Data warehouse .STAT
User Guide
# Table of Contents

## INTRODUCTION

Main Concepts
- Data Organisation
- Search
- The Web Browser

## USER GUIDE

Getting Started Page
Language Selection
Locating Data through the Search
Locating Data by Browsing Themes
Locating a dataset or a query in the theme list

### Using the Table Features
- Selecting Dimensions and their Variables
- Dimension Member Colour Coding
- Changing the Time Dimension
  - Date Range Selection
  - Time Period Selection
- Rotating Dimensions by Adjusting the Dataset View
- Table Options
  - Showing Codes Instead of Names
  - Hiding Empty Rows or Columns
- Exporting to Excel
- Exporting to a CSV file
- Downloading related files and large data selections
- Data charts

Viewing Metadata
- Flags

### Working with Queries
- Saving a Query
- Merge Queries
  - Combining Queries
  - Viewing the Query
  - Modify a selection
  - Remove a selection
  - Saving a Multi-Dataset Query

Viewing Frequently Requested and Saved Queries

## CONTACT US
Introduction

ISTAT has implemented I.Stat (the corporate statistical data warehouse) starting from a set of statistical software components and services that OECD has developed for the delivery of its statistical data warehouse applications (OECD Stat Product).

The activities have been conducted under a Memorandum of Understanding ("MOU") and everything will be carried out in accordance with the respective rules and practices of both Organizations. ISTAT will use the OECD Stat Product for the sole purpose of its statistical data warehouse applications.

A series of sub-products on specific themes are connected to I.Stat and used the same technology. In the following guide we report generally "web platform", identifying any product related to the technology used.

The web platform provides a single online system where users can discover and access statistical databases from the Italian National Institute of Statistics. You will be able to build tables and extract data from across databases as well as work within individual databases. Use Browse Themes or Search to find the statistical information you need.

Each database includes detailed metadata to help you understand the numbers.

The platform offers a number of opportunities for improving data management, in particular in the following areas:

- General data identification, taxonomy and classification
- Data accessibility and security
- Data relevance by the provision of related metadata
• Fostering coherency of statistical data and metadata
• Eliminating duplication of effort by maintaining a single copy of key series.

Main Concepts

The platform features the following main concepts:
• Catalogues of information on themes, datasets, dimensions, dimension members and reference series
• Storage of metadata at all levels from dataset through to cell level
• Storage of numeric data figures, as well as cell-level flags
• A data warehouse searching mechanism

Data Organisation

Datasets

Within the platform, statistical data are organised in the form of datasets. A dataset is a collection of numerical values and their associated textual information, with all values sharing a common set of dimensions. Each dataset is attached to a specific statistical activity or subactivity.

Dimensions

The dimensions of a dataset are the axes on which the data are described. Territory and Year are two common examples of dimensions. Dimensions can be presented as either a flat list or as a hierarchy.

Dimension Members

Every dimension contains a pre-defined list of items, called dimension members. In the Territory dimension for example, the dimension members are the list of Regions, Provinces etc.

Metadata

Qualitative data, attached to the quantitative data, is called metadata. Metadata can be found at various levels, ranging from an abstract found at the level of an entire dataset, territory- or series-level footnotes which apply to a broad selection of data within a dataset, and down to the level of a single cell.

Flags

When a qualitative note is recurrent in a dataset, and the exact same note can be attached to many individual data figures, this may be stored as a flag. A flag is a letter, which appears alongside the numerical data in any table cell where it applies. A legend appears beneath the
table, explaining the signification of each flag present in the table. Flags are frequently used to note general data properties such as "confidential", "provisional" and "estimated".

**Search**

A text-based search facility, which searches through dataset names, dimension names, dimension member names and metadata, can be used to quickly locate relevant datasets and view tables.

**The Web Browser**

The **Web Browser** provides the online user environment to access the Italian National Institute of Statistics data. The browser has been designed to allow both experienced and new users to rapidly locate and retrieve statistical data and related metadata.

Main features include:
- Links to frequently requested tables - access the latest versions of the most frequently requested ISTAT data.
- "Table View" – view data, flags and metadata from a table that includes a number of features for customising the layout of the data.
- "Data Basket" - save data queries for later viewing, and share them with others.
- Merged Queries – combine queries to allow cross-dataset comparisons along common dimension(s).
- Metadata access - access all metadata from dataset level down to series footnotes and data flags.
- Full Text Search - rapidly locate series across data sets, data dimensions and metadata.
- Dynamic graphics – visualize data dynamically in a line chart, bar chart or other.
User Guide

The Web Browser has been designed to be straightforward and intuitive in use. It does, however, contain a number of advanced features, which reflect the diversity of the ISTAT statistical resources. This User Guide has been prepared to help acquaint first-time users with these features, by stepping through them with the aid of selected screen shots.

**Getting Started Page**

From the *Getting Started* page, the user has the option of finding ISTAT statistical data and metadata through a number of different avenues.

Data can be located in various ways:

- **Search**
- **Browse Themes**
- **Browse Queries**

*Each of these options will be explained in more detail in the remainder of the User Guide.*

The browser is best used in full-screen view at 1024x768 pixel screen resolution or higher.

![Figure 1: Browser Getting Started page](image)

Click here to open/close Data by theme or Popular queries.

Click on arrow/dash to expand or collapse the respective theme.

Click here to expand or collapse the respective panel.
Language Selection

The interface and data can be viewed in Italian or English. The language can be changed by clicking on the language option, located at the top of the screen.

Figure 2: Changing the language of the browser
**Locating Data through the Search**

Data can be located by entering keywords in the Search box, at the top right of the page. This will return a list of links giving the name of each dataset in which the keywords were found, and providing further details as to where the keywords are present within each dataset. The most relevant links are shown first based on the frequency with which the keywords appear in each dataset, and at which level they were found.

A keyword found in the name of a dataset is considered more relevant than a keyword found only in one of the dimension members of a dimension of the dataset or in the metadata.

Clicking on the link will open the dataset with the default view.

The tables resulting from a search will contain any references to the search keywords entered that could be found in the data.

These tables can be modified and used as a starting point to find data related to the keywords entered.

---

Figure 3: Results from the search ranked in order of relevance
Locating Data by Browsing Themes

In the BROWSE THEMES panel, click a theme title, click a sub-theme title and so forth, until you can select a dataset. Datasets can be recognised by the table icon in front of their name. A pre-defined, default view of the table will open. This view can be customized (see next section).

Figure 4: Browsing through themes to view data
Locating a dataset or a query in the theme list

At the top of the BROWSE THEMES panel click the text box and enter a key term to search for within the theme and data set list. Click the “>>” sign or press ENTER to execute the search. All found entries will be highlighted in yellow.

To start over, simply click RESET to the right of the text box and click the text box again in order to enter a new key term.

Figure 5: Find a theme or data set

This search routine will look for exact matches of the entered text string, including spaces but irrelevant of miniscule or majuscule characters.

This method also allows a quick lookup for well known Istat data sets.
Using the Table Features

When a dataset is first opened, a table in a default view is presented. This default data selection can be modified using the features found atop the table.

Clicking on [Customise] is shown:

the key Selection:

Figure 6: Default dataset view and Current Selection values

Figure 7: The Selection key
that shows information on the dimensions of the dataset with an indication, in parentheses, of the number of items, or members, selected in each dimension;

the key **Layout:**

Figure 8: The Layout key

where you can filter and customize the size of the table view;

the key **Table Option:**

Figure 9: The Table Option key
where you can take a series of steps to better display the information of the selected dataset.

**Selecting Dimensions and their Variables**

The items selected for each dimension, often referred to as dimension members, can be modified by clicking on the button to *Select Dimensions* at the top of a table.

![Image: Selecting Dimensions and their Variables](image.png)

**Figure 10: Using the Dimension Selector**

To select or deselect an item, click the check box to the left of the item. In hierarchical dimensions, a plus (+) or minus (−) will appear next to parent-level items. Clicking on this symbol will allow you to show or hide the child items.

At the top of the screen, a series of options appear.

To apply your modified selection of items for the current dimension and see a table of data corresponding to this new selection, click *View Data*. Alternatively, after changing the selection for one dimension, you can click on another dimension to modify, before viewing the new table of data.

For hierarchical dimensions, clicking on *Expand Branch* or *Collapse Branch* will show or hide all the child-level items of the dimension.

*Note: If all members of a hierarchical dimension are selected, the resulting table could contain empty rows or columns corresponding to the higher levels (see below).*
Dimension Member Colour Coding

The Web Browser uses colour coding to help you identify for which dimension members data exists. This is especially important for sparse datasets where data can be found in only a small proportion of dimension combinations. This function has been provided to save time by avoiding searching for non-existent data.

The colours for each dimension member indicate the following:

**Dark Blue**: The item has data.

**Blue**: The item has no data but some of its children have data.

**Grey**: The item has no data.

---

**Figure 11: Dimension Member Colour Coding**

*Remember: The colour coding takes into account the current selection for all other members. In other words, the colours indicate the existence of data for the territory, years, etc. currently selected.*

Another option is to search in the dimension selector to display dimension members matching the search criteria.
Figure 12: Searching in the dimension selector

All matching dimension members will be displayed and automatically selected.

Figure 13: Results of a Search in the dimension selector
Changing the Time Dimension

The options available for changing the selected dates of a time dimension are different from those presented for other dimensions. Dates can be selected either using the **DATE RANGE SELECTION** method or the **TIME PERIOD SELECTION** method.

**Date Range Selection**

All frequencies available for the selected dataset are shown on the screen. If the data is only annual, for example, the controls for selecting quarters, semesters and months will not appear.

First tick the boxes next to the frequencies to be included in the selection. Then choose either a specific range of dates (from 1980 to 2000, for example) by clicking **SELECT DATE RANGE**, or indicate that you wish to have the most recent data available (last 5 years, for example) by clicking **SELECT LATEST DATA**.

When selecting a date range, it is possible to select a start year, quarter, month, etc. and automatically include all available data from that date onward by ticking the box next to **LATEST AVAILABLE DATA**.

![Figure 14: Date Range Selection](image)

**Time Period Selection**

It is also possible to select individual years, quarters, months, etc. by ticking the box next to each item. This is necessary if the selection of dates you wish to include is not a continuous series, but instead contains breaks or time periods outside the main date range.

This works in a manner identical to the general dimension selector described above.

*Note: It is advisable to select a general range of dates using the DATE RANGE SELECTION first, and then include or exclude individual time periods using the TIME PERIOD SELECTION screen.*
The inverse, making changes to a **TIME PERIOD SELECTION using the DATE RANGE SELECTION screen**, may result in some of the individual time periods selected being lost, as the DATE RANGE SELECTION screen is only able to manage continuous, non-broken series of dates.

**Rotating Dimensions by Adjusting the Dataset View**

The different dimensions of a dataset can be displayed on either the horizontal or vertical axis of a table, or in the filter area at the top of the table. To move a dimension from one place to another, click once on the **LAYOUT** entry, under the **CUSTOMISE** menu, above the table. A dialog box will pop up that allows you to visually drag and drop a dimension to the desired location.
If a dimension for which more than one item has been selected is placed into the filter area at the top of the page, a drop-down list will become available. To change the filter currently applied to the data, simply select the appropriate value in the drop-down list.

If the dimension in question does not have a drop-down list, then only one item has been selected. To include other items, either click on the dimension name on the left, or select the dimension in **CUSTOMISE -> SELECTION...** submenu and follow the instructions provided above to select more items.
Figure 17: Changing the view using Territory drop-down list

Table Options

A set of options is available to allow you to format the view of the displayed table, e.g. you can select the number of digits after the decimal point to be displayed, or select a “scaling”, which allows you to display the figures in unit value, hundreds, thousands, hundredths, thousandths, etc.
Showing Codes Instead of Names

All datasets, dimensions and dimension members have short codes as well as names. If it is useful, the codes can be displayed instead of names, by ticking the option **USE CODES** instead of full descriptions.

Hiding Empty Rows or Columns

If a table of data has many rows or columns containing no data, it can be useful to hide such empty spots, in order to condense the table and improve readability.
Figure 20: Using the **HIDE EMPTY ROWS AND COLUMNS** options

**Exporting to Excel**

Click on the **EXPORT** and then to **EXCEL** button to export your data selection to a MS Excel file. You will be prompted to either save the MS Excel file to disk or open it directly. As your computer settings may not allow for the file to be directly opened, it is recommended that you save the file to your desktop and then open it from there.

Figure 21: Using the **EXPORT TO EXCEL** option to save data in Excel format

Data flags saved as Excel footnotes
Exporting to a CSV file

Larger tables can be saved to a CSV file; to do so simply click on the Text File (CSV) button above the table.

You will be prompted to select a dimension label format and a column separator. The dimension label format determines what information will be included in your file for each territory, variable, etc. The column separator determines which character will be used to separate columns in the file.

Enter your e-mail address in the box in the middle of the dialog box and click the Export to Text File button. The file will be created and an e-mail sent to you informing you how to retrieve it. The file will not be included as an attachment to the e-mail, as it could potentially be very large. Instead the e-mail will contain a link allowing you to download the file. The file is in compact format “.gz” and the free SevenZip software is needed to read it. All the instructions are in the mail.

Downloading related files and large data selections

Any number of documents related to a given dataset may be available for download through the data browser. To view the list of files related to some dataset, choose the Related Files entry under the Export menu, located at the top of the table.
A dialog window containing a list of files related to the dataset will be shown up.

The list of files may contain documents in any format, or large prepared selections of the data in text format. Click on the name of a file to begin downloading.
Data charts

The browser platform includes a feature for creating dynamic charts ( ) allowing the visualization and animation of Time Graph, Bar Chart and Scatterplot, according to the current selection data.

The Time Graph component is an enhanced line graph that displays the time evolution of a dimension for one or more data items of a second dimension. The x-axis represents the year and the y-axis the indicator chosen, each line represents the trend over the time of the chosen indicator for each data items of the second dimension, allowing the comparison. The indicator displayed can be change from the combo box on the left side.

The navigation options in the dialog box graphic include:
- a button to animate the graph by regular time intervals;
- a combo box to change the language;
- a window for setting the layout parameters and that can be accessed by clicking the button;
- icon to access the online help.

![Time Graph](image)

Figure 25.a: Time Graph

The Bar Chart component is a good component to display a set of indicators with data sorted and easily comparable. Each indicator for each element of a given size is represented by a bar with a height proportional to its value. This component is enriched by some of the features
(change language, possibility to set parameters of layout, online help) which we have spoken above. Moreover through the icon the user have the opportunity to export the initial view to jpg or png file format.

Italy, number of active enterprises, total, totale, total, total, total, all territorial distribution

The ScatterPlot component, in its simplest version, shows the relationship between two indicators, positioning them on X and Y axes. In the advanced version up to four dimensions can be represented. Time dimension is represented activating the animation through the button; two dimensions are represented on the X and Y axes, as we already said for the basic version, and the fourth is represented by the size of the bubbles. It is activated using the combo box on top of the graph and the maximum size of the bubbles can be adjusted by a slider on the right side.

Figure 25.b: Bar Chart
Figure 25.c: ScatterPlot
Viewing Metadata

Metadata can be viewed at all levels, from the dataset-level abstract down through to cell level footnotes. The presence of metadata is flagged by a small, red "i".

Clicking on the red "i" will show the related metadata panel located at the right-hand area of the screen. The panel can thus be newly hidden by clicking the tab on its edge.

Flags

Certain pieces of metadata, which are recurrent within the dataset, are treated differently by the system and referred to as flags. These flags appear as letters in parentheses in the cell alongside the piece of data to which they correspond. Below any data table containing flags will appear a legend, explaining the meaning of each letter used as a flag in the table.
Figure 28: Viewing the Data Flag Legend

**Working with Queries**

**Saving a Query**

When working with queries most features require you to log in, e.g. the **MY QUERY -> SAVE** button is only displayed when you logged in via the **LOGIN** link at the top right of the page. Once logged in your name will be displayed at the top right of the page. Your data selection, or query, can be saved at any time by clicking **My Query** and then **Save** button above the table. On subsequent visits to the browser, this saved query can be opened to view the latest data available within the selection.
Merge Queries

Data from two or more datasets can be viewed in a combined table, using the **Merge Queries** function. Any common dimensions between the different datasets should be automatically combined. Other dimensions will automatically be displayed along the vertical axis, underneath section headings indicating which dataset they come from.

-Combining Queries

To create a Combined Query, first create a customized view in one dataset. Then click **Combine** entry in the **My Queries** menu located above the table to put this query into your "data basket".
Repeat this for each dataset that you wish to include in the multi-dataset query.

**Viewing the Query**

When all selections of data have been added, locate the **Combine Queries** panel in the top left area of the screen and click **View** to see the combined data table.

**Figure 32: Viewing a Combine Query**

**Modify a selection**

To modify any of the selections of data in the view, click the name of the dataset in the **Merger Queries** panel, and then click **Edit**. The selection will appear on screen in the same format as a single-dataset query. Make your changes and click once again the **Combine** entry under the **My Query** menu. Click **View** in the **Combine Queries** panel to see the modified query.

**Remove a selection**

To remove one of the selections in a Merged Query, click the name of the dataset in the **Combine Queries** panel, and then click **Remove**. Click **View** in **Combine Queries** panel to see the modified query.
Saving a Multi-Dataset Query

You can save your Multi-Dataset Query by clicking **Save** in the **Combine Queries** panel; typing a name and then clicking **Save** once again.

Viewing Frequently Requested and Saved Queries

Clicking on **My Queries** in the left panel displays the list of all queries which you have personally saved together with a list of recent queries. This feature requires the user to login.

![Figure 33: Viewing saved and recent queries](image-url)
The **Query Manager** allows you to re-sort and save the list of queries as well as open and delete any of them.

Additionally, clicking the **Share** button displays a reference URL pointing to your saved query which can be shared with colleagues, e.g. sent via email. Appropriate access rights will be required to view the query.

**Contact us**

The user can find Istat contacts using the Contact Us link, at the top of the screen.

You can contact:

- the contact centre: for information on data, publications, certifications, microdata files, cartographies, historical researches and customised elaborations;
- **comunica@istat.it**: to contact the website editorial office;
- **webmaster@istat.it**: to report technical problems or improvement suggestions

or you can consult the **Organisation chart** for contact details of other offices.
Figura 35: Opening the ISTAT contact page