



SURVEY MONITORING WITH PARADATA: LEVERAGING THE SURVEY SOLUTIONS PARADATA VIEWER APPLICATION IN R

Michael Wild, World Bank, Austria

Ciprian Alexandru, Ecological University of Bucharest, Romania

AGENDA

Introduction to paradata

Key elements

Application

Key Features

More details ...

PARADATA – AN IMPORTANT VECTOR FOR QUALITY

- data quality
- nonresponse rates on nonresponse bias
- survey performance
- operational efficiency

- Census & Surveys



DALL·E

KEY ELEMENTS



Survey Solutions is a free software for data collection and survey management developed in the Data group of The World Bank.

Survey Solutions Paradata Viewer Application

SURVEY SOLUTIONS PARADATA VIEWER APPLICATION

- The Survey Solutions Paradata Viewer Application is part of a comprehensive toolkit designed to enhance the process of survey implementation through Survey Solutions.
- Developer: Michael Wild (Author, maintainer)
- Data analyst: Ciprian Alexandru

- <https://github.com/michael-cw/susoparaviewer>

- `devtools::install_github("michael-cw/susoparaviewer")`

- General Agricultural Census
- Population & Housing Census



KEY FEATURES

1. Visualization at the questionnaire level: Review a questionnaire's average time profile, average timings and number of invalid questions.

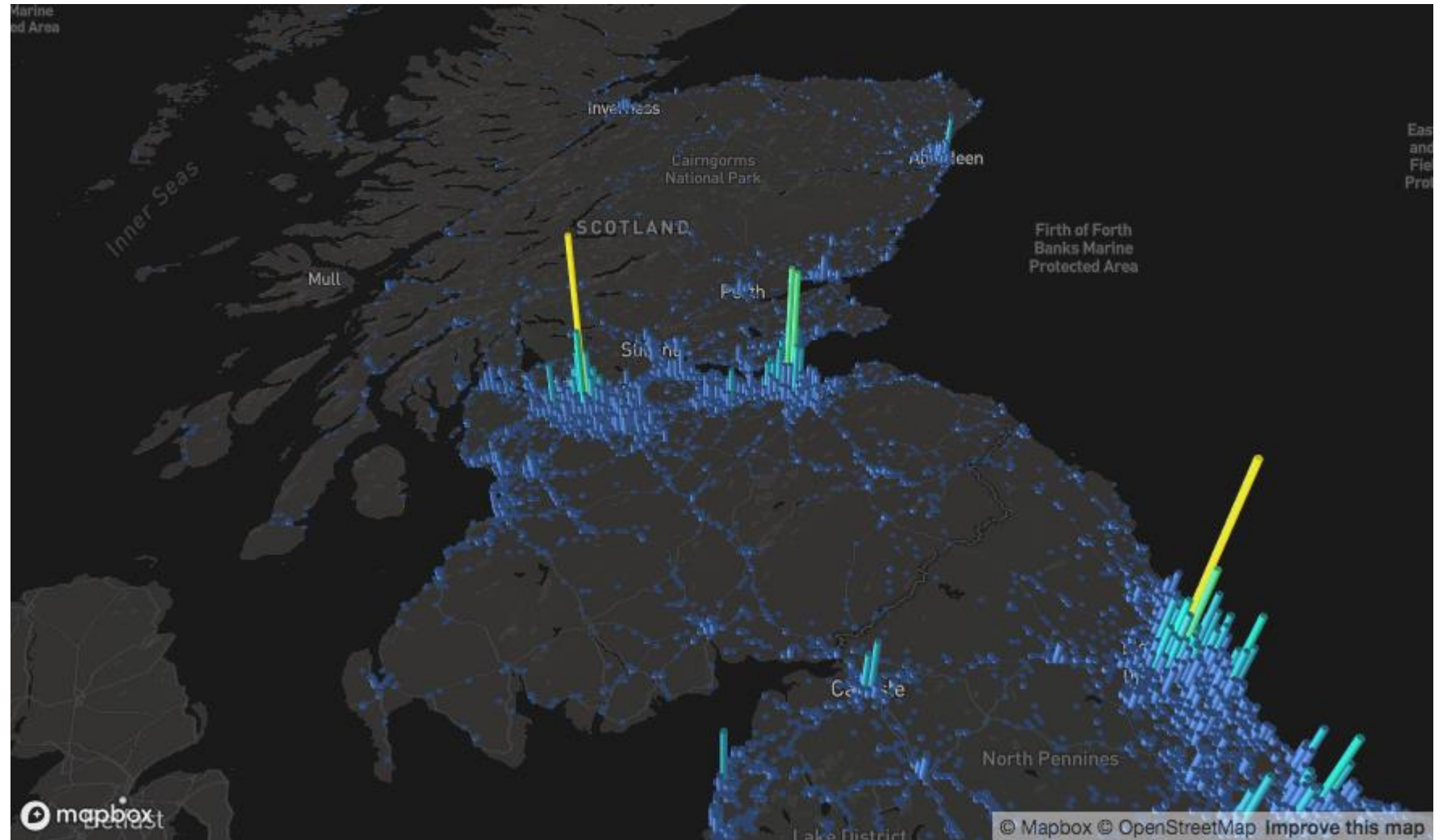
2. Visualization at the interviewer level: Review average completion times, average pace and number of response removals at the level of the interviewer

...3-8

3. VISUALIZATION AT THE GEOGRAPHIC LEVEL

Review average completion times, average pace and number of response removals by geographic regions.

Visualization is either through a [mapdeck](#) (requires an API key) or a [leaflet](#) map.



KEY FEATURES

4. Admin Interface: This intuitive interface lets you:

1. Directly connect to your Survey Solutions server to load the paradata.
2. Specify the time interval when new data should be downloaded.

5.Flexibility in data input: Data can be provided as:

1. File upload.
2. Directly from the server through the API.
3. From a local directory.

6.Download the data: The application also allows to download the data in a more convenient format and by event type.

7.Direct review of interviews/interviewers: Allows you to directly connect to the Survey Solutions server and review interviews and interviewers.

8.Report creation: Generates a full paradata report as a word document at the:

1. *Questionnaire* level,
2. *Interviewer* level,
3. *Map* level³.

INSTALLATION

1. Install R: <https://cran.r-project.org/mirrors.html> (version 4.1.1 or greater)
2. Install R Studio: <https://rstudio.com/products/rstudio/download/> (version 1.2.5001-3 or newer)
3. Make sure the devtools package is installed, if not install it with:
 - `install.packages("devtools")`
4. After that install the actual package:
 - `devtools::install_github("michael-cw/susoparaviewer")`

RUNNING THE APPLICATION

- Running the application interactively

```
library(susoparaviewer)
```

```
susoparaviewer::runParaApp()
```

- Running the application on a Shiny Server

...or run the application on your own shiny server. In this case, create a directory, i.e. susoparaviewer and put a file called app.R with the following code into it:

```
library(susoparaviewer)
```

```
susoparaviewer::runParaAppServer()
```

GITHUB

- <https://michael-cw.github.io/susoparaviewer/index.html#experimental-optional-feature-gpslogger-tracking>

THE APP ITSELF...

The screenshot displays the RPL21 Visualizer & Converter Paradata application. It features three main data visualization panels:

- QUESTIONNAIRE:** A horizontal bar chart titled 'Average Questionnaire Completion Time' showing completion times for various questionnaires (int_0001 to int_0010).
- INTERVIEWER:** A horizontal bar chart titled 'Average Pace' showing the deviation from the question/answer time for different interviewers (int_0001 to int_0010).
- QUESTIONS:** A horizontal bar chart titled 'Number of Response Removals' showing the number of removals for various questionnaires.


Below the charts is a 'Detaili interviewator' section with a table:

responsibile	Ar (DuratiaQuestionnaire)	NumarInterviu	Fast	Slow	MaxRata
int_0008	0.40	07	1000	070	700

THE APP ITSELF...

The screenshot displays the 'RPL25 Monitor National' application. The top navigation bar contains the title 'RPL25 Monitor National' and a 'Home' button. Below the navigation bar is a map of a region with several highlighted areas in yellow and orange. To the left of the map is a sidebar menu with a search bar and a list of items. At the bottom right, there is a small 'Loading...' indicator.

THE APP ITSELF...



THE WORLD BANK
IBRD · IDA | WORLD BANK GROUP
Development Economics + Data

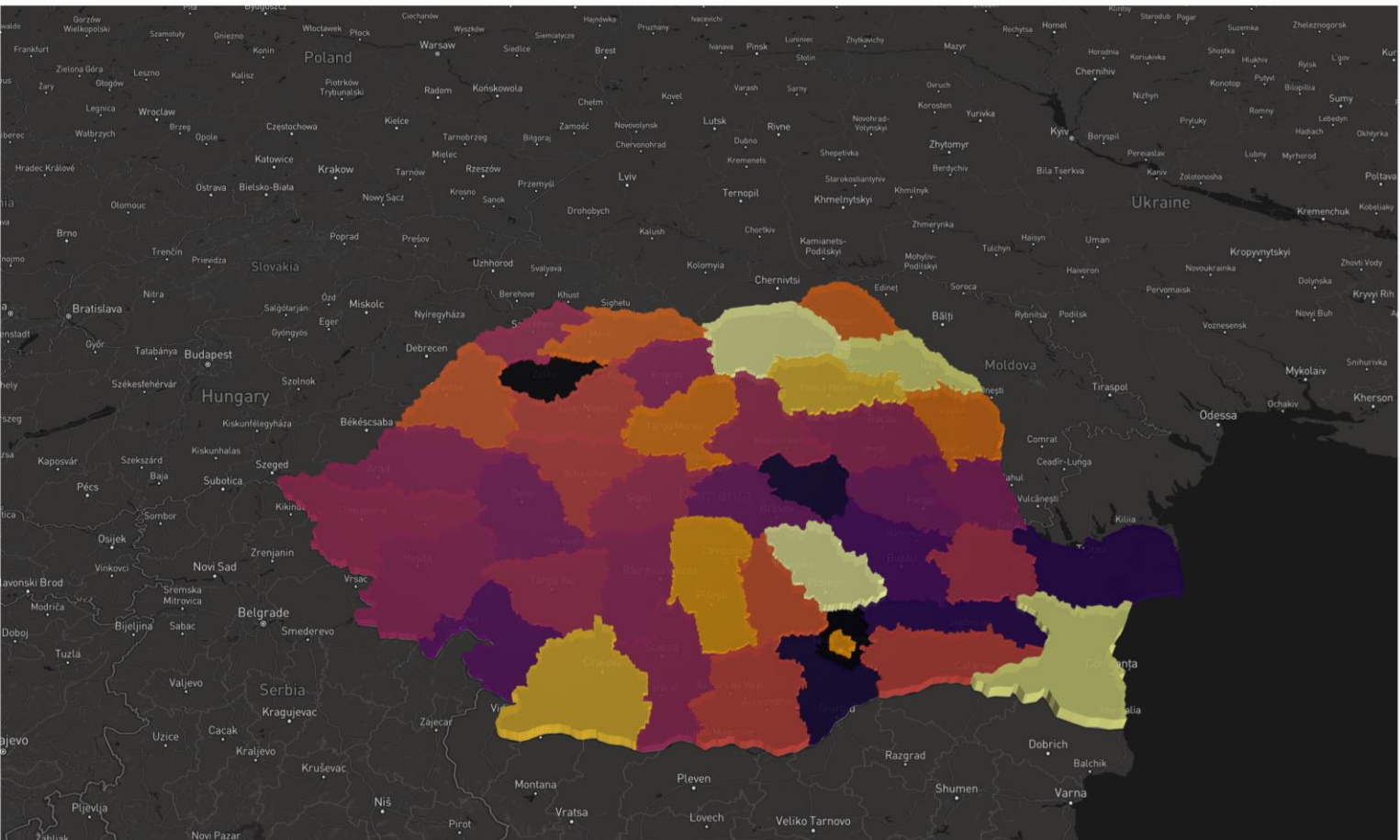
RPL21 Monitor National

Logged in as michaelow [Logout](#)

Pentru a schimba unghiul de vizualizare al hărții, utilizați CTRL + MOUSE POINTER

Pentru a primi mai multe informații despre un județ, faceți clic pe județul de pe hartă!

	n_adresa	n_sector	cawi_gospodării	capi_gospodării	total_gospodării	cawi_durata_medie	capi_durata_medie	n_terminat	n_in_curs	n_nu_inceput	
1	Toată România	39413	568	1564	38210	39774	55.79	28.35	36715	1525	1173



N_adresa SUM of Count of addresses in segment files

N_locuinte.... SUM of Count of dwellings in segment files

CAWI ... Count of INTERVIEWS collected via web-mode (sums)

CAPI ... Count of INTERVIEWS collected via tablet (sums)

T../In../Nu...Count of the building, one of not-started/in progress/completed*

tempgr	N_adresa	N_locuinte	CAPI	CAWI	TOTAL
TOTAL	1201	2051	1133	43	1176

g_jud_code	sector	g_sat_code	N_adresa	N_locuinte	CAPI	CAWI	TOTAL
403	1	179141	1201	54	40	3	43

_responsibile	g_jud_code	sector	g_sat_code	CAWI_dur	CAPI_dur	Terminat	In_curs	Nuinceput
int_s1001	403	1	179141	24.00	3.57	30	10	0

CAWI_dur..... Average duration in the segment

CAPI_dur..... Average duration in the segment



Tools Directory ../phc_paradata

The image displays a collage of screenshots from the RPL21 Visualizator & Converter Paradata software. The main interface features three primary visualization panels: 'Average Questionnaire Completion Time' (a stacked area chart), 'Average Pace' (a line graph showing pace over time), and 'Number of Response Removals' (a horizontal bar chart). A 'Detalii interviator' (Interviewer Details) section is also visible, containing a table of user information.

Table: Detalii interviator

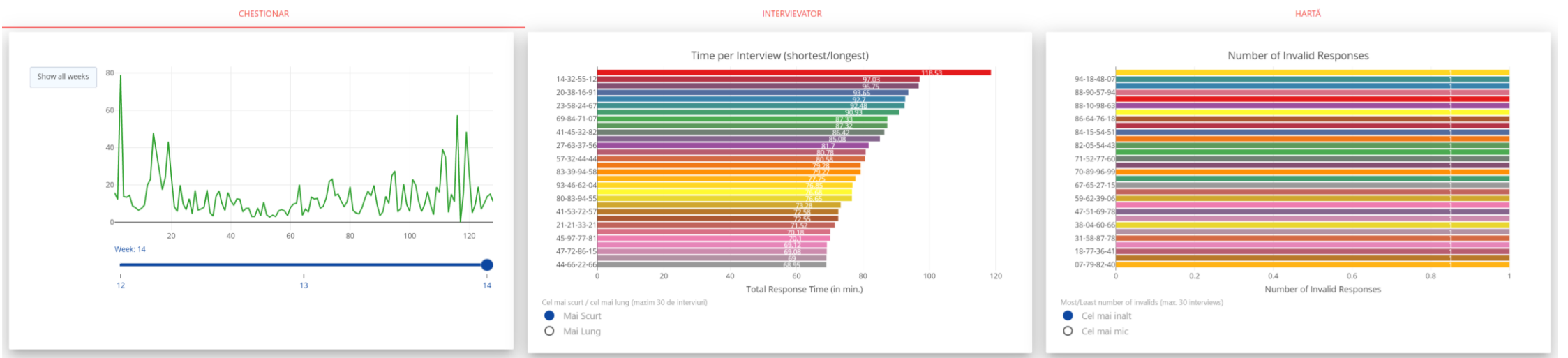
Key	Username	Avansat/Avanzat	Completat/Complet	Validat	Validat/Validat
1	AV	AVAN	AVAN	AVAN	AVAN

Table: Data Viewer

Key	Avatar	Avatar	Avatar	Avatar	Avatar	Avatar	Avatar	Avatar	Avatar	Avatar
00-00-00-00	AvatarSet	AvatarSet	AvatarSet	AvatarSet	AvatarSet	AvatarSet	AvatarSet	AvatarSet	AvatarSet	AvatarSet
00-00-00-00	AvatarSet	AvatarSet	AvatarSet	AvatarSet	AvatarSet	AvatarSet	AvatarSet	AvatarSet	AvatarSet	AvatarSet
00-00-00-00	AvatarSet	AvatarSet	AvatarSet	AvatarSet	AvatarSet	AvatarSet	AvatarSet	AvatarSet	AvatarSet	AvatarSet
00-00-00-00	AvatarSet	AvatarSet	AvatarSet	AvatarSet	AvatarSet	AvatarSet	AvatarSet	AvatarSet	AvatarSet	AvatarSet
00-00-00-00	AvatarSet	AvatarSet	AvatarSet	AvatarSet	AvatarSet	AvatarSet	AvatarSet	AvatarSet	AvatarSet	AvatarSet
00-00-00-00	AvatarSet	AvatarSet	AvatarSet	AvatarSet	AvatarSet	AvatarSet	AvatarSet	AvatarSet	AvatarSet	AvatarSet
00-00-00-00	AvatarSet	AvatarSet	AvatarSet	AvatarSet	AvatarSet	AvatarSet	AvatarSet	AvatarSet	AvatarSet	AvatarSet
00-00-00-00	AvatarSet	AvatarSet	AvatarSet	AvatarSet	AvatarSet	AvatarSet	AvatarSet	AvatarSet	AvatarSet	AvatarSet
00-00-00-00	AvatarSet	AvatarSet	AvatarSet	AvatarSet	AvatarSet	AvatarSet	AvatarSet	AvatarSet	AvatarSet	AvatarSet
00-00-00-00	AvatarSet	AvatarSet	AvatarSet	AvatarSet	AvatarSet	AvatarSet	AvatarSet	AvatarSet	AvatarSet	AvatarSet

The interface also includes a 'Chestionar/Detalii intrebare' (Questionnaire/Question Details) section and a 'Data Viewer' table showing a list of data entries with columns for key, avatar, and other attributes.

Tools Directory ...



Chestionar/Detalii întrebare

key	Questions	AvTimeQuestion	CompletionTime	MainRole	MainResponsible
07-79-82-40	97	6.568	21.57	1	int_ms007

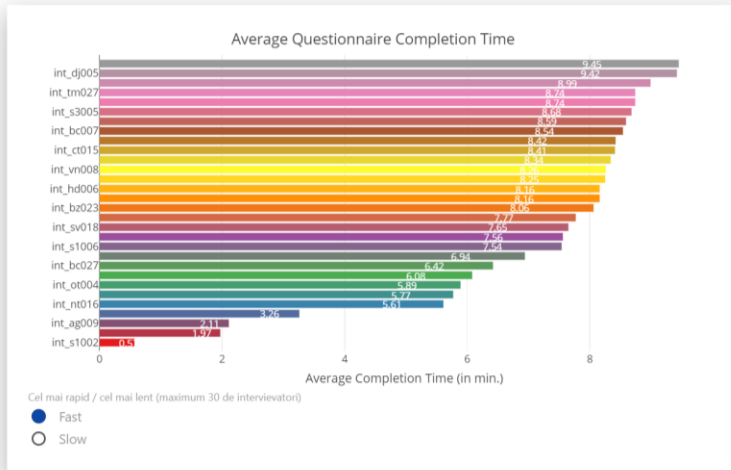
Dacă faceți clic pe TASTĂ, chestionarul se va deschide într-o fereastră nouă (trebuie să fiți conectat pentru a efectua această acțiune)

[DESCĂRCAȚI RAPORTUL COMPLET](#)

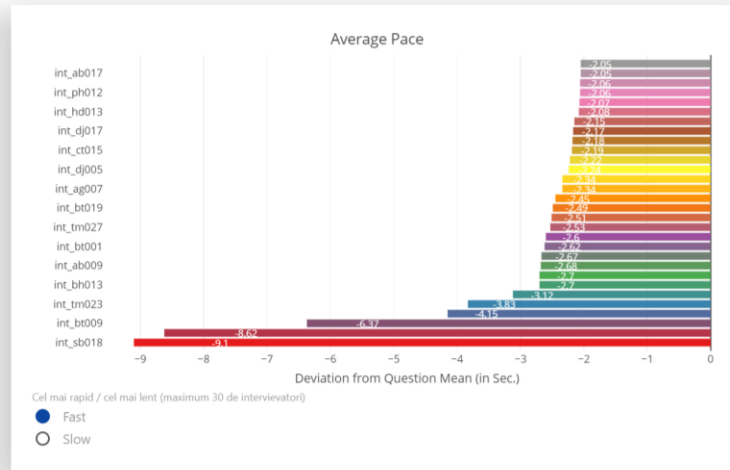
Tools Directory ...

RPL21 Vizualizator & Convertor Paradata

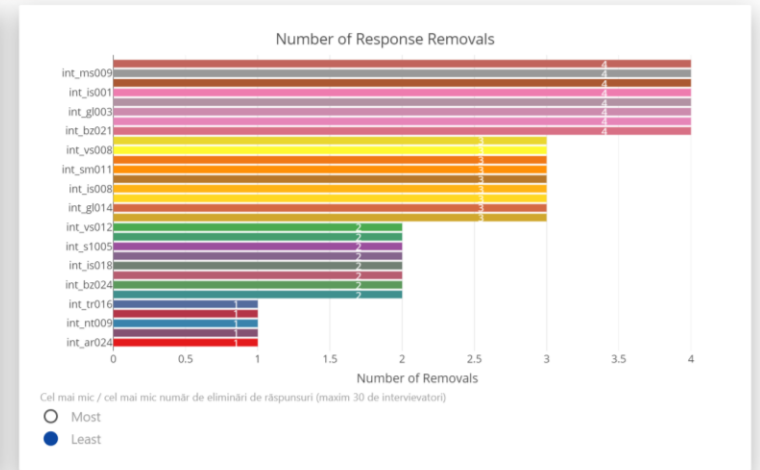
CHESTIONAR



INTERVIEWATOR



HARTĂ



Detalii interviator

[DESCĂRCAȚI RAPORTUL COMPLET](#)

Tools Directory ...

≡ RPL21 Vizualizator & Convertor Paradata

How do you want to load the data?

- File
- Server
- LocalFile

Please select the paradata file

- KeyAssigned
- CommentSet
- Completed
- AnswerSet**
- AnswerRemoved
- QuestionDeclaredInvalid
- QuestionDeclaredValid
- Restarted
- Reject

Data Viewer

Show

10

Search: _____

entries

key	counter	action	responsible	role	tz	var	response6	response7	response8
60-08-63-36	4	AnswerSet	int_mh003	1	03:00:00	g_jud	<NA>	<NA>	<NA>
60-08-63-36	5	AnswerSet	int_mh003	1	03:00:00	g_muni	<NA>	<NA>	<NA>
60-08-63-36	6	AnswerSet	int_mh003	1	03:00:00	g_sat	<NA>	<NA>	<NA>
60-08-63-36	7	AnswerSet	int_mh003	1	03:00:00	g_strada	<NA>	<NA>	<NA>
60-08-63-36	8	AnswerSet	int_mh003	1	03:00:00	g_numar	<NA>	<NA>	<NA>
60-08-63-36	9	AnswerSet	int_mh003	1	03:00:00	g_bloc	<NA>	<NA>	<NA>
60-08-63-36	10	AnswerSet	int_mh003	1	03:00:00	g_geoLocation	<NA>	<NA>	<NA>
60-08-63-36	11	AnswerSet	int_mh003	1	03:00:00	g_jud_code	<NA>	<NA>	<NA>
60-08-63-36	12	AnswerSet	int_mh003	1	03:00:00	g_muni_code	<NA>	<NA>	<NA>
60-08-63-36	13	AnswerSet	int_mh003	1	03:00:00	g_sat_code	<NA>	<NA>	<NA>

Showing 1 to 10 of 5,168,426 entries

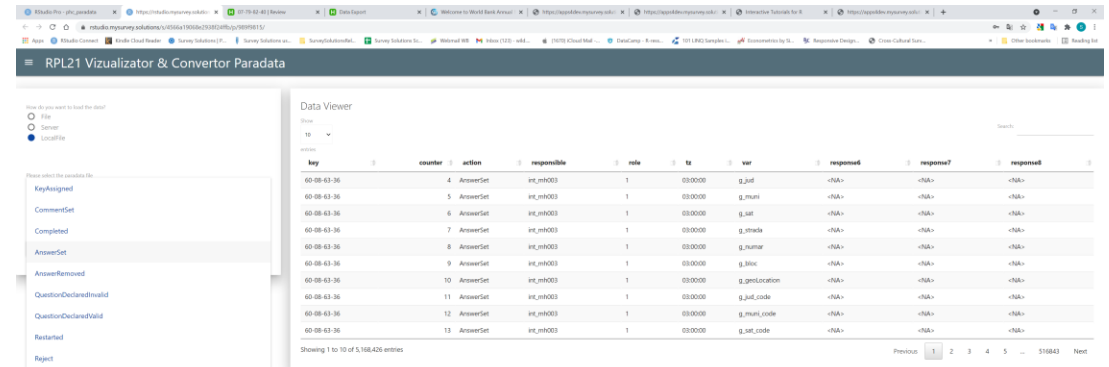
Previous **1** 2 3 4 5 ... 516843 Next

Load data into ../phc_paradata

Upload file exported from survey solutions → **FILE**

Directly from Survey Solutions via API → **SERVER**

From previous (manual) upload → **LOCALFILE**



RPL21 Vizualizator & Converter Paradata

How do you want to load the data?

- File
- Server
- Localfile

From which file do you want to load the data?

- KeyAssigned
- CommentSet
- Completed
- AnswerSet
- AnswerRemoved
- QuestionDeclaredInvalid
- QuestionDeclaredValid
- Restarted
- Rejected

Data Viewer

key	counter	action	responsible	role	ts	var	response6	response7	response8
60-08-63-36	4	AnswerSet	HT_MH003	1	03:00:00	g_jud	<NA>	<NA>	<NA>
60-08-63-36	5	AnswerSet	HT_MH003	1	03:00:00	g_num1	<NA>	<NA>	<NA>
60-08-63-36	6	AnswerSet	HT_MH003	1	03:00:00	g_jst	<NA>	<NA>	<NA>
60-08-63-36	7	AnswerSet	HT_MH003	1	03:00:00	g_trade	<NA>	<NA>	<NA>
60-08-63-36	8	AnswerSet	HT_MH003	1	03:00:00	g_numar	<NA>	<NA>	<NA>
60-08-63-36	9	AnswerSet	HT_MH003	1	03:00:00	g_bloc	<NA>	<NA>	<NA>
60-08-63-36	10	AnswerSet	HT_MH003	1	03:00:00	g_geolocation	<NA>	<NA>	<NA>
60-08-63-36	11	AnswerSet	HT_MH003	1	03:00:00	g_jud_code	<NA>	<NA>	<NA>
60-08-63-36	12	AnswerSet	HT_MH003	1	03:00:00	g_num_code	<NA>	<NA>	<NA>
60-08-63-36	13	AnswerSet	HT_MH003	1	03:00:00	g_sst_code	<NA>	<NA>	<NA>

Showing 1 to 10 of 5,168,426 entries



.	mean_durati ion	mean_durationNOB REAK	mean_startH our	mean_RespTi me	N_o bs
Toată Român ia	59.82	11.32	10.68158	5.58	961

responsi ble	mean_durati on	mean_durationNOBR EAK	startHo ur	mean_RespTi me	N_ob s
int_bt008	46.84	7.50	13	3.78	56

.	Av_ResponseTime	Av_Duration	N_questions
Toată România	8.190066	8.190066	151

counterMedian	var	Av_ResponseTime	Av_Duration	N_obs	tot
152.5	AA_ALM	13.28	11.96	1075	Toată România

mean_duration (min) Mean duration of interview, calculated from 1st ANSWER SET TO LAST.

mean_durationNOBREAK (min) Mean duration of interview, calculated from 1st ANSWER SET TO LAST but without BREAKS (breaks as signalled by tablet, or when response time >3min).

mean_startHour ... (24h) mean of the hour when most of the interviews did their first question

mean_RespTime ... (sec) mean of the average response time to all questions in seconds

N_obs...All units represented in the report (i.e. interviewers, or questions OR also interviews, like in N_obs for interviewer)

Av_ResponseTime..... Average response time in the segment (excludes breaks)

N_questions..... Count of represented questions

counterMedianposition in the interview process for the median of interviews.

N_obs...how many times used in the interview process.

REFERENCES

- Biemer, P.P., P. Chen, and K. Wang. 2011. “Errors in the Recorded Number of Call Attempts and Their Effect on Nonresponse Adjustments Using Callback Models.” In *Paper Presented at the 58th World Statistics Congress of the International Statistical Institute*. Dublin, Ireland.
- Callegaro, Mario. (2013). *Paradata in Web Surveys*. 10.1002/9781118596869.ch11.
- Groves, R.M. 2006. “Nonresponse Rates and Nonresponse Bias in Household Surveys.” *Public Opinion Quarterly* 70 (5): 646–75.
- Groves, R.M., and E. Peytcheva. 2008. “The Impact of Nonresponse Rates on Nonresponse Bias: A Meta-Analysis.” *Public Opinion Quarterly* 72 (2): 167–89.
- West, Brady T. (2011). “Paradata in Survey Research.” *Survey Practice* 4 (4). <https://doi.org/10.29115/SP-2011-0018>.



THANK YOU

Ciprian Alexandru

alexciopro@gmail.com