**Datasets**

|  |  |  |  |
| --- | --- | --- | --- |
| **File name** | **Based on** | **Dimensions** | **Notes** |
| rss.csv | Reintegration Cases M&E RSS June 2023 (liste complète).xlsx [Mimosa]  and  RSS data cleaned to use for analysis V3 doublons corriges.xlsx [Kobo] | 1,385 x 215 | The RSS survey data, a merger done by the consultant. Contains all data, including scores, weights, etc.  All duration variables should be used with caution due to some inconsistencies in the raw data. |
| rss\_slim.csv | 1,385 x 34 | Same as previous, with only the composite scores and only our variables of interest  All duration variables should be used with caution due to some inconsistencies in the raw data. |
| rss\_slim.csv | 1,385 x 20 | Same as previous, except that data were processed for regression analysis (i.e., some categories were recoded to meet sample size criteria, outliers were replaced or removed, etc.), and that variables that could not be recoded were excluded.  All duration variables should be used with caution due to some inconsistencies in the raw data. |
| res.csv | RE\_Economic\_Survey\_clean for data analysis\_final.xlsx [kobo][[1]](#footnote-1) | 2,026 x 23 | The RES data. |
| res\_slim.csv | 1,917 x 23 | Same as previous, except that data were processed for regression analysis (i.e., some categories were recoded to meet sample size criteria, outliers were replaced or removed, etc.), and that missing data that could not be recoded were excluded. |
| res\_training.csv | Reintegration Economic\_clean 6.6.23\_doublons corrigesV2.xlsx | 1,949 x 32 | The RES data with the training variables from Kobo, a merger done by the client. This dataset is possibly corrupt and should be used with caution. In particular, the computed variables SupportDuration and TrainingDuration are likely corrupt. |
| res\_training\_slim.csv | 1,836 x 32 | Same as previous, except that data were processed for regression analysis (i.e., some categories were recoded to meet sample size criteria, outliers were replaced or removed, etc.), and that missing data that could not be recoded were excluded. This dataset is possibly corrupt and should be used with caution. In particular, the computed variables SupportDuration and TrainingDuration are likely corrupt. |

Except that… you have not recoded and cleaned rss yet!!!

1. This file was received 23 June 2023. It is exactly the same file as RE\_Economic\_Survey\_clean for data analysis.xlsx received on 21 June 2023, except that I the suffix “\_final” was added. [↑](#footnote-ref-1)