**PrePost\_NIH\_ListSorting.R\_README**

Robert Toms, robert.toms@utdallas.edu

README SECTIONS:

1. Purpose of Code
2. Researcher-facing Explanation of Code
3. Explanation of Variables
4. Methods Section Summary
5. References

PURPOSE OF CODE:

The purpose of this code is to match Pre & Post menarche timepoints for each subject with available NIH Toolbox score data for the List Sorting Working Memory (LSWM) Task.

RESEARCHER-FACING EXPLANATION OF CODE:

1. Derivative ABCD 6.0 menarche data were imported for youth (Youth\_PrePost\_Menarche\_9\_15\_25.csv) and parent (Parent\_PrePost\_Menarche\_9\_15\_25.csv) reports. These were used rather than raw data because they contain predetermined variables for when participants switched from premenarche to postmenarche (last\_pre\_session, first\_post\_session). Additionally, NIH Toolbox data (nc\_y\_nihtb.tsv) and interview date data (ab\_g\_dyn.tsv) were imported.
2. Clean and Prepare Data
   1. Useful columns were kept and renamed for human readability, if necessary.
   2. Data were converted to facilitate downstream transformations
   3. Sessions/Rows with no LSWM Task score data were removed
      1. n = 11466 timepoints removed
   4. Make Human-readable list vectors of timepoints with available LSWM score data and interview dates
      1. This is done for consolidation purposes, and so that the LSWM list vector can be algorithmically searched through to determine which scores best align with premenarche and postmenarche reports. It is saved to “has\_scores”.
         1. e.g. has\_scores = c(“0”, “4”, “6”) ; has scores for ses-00A, ses-04A, and ses-06A
      2. The interview dates list vector is created so we can easily calculate the number of days from pre/post-menarche reports to any other date of interest, like date of LSWM assessment
   5. Data are pivoted, formatted, and organized for readability and to facilitate future transformations
3. NIH Toolbox data and menarche data are merged
4. Find PreMenarche and PostMenarche Scores
   1. Wrote 2 functions:
      1. find\_premenarche\_score: finds the location index in the has\_scores list vector (1-4) with the session (with score data) that matches or most closely precedes a participant’s last premenarche report, saves it to PreScoreIndex
         1. Example
            1. last\_pre\_session = 3
            2. has\_scores = c(“0”, “4”, “6”)
            3. finds “0” value, 1st item in vector, PreScoreIndex = 1
      2. find\_postmenarche\_score: finds the location index in the has\_scores list vector (1-4) session (with score data) that matches or most closely follows a participant’s first postmenarche report, saves it to PostScoreIndex
         1. Example
            1. first\_post\_session = 4
            2. has\_scores = c(“0”, “4”, “6”)
            3. finds “4” value, 2nd item in vector, PostScanIndex = 2
   2. Applied the find\_premenarche\_score & find\_postmenarche\_score functions to youth and parent reports
      1. Subjects with no score data cause the functions to error (and we they don’t have useful data) so they were removed
         1. n = 20 removed
   3. Used the indexes (PreScoreIndex, PostScoreIndex) to identify the timepoints with score data most closely associated with premenarche and postmenarche reports
      1. Example
         1. has\_scores = c(“0”, “4”, “6”)
         2. PreScoreIndex = 1, 1st item is “0”, ses-00A has the best-fit premenarche scan, “0” is saved to PreScore\_session
         3. PostScoreIndex = 2, 2nd item is “4”, ses-04A has the best-fit postmenarche scan, “4” is saved to PostScore\_session
   4. Pulled the LSWM Scores associated with PreScore\_session (Premenarche score) and PostScore\_session (Postmenarche score), saved to PreScore and PostScore, respectively
5. Calculate Days between Data collection and/or days between menarche report and data collection
   1. Find Premenarche and Postmenarche LSWM task administration dates
      1. PreScore\_session and PostScore\_session were used to systematically save their respective administration dates to PreScore\_date and PostScore\_date
   2. Calculating Date Differences
      1. Number of days between best-fit premenarche and postmenarche scores, saved to “between\_scores”
      2. Number of days between last premenarche and first postmenarche reports, saved to “PreMen->PostMen”
      3. Number of days from best-fit premenarche Score to last premenarche report, saved to “PreMenScore->PreMenReport”
      4. Number of days from first postmenarche report to best-fit postmenarche score, saved to “PostMenReport->PostMenScore”
6. Clean and Reorder dataframes for CSV export
   1. has\_scores was converted to a string to retain readability while affording the ability to be written into a csv
   2. Necessary columns are kept and reordered
   3. Itemized menarche reports were renamed and merged in for reference
7. Write to CSV
   1. parent\_nih -> Parent\_PrePost\_NIHtbx\_Menarche\_9\_16\_25.csv
   2. youth\_nih -> Youth\_PrePost\_NIHtbx\_Menarche\_9\_16\_25.csv

