



# Integrated Automated (Statistical) Information System

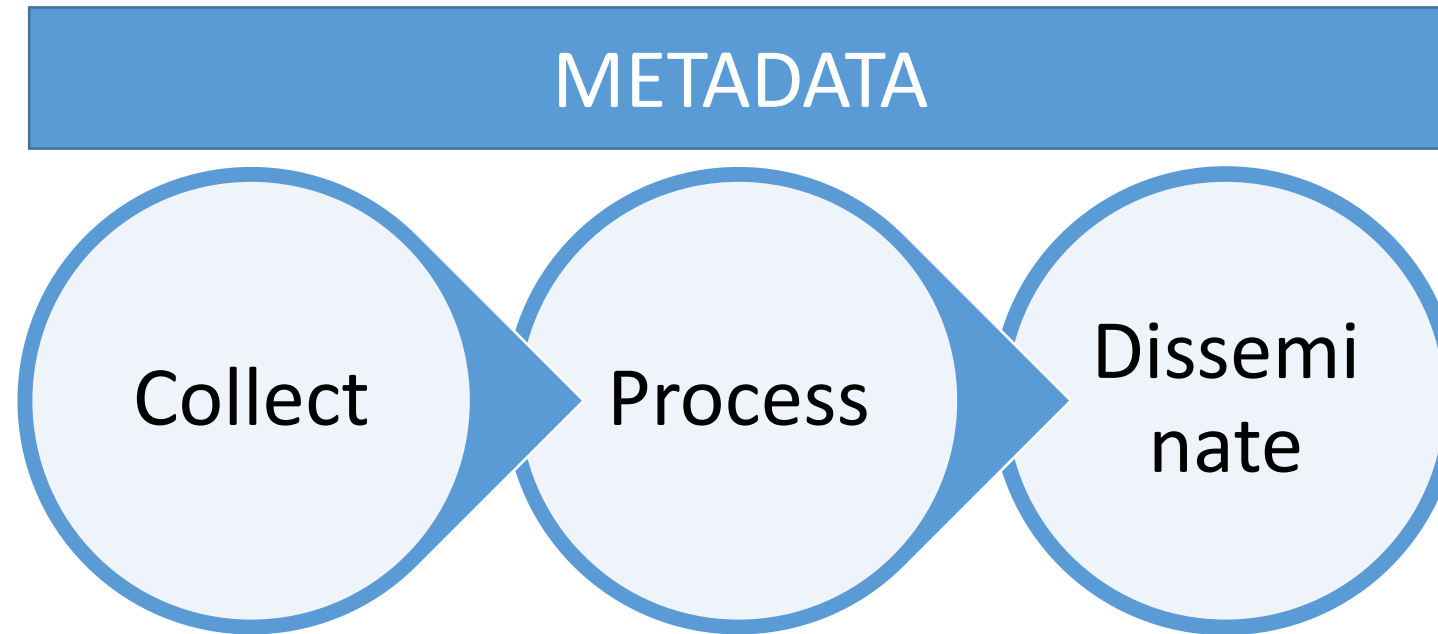
*Vasile Bujor, NBS*

*Presentation of Project Results*

*5 July 2022, Palace of the Republic, Chisinau, Republic of  
Moldova*



# IAIS Data Flow





# IAIS Principles

- Metadata driven
- Processes and methods standardization
- Data storage centralization
- Data process centralization
- Governance (process harmonization between statistical programs)
- Security



# IAIS Modules

- Metadata
  - Referential
  - Structural
- Collection
  - Data entry builder
  - Fieldwork monitoring dashboard
  - Survey Management system
- Process
  - Centralized data processing system
- Dissemination
  - Statistical database (statbank)



# Metadata

- Standardized documentation of surveys (statistical programs) using GSBPM and GSIM

- [Link](#)

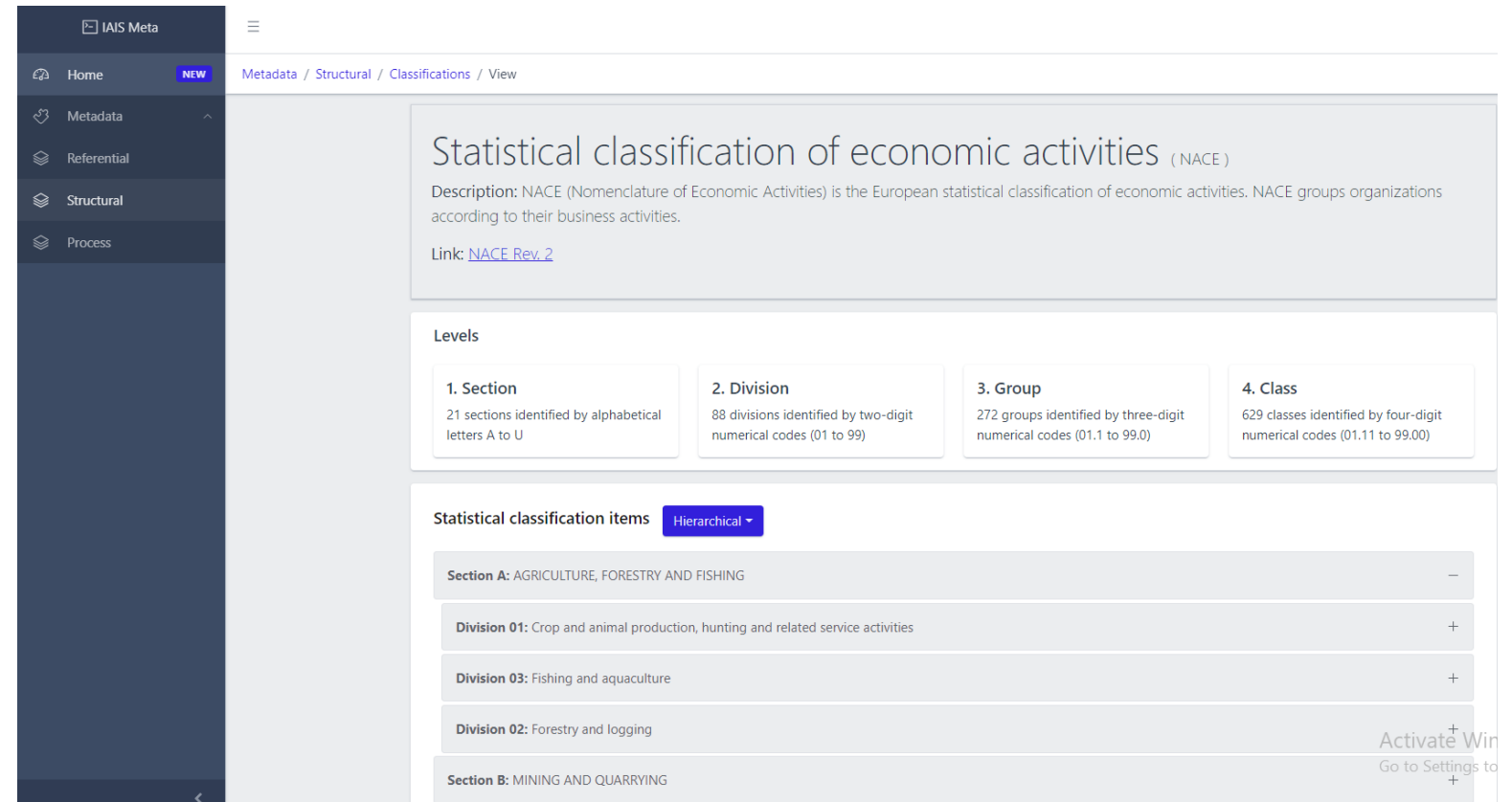
The screenshot shows the IAIS Meta web application interface. The main content area displays the 'International Standard Classification of Occupations' (ISCO) metadata page. The page includes a search bar, a description of ISCO, and a 'Process documentation' table. The table lists various steps in the process, such as 'Identify needs', 'Consult and confirm needs', 'Establish output objectives', and 'Identify concepts', each with associated GSBPM, Name, Frequency, Next step, and Version information. A navigation bar at the bottom indicates the current page is 1 of 2, and the application is powered by CoreUI for Vue.

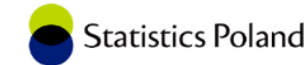
ID	GSBPM	Name	Frequency	Next	Version	Actions
1.1	Identify needs	Identify needs for LFS	ONCE	1.2	1.0	<a href="#">?</a>
1.2	Consult and confirm needs	Consult and confirm needs for LFS	ONCE	1.3	1.0	<a href="#">?</a>
1.3	Establish output objectives	Establish output objectives for LFS	ONCE	1.4	1.0	<a href="#">?</a>
1.3	Establish output objectives	Establish output objectives for LFS	ONCE	1.4	1.1	<a href="#">?</a>
1.4	Identify concepts	Identify concepts for LFS	ONCE	2.1	1.0	<a href="#">?</a>



# Metadata

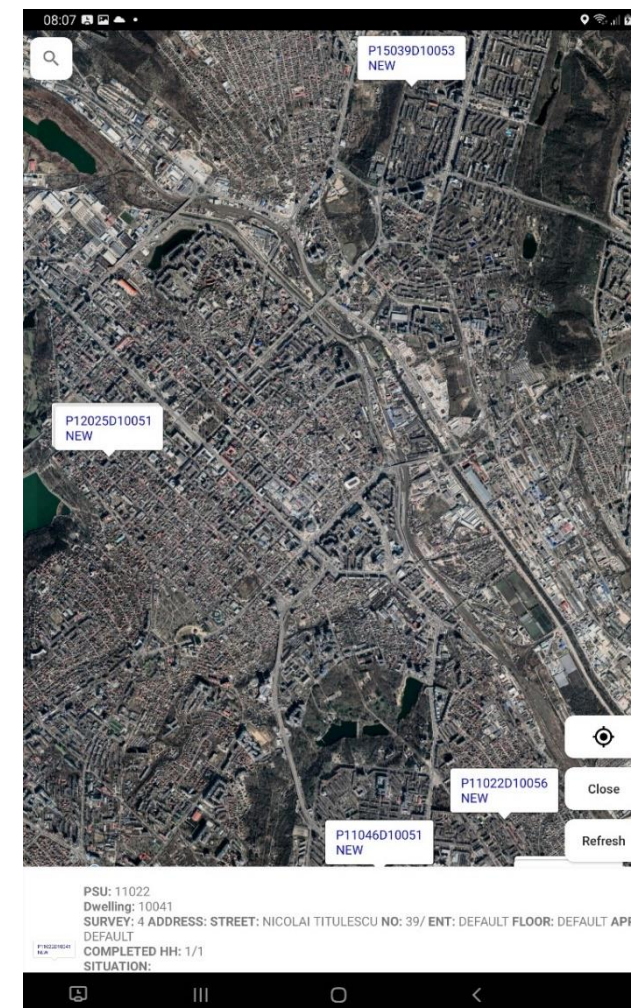
- Standardized documentation of dataset, indicators, variables, statistical classification, statistical concepts etc.
- [Link](#)





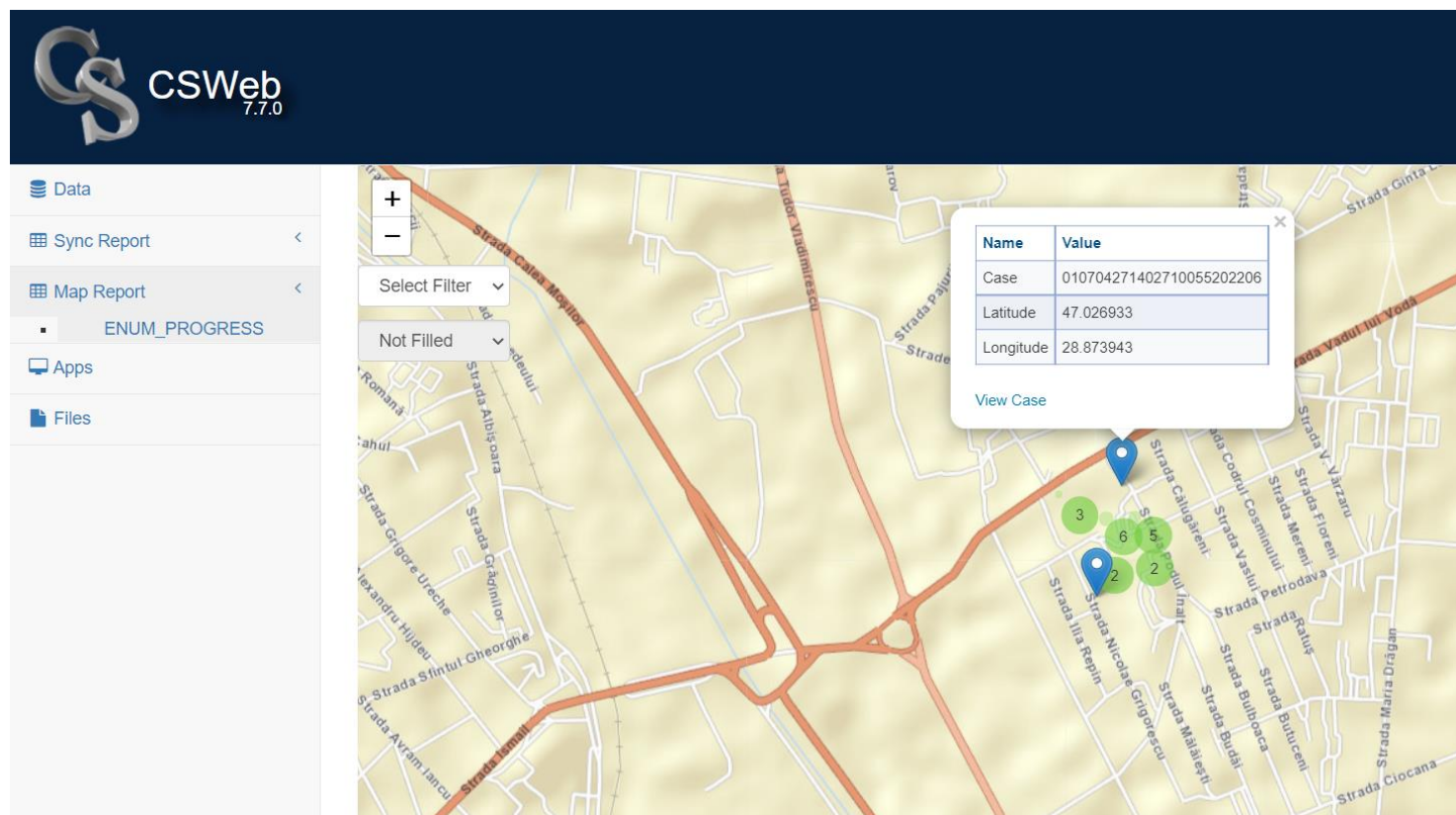
# Data Collection

- Transition to CAPI for household surveys using CPro
- CAPI (Digital data collection) advantages
  - More question types supported
  - Many devices into one
  - Automatic questionnaire flow
  - Real time consistency checking
  - Real time data synchronization & evaluation
  - Field work monitoring



# Data Collection

- Real time data synchronization
- Data server dashboard



The screenshot displays the CSWeb 7.7.0 interface. On the left is a navigation menu with options: Data, Sync Report, Map Report, ENUM\_PROGRESS, Apps, and Files. The main area shows a map with several data points (green circles with numbers 2, 3, 6) and a blue location pin. A popup window is open over the pin, displaying the following data:

Name	Value
Case	010704271402710055202206
Latitude	47.026933
Longitude	28.873943

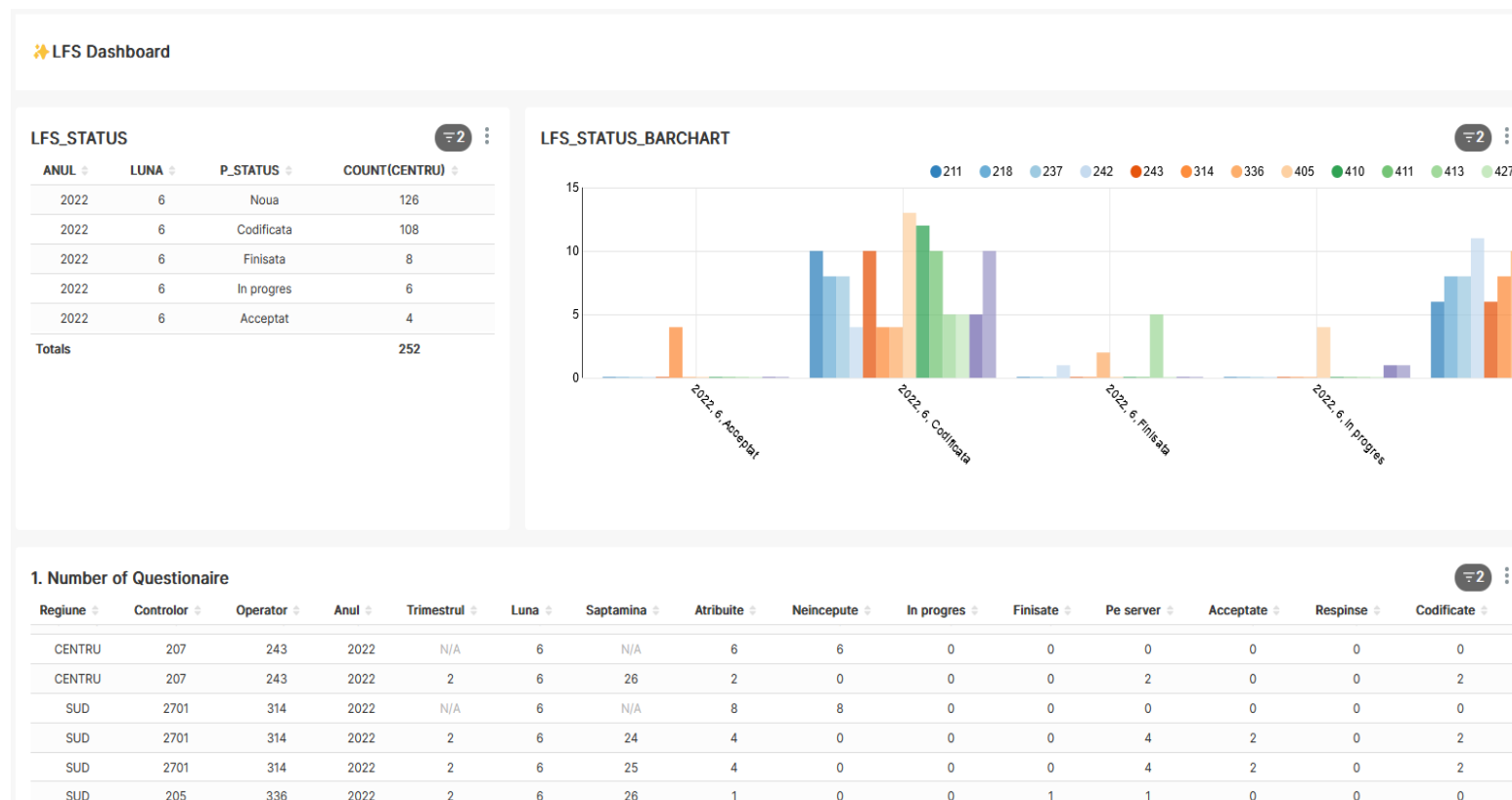
Below the table is a "View Case" link. The map also includes a "Select Filter" dropdown and a "Not Filled" dropdown.





# Data Collection

- Fieldwork monitoring dashboard
- Fieldwork indicators calculation
- Real-time data quality monitoring



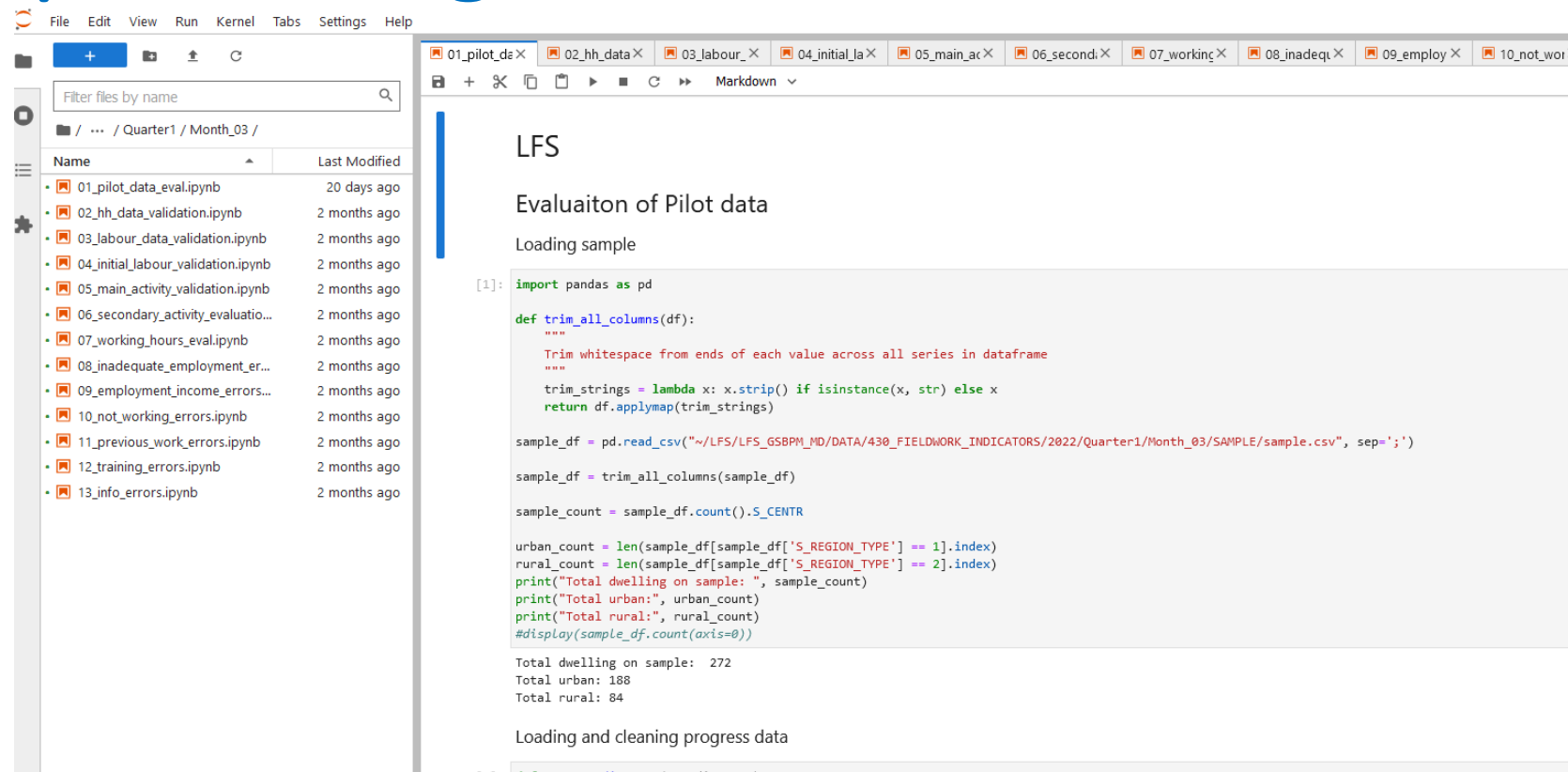


# Data processing

- Centralized data processing system using Jupyter notebooks:
  - Governance
  - Standardization of processes
  - Harmonization of processes
  - Organization of processes (i.e. according to GSBPM sub-phases)
- Support for all open source statistical programming languages (R, Python, Julia, SQL, Spark... etc)
- Server side data processing.
- Security and data loss prevention



# Data processing



File Edit View Run Kernel Tabs Settings Help

Filter files by name

... / Quarter1 / Month\_03 /

Name	Last Modified
01_pilot_data_eval.ipynb	20 days ago
02_hh_data_validation.ipynb	2 months ago
03_labour_data_validation.ipynb	2 months ago
04_initial_labour_validation.ipynb	2 months ago
05_main_activity_validation.ipynb	2 months ago
06_secondary_activity_evaluatio...	2 months ago
07_working_hours_eval.ipynb	2 months ago
08_inadequate_employment_er...	2 months ago
09_employment_income_errors...	2 months ago
10_not_working_errors.ipynb	2 months ago
11_previous_work_errors.ipynb	2 months ago
12_training_errors.ipynb	2 months ago
13_info_errors.ipynb	2 months ago

## LFS

### Evaluaiton of Pilot data

#### Loading sample

```
[1]: import pandas as pd

def trim_all_columns(df):
    """
    Trim whitespace from ends of each value across all series in dataframe
    """
    trim_strings = lambda x: x.strip() if isinstance(x, str) else x
    return df.applymap(trim_strings)

sample_df = pd.read_csv("~/LFS/LFS_GSBPM_MD/DATA/430_FIELDMETHODS/2022/Quarter1/Month_03/SAMPLE/sample.csv", sep=';')

sample_df = trim_all_columns(sample_df)

sample_count = sample_df.count().S_CENTR

urban_count = len(sample_df[sample_df['S_REGION_TYPE'] == 1].index)
rural_count = len(sample_df[sample_df['S_REGION_TYPE'] == 2].index)
print("Total dwelling on sample: ", sample_count)
print("Total urban:", urban_count)
print("Total rural:", rural_count)
#display(sample_df.count(axis=0))

Total dwelling on sample: 272
Total urban: 188
Total rural: 84
```

#### Loading and cleaning progress data

```
[2]: def convert_dataframe(dataframe):
```



# Dissemination

- Automatization of output table creation
- PxWeb/Statbank
- New website



## IAIS future plans

- Improving the metadata system.
- Integrating all LFS processes into IAIS.
- Implementing CAPI for all social surveys.
- Integrating all social surveys into IAIS.
- Improving the entire system.

EU funded Project ENI/2019/406-262

“Technical Assistance to Support the National Bureau of Statistics of the Republic of Moldova”



# Thank you for the attention!

*Project implemented by the Italian Geographical Society and Roma Tre University (SGI URS), the Italian National Institute of Statistics (Istat), Statistics Denmark and Statistics Poland in partnership with the National Bureau of Statistics of the Republic of Moldova*