Codebook | Time Study Sleep Report

TIME study website

6/30/23: This codebook is currently being improved as more researchers use the TIME data. If information in the codebook is not clear, or if you feel information is missing, including key variables that might simplify your data analysis, you are encouraged to use the Google Docs comment feature to insert your question(s) or comment(s) directly in this document. Alternatively, you can use the project forum. Thank you for your help! - The TIME Study Team.

Table of Contents

About the TIME Study

The TIME study used real-time mobile technologies to collect intensive longitudinal data examining differences in micro-temporal processes underlying the adoption and maintenance of physical activity, low sedentary time, and sufficient sleep duration.

To learn about TIME Study goals, methods, and data, please begin by reviewing the TIME Study website.

About this Codebook

This codebook describes the data format of the combined sleep report CSV file of TIME study. The sleep report consists of prospective and retrospective self-reported sleep data, accelerometry-detected sleep data, and aggregated sleep metrics for each participant, by date. These are meta-data obtained from processing the raw sensor and logs data saved by the apps that were running on the smartphone and smartwatch. These data include **prospective and retrospective sleep/wake time, personal daybreak time and personal daybreak days in study, sleep metrics (total sleep time, sleep period time window, wake after sleep onset), etc.**

The first page has a glossary that helps readers direct to the page for particular columns in the sleep report. Each column consists of one summary section and one metadata column section. The summary section has the general description about data source and nature of the data. The metadata column section includes description, format, and value options for each column in the intermediate data file.

This is one of several codebooks for the TIME Study. Many projects will require combining data from multiple codebooks. The TIME Study website has links to all codebooks.

Researchers using these data should ask questions or report problems on the project forum.

About Sleep Report File

The sleep report file is a user-friendly CSV file obtained from processing the raw sensor and logs data saved by the apps that were running on the smartphone and smartwatch. It is intended to collect sleep-related information for all participants from raw data and be used to create personal daybreak time based on participants' wake-sleep schedule. Besides, we computed sleep-related metrics by using both self-reported sleep information and accelerometer-detected sleep data for advanced sleep analysis.

About Data Generation

The data extraction and processing scripts that generate these meta-data from raw app data are written in Python and are computed using Northeastern University's Discovery Cluster. The Python code and scripts required to run the code on the Discovery Cluster can be found in the TIME Study