files in tab-delimited format) complete with the relevant metadata. In addition, a rescue area is made available to users, suitable for maintaining partial processing for up to 6 months from the last use or email contact.

5. The user autonomously reads data, uses software and hardware, elaborates and analyses the results.

6. Photograph, copy or reproduce in any way the data made available is not permitted inside the Laboratory; using electronic equipment, including mobile phones, to connect to the Internet or communicate outside the Laboratory, is forbidden as well.

7. It is absolutely forbidden to introduce into the user room data that could allow the identification of the statistical units to which the made available microdata refer to, and to make any attempt to identify them. No data or material can leave the user room out of those permitted pursuant to the following point 8.

8. Any kind of input and output must be previously viewed by the Laboratory staff. If the staff find neither violations of the statistical nor of the interested parties' confidentiality, they authorize entry or exit of input and output respectively. As a consequence, the release of intermediate output of the research is not expected and the results of the final elaborations will be delivered to the users a few days after they are finalized, in order to allow to perform the necessary checks. The user also undertakes to collaborate with the Laboratory Staff, documenting the elaborations performed and the output required according to the indications of the Staff and providing anyway all the information necessary for a correct interpretation of the output.

9. Whoever violates the commitments undertaken in relation to ISTAT will be liable to report to the competent judicial authority, in addition to exposing themselves to the following penalties:

    a) service interruption;

    b) denial of the right to access the Laboratory for any possible future requests;

    c) reporting of facts to the institution of affiliation for the adoption of the consequent disciplinary actions.

10. The ADELE Laboratory can be accessed by appointment, to be agreed by email with the chosen access point at least three days in advance, and after signing a responsibility undertaking. The email addresses of the regional access points to the Laboratory and their timetables will be communicated to the user by email. For more information, write to adele@istat.it.

11. The research project is concluded when the output of the elaborations is released to the user. Furthermore, a project is considered as concluded after six months without communications from the user.

<div style="text-align:center">

RESPONSIBILITY UNDERTAKING
BY THE PERSON AUTHORIZED TO ACCESS THE ADELE LABORATORY

</div>

The subscribed , in his/her capacity of

☐   Person in charge of the research project described in the "Application form for the access to the Laboratory of Microdata Analysis" undersigned on the ………………….

☐   Duly designated person (by the person in charge)……………………………………………….., to carry out data processing for the research project described in the "Application form for the access to the Laboratory of Microdata Analysis" undersigned on the ……………….....
by……….

<div style="text-align:center">

UNDERTAKES TO

</div>

- -use the data made available and the results of the elaborations carried out for the sole research purposes declared in the application for access to the ADELE Laboratory;
- do not attempt in any way and for any reason to identify the statistical units to which the made available data refer to, and do not allow others to carry out similar attempts at identification;
- do not attempt in any way and for any reason to partially copy the made available data or the results of their processing;
- comply with the rules on statistical confidentiality and personal data protection, as well as, with regard to data that refer or may be referred to enterprises, the rules on industrial and company secrecy;
- operate in compliance with the *Code of Deontology and Good Conduct for the processing of personal data for statistical and scientific purposes* (Annex A.4 of Legislative Decree 30 June 2003, No. 196 - *Code for the protection of personal data*);
- quote the source (e.g. Istat, the Survey the data contained in the files refer to) in texts, in publications and in any other form of dissemination of studies, analysis and processing of any type for which they used all or part of the results of the processing conducted at the ADELE Laboratory. They should specify that the processing was conducted at the Istat Laboratory for the microdata analysis, ensuring in any case compliance with the legislation on the protection of statistical confidentiality and of personal data, and also indicating that the results of the analysis carried out at the Laboratory are the sole responsibility of the author, do not constitute official statistics and do not in any way implicate Istat. Eventually, in case analysis were conducted without using weight normalization, this fact must be clearly specified in the disclosure of the results;
- send to Istat (adele@istat.it) all information concerning activities and initiatives of scientific dissemination involving reference to or use of the results of processing performed at the ADELE Laboratory.

<div style="text-align:center">

DECLARES

</div>

- in the case of access to data referring to or referable to companies, not to be in any of the situations of incompatibility or of current or potential conflict of interest provided by the current legislation on the protection of competition, as well as the protection of industrial and corporate secrecy ;
- to have read, understood and approved the "General conditions of use of the ADELE Laboratory" mentioned below.

Surname……………………………………Name…………………………………………………..

Date…………………………….Signature……………………………………………………………………

Identity document…………………………………………………………………………………….

Information pursuant to art. 13 of the legislative decree 30 June 2003, n. 196

Personal data collected by means of this form will only be used to perform the necessary operations for the provision of the requested service. The same data will also be kept for monitoring the management system.

Istat is the data controller of the personal data; the Director of the Central Directorate for the development of information and statistical literacy (dcsi@istat.it) is the responsible for the treatment, to the DCSI Director is also possible to apply for the exercise of the rights referred to in art. 5-ter of the legislative decree 25 May 2016, n. 97.

GENERAL CONDITIONS OF USE FOR THE ADELE LABORATORY

The following are the "General conditions of use for the ADELE Laboratory"; further details about the provision of the service are included in the "User Guide" available online on the Institute's website, at:

https://www.istat.it/en/information-and-services/researchers/laboratory-for-elementary-data-analysis

1. For exclusive statistical and scientific research purposes, Istat allows specialized external users, based on a specific request and prior presentation of a research project, to analyze the basic data of the surveys available at the ADELE Laboratory. The data communicated are the result of findings are owned by the Institute, they are free from direct identifying references and sensitive and judicial data, in compliance with the Code regarding the protection of personal data.

2. Access to the Laboratory is subject to verification of the existence of the conditions and requirements established by current legislation (Article 7, paragraph 2, of the Code of Ethics for the processing of personal data carried out within the National Statistical System) and to the assessment of the submitted research project. In order to allow this assessment, the description of the project must highlight both the characteristics of the analysis that make the use of microdata necessary and limit the scope of use of the results.

3. The Laboratory is not authorized under any circumstances to release microdata. The objective of the Laboratory is to enable statistical analyzes to be carried out when, due to their nature, they require access to microdata (for example, statistical models). Tables and common descriptive statistics, when not published by Istat and if not excessively complex, should be requested with the personalized processing service by contacting the Cont@ct Center (https://contact.istat.it//index.php) The tables deriving from the general population census (years 1971-2001) are an exception and access can be requested at the ADELE Laboratory, according to the rules for release provided in the User Guide (section 5.2).

4. The Laboratory makes available to authorized users, in a special room with access controlled by a specialized staff, a workstation for processing. Each user finds in the dedicated station the required microdata (ASCII files in tab-delimited format) complete with the relevant metadata. In addition, a rescue area is made available to users, suitable for maintaining partial processing for up to 6 months from the last use or email contact.

5. The user autonomously reads data, uses software and hardware, elaborates and analyses the results.

6. Photograph, copy or reproduce in any way the data made available is not permitted inside the Laboratory; using electronic equipment, including mobile phones, to connect to the Internet or communicate outside the Laboratory, is forbidden as well.

7. It is absolutely forbidden to introduce into the user room data that could allow the identification of the statistical units to which the made available microdata refer to, and to make any attempt to identify them. No data or material can leave the user room out of those permitted pursuant to the following point 8.

8. Any kind of input and output must be previously viewed by the Laboratory staff. If the staff find neither violations of the statistical nor of the interested parties' confidentiality, they authorize entry or exit of input and output respectively. As a consequence, the release of intermediate output of the research is not expected and the results of the final elaborations will be delivered to the users a few days after they are finalized, in order to allow to perform the necessary checks. The user also undertakes to collaborate with the Laboratory Staff, documenting the elaborations performed and the output required according to the indications of the Staff and providing anyway all the information necessary for a correct interpretation of the output.

9. Whoever violates the commitments undertaken in relation to ISTAT will be liable to report to the competent judicial authority, in addition to exposing themselves to the following penalties:

    a) service interruption;

    b) denial of the right to access the Laboratory for any possible future requests;

    c) reporting of facts to the institution of affiliation for the adoption of the consequent disciplinary actions.

10. The ADELE Laboratory can be accessed by appointment, to be agreed by email with the chosen access point at least three days in advance, and after signing a responsibility undertaking. The email addresses of the regional access points to the Laboratory and their timetables will be communicated to the user by email. For more information, write to adele@istat.it.

11. The research project is concluded when the output of the elaborations is released to the user. Furthermore, a project is considered as concluded after six months without communications from the user.

OUTPUT RECORD

DATA USED

Specify, among the data provided, those actually used in the processing for which release is requested: indicate the name and the reference period of the survey(s) used and specify any external data files used in the processing.

VARIABLES/INDICATORS

Report the name and a brief description of the variables used. In the case of variables not present in the original data bases (reclassifications made by the user, external variables etc.) in addition to the name and description, report the meaning of the methods assumed (or the construction method, especially in the case a variable takes values as a function of other variables).

TRANSFORMATIONS OPERATED ON VARIABLES

For each variable supplied by the laboratory and subjected to transformation, indicate the function used to obtain its transformation. For each variable created by the user, indicate in detail the construction procedure.

OUTPUT FILES

Report the name and structure (example: Excel file with a sheet per each considered year) of the output files for which release is requested, providing a summary description of the content.

PROCESSING CARRIED OUT

Describe each elaboration carried out, providing a brief but complete description.

It is useful to associate a denomination to each processing and report it in the output file, so as to ensure an unequivocal identification and interpretation.

FILTERS ON UNITS

For each processing (or processing group) specify the filters applied to the starting population and the number of observations involved.

Note that it is necessary to specify exactly the effective number in each processing, even in the case of reductions in the number due to the presence of missing values in one or more of the used variables.

WEIGHT SYSTEM

Specify the weight system you used and if this varied in the different processes.

If standardized (normalized) weights were used, specify whether the normalization was with respect to the total population or specific subpopulations.

Note that if you request the release of weighted output, the same must also be presented in the unweighted version to allow evaluation.

NOTES

Report any other information you may consider useful for a correct interpretation of the output files.

The applicant: _____   Date: ___/___/_____

N.B.: the output description should be sufficient to understand it; reference to other sources (i.e. the used syntax files)is not allowed.

RECORD ON EXTERNAL DATA FILES

FILE DESCRIPTION

Indicate the name, the type and possibly the structure (example: excel file with a sheet per each year) of the data file for which loading is requested. Describe the contents in a synthetic way.

DATA SOURCES

Report in this section the sources of the data contained in the file (institution, survey, year etc.).

VARIABLES' DESCRIPTION

Provide a list of the variables of the file. Provide a brief description of the variable next to the name, specifying whether it is a reclassification of variables found in the data provided by the Laboratory. In the case of indicators, report the meaning of the methods assumed or the procedure followed for the construction.

RELATIONS AMONG VARIABLES

Specify, in this section, any relationships between the variables in the file.

USE NEEDS

Motivate the need to have the mentioned file and specify in which analysis / processing you intend to use it.

NOTES

Write any other information you may consider useful for a correct interpretation of the above-mentioned files.

Applicant: _____ Date: ___/___/_____

QUESTIONNAIRE ON ADELE LABORATORY ASSESSMENT

Tick the box or the letter
Key: I = Insufficient; S = Sufficient; G = Good; E = Excellent

ACCESS PROCEDURE

1) How did you get to know the Laboratory?

    ☐ Istat/Sistan events
    ☐ Istat website
    ☐ Other users of the Laboratory
    ☐ Other …………………………………….

2) Clarity of the objectives of the Laboratory service   I   S   G   E

3) Clarity on:

    a) Application form for access   I   S   G   E
    b) User guide   I   S   G   E
    c) Tool for assisted compilation   I   S   G   E
    d) Web-page on the available surveys   I   S   G   E
    e) Information by the Laboratory staff   I   S   G   E

4) Remarks and suggestions

_____
_____
_____
_____
_____

SERVICE USE

1) Service accessibility at the chosen access point   I   S   G   E

2) Information clarity at first access   I   S   G   E

3) Clarity of metadata files   I   S   G   E

4) Assessment of the working area:

    a) comfort of the work station   I   S   G   E
    b) lighting   I   S   G   E
    c) heating/conditioning   I   S   G   E
    d) availability of the Laboratory staff   I   S   G   E

5) Remarks and suggestions

_____
_____
_____
_____

Istat

IT COMPONENT IN THE LABORATORY

1) Assessment of the Laboratory IT component

       a) Operating environment   I   S   G   E

       b) Processing capability   I   S   G   E

       c) Screen size   I   S   G   E

2) Software compliance with your needs   I   S   G   E

Mention any other software you may consider useful

_____

_____

_____

1) Remarks and suggestions

_____

_____

_____

_____

_____

GENERAL QUESTIONS

2) Availability of the Laboratory staff   I   S   G   E

3) Service compliance with your analysis needs   I   S   G   E

4) Overall assessment on the ADELE Laboratory   I   S   G   E

5) Remarks and suggestions

_____

_____

_____

_____

_____

Date: ____/____/_____