EXPLANATION OF VARIABLES

|  |  |
| --- | --- |
| participant\_id | ABCD Participant Identifier |
| session\_id | ABCD Visit Identifier |
| nc\_y\_nihtb\_\_lswmt\_dtt | datetime of when list-sorting task was completed |
| nc\_y\_nihtb\_\_lswmt\_\_agecor\_score | NIH Toolbox, List Sorting Working Memory(LSWM) Task, age corrected score |
| ab\_g\_dyn\_\_visit\_dtt  interview\_date | Datetime for interview |
| session\_num | Numeric version of session\_id |
| has\_scores | List vector of session numbers with usable LSWM score data; e.g. c(“0”, “4”, “6”) = usable ses-00A, ses-04A, ses-06A scores |
| interview\_dates | List vector of each participant’s interview dates |
| 00A\_date / ses-00A\_score | Date or score for ses-00A LSWM task |
| 02A\_date / ses-02A\_score | Date or score for ses-02A LSWM task |
| 04A\_date / ses-04A\_score | Date or score for ses-04A LSWM task |
| 06A\_date / ses-06A\_score | Date or score for ses-06A LSWM task |
| last\_pre\_session | Numeric version of session with last premenarche report |
| first\_post\_session | Numeric version of session with first postmenarche report |
| PostMenarche\_at\_Baseline\_Y1N0 | Flags everyone who was postmenarche at first visit, Yes = 1 and No = 0 |
| PreMenarche\_at\_LastReport\_Y1N0 | Flags everyone who was premenarche through 6.0, Yes = 1 and No = 0 |
| Inconsistent\_Reporting\_Y1N0 | Flags everyone who reported being premenarche after reporting bein postmenarche  e.g. c(“0”, “0”, NA, “1”, **“0”,** “1”, “1”) = c(Pre, Pre, NA, Post, **Pre**, Post, Post) |
| find\_premenarche\_score | finds the index in the has\_scores list vector (1-4) with the session (with score data) that matches or most closely precedes a participant’s last premenarche report, gets saved to PreScoreIndex |
| find\_postmenarche\_score | finds the index in the has\_scores list vector (1-4) session (with score data) that matches or most closely follows a participant’s first postmenarche report, gets saved to PostScoreIndex |
| PreScoreIndex | Location in has\_scores list vector (1-4) of the value that matches or is the best-fit for last premenarche report |
| PostScoreIndex | Location in has\_scores list vector (1-4) of the value that matches or is the best-fit for last postmenarche report |
| PreScore\_session | PreMenarche LSWM session: Numeric version of the session with score data best-fit matched to last premenarche report |
| PostScore\_session | PostMenarche LSWM session: Numeric version of the session with score data best-fit matched to first postmenarche report |
| PreScore | PreMenarche Score: LSWM task score from the session best-fit matched to last premenarche report |
| PostScore | PostMenarche Score: LSWM task score from the session best-fit matched to first postmenarche report |
| ChangeScore | PostScore – PreScore; change in LSWM task performance from premenarche to postmenarche |
| PreScore\_date | PreMenarche Date: date of the LSWM task session best-fit matched to last premenarche report |
| PostScore\_date | PostMenarche Date: date of the LSWM task session best-fit matched to first postmenarche report |
| between\_scores | Number of days between PreMenarche LSWM task date (PreScore\_date) and PostMenarche LSWM task date (PostScore\_date) |
| PreMen->PostMen | Number of days between PreMenarche report and PostMenarche report |
| PreMenScore->PreMenReport | Number of days from PreMenarche LSWM task date (PreScore\_date) to PreMenarche report |
| PostMenReport->PostMenScore | Number of days from PostMenarche report to PostMenarche LSWM task date (PostScore\_date) |
| 00A\_menarche\_report | ses-00A menarche report:  Premenarche = 0, Postmenarche = 1 |
| 01A\_menarche\_report | ses-01A menarche report:  Premenarche = 0, Postmenarche = 1 |
| 02A\_menarche\_report | ses-02A menarche report:  Premenarche = 0, Postmenarche = 1 |
| 03A\_menarche\_report | ses-03A menarche report:  Premenarche = 0, Postmenarche = 1 |
| 04A\_menarche\_report | ses-04A menarche report:  Premenarche = 0, Postmenarche = 1 |
| 05A\_menarche\_report | ses-05A menarche report:  Premenarche = 0, Postmenarche = 1 |
| 06A\_menarche\_report | ses-06A menarche report:  Premenarche = 0, Postmenarche = 1 |

Methods Section Summary:

To facilitate analysis of changes in working memory ability from premenarche to postmenarche, we compared NIH Toolbox List Sorting Working Memory task age-corrected scores (nc\_y\_nihtb\_\_lswmt\_\_agecor\_score ; nc\_y\_nihtb.tsv) to predetermined last premenarche reports and first postmenarche reports to find matching sessions or sessions of best-fit, then calculated differences. To accomplish this for each participant, lists of sessions with LSWM score data were compiled. Sessions matching and/or preceding the participant’s last premenarche report were separated out, and the score associated with the latest of these sessions was deemed the “best-fit” premenarche score. Similarly, sessions matching and/or following the participant’s first postmenarche report were separated out, and the score associated with the first of these sessions was deemed the “best-fit” postmenarche score. The premenarche LSWM score was then subtracted from the postmenarche LSWM score to get a Change Score.

References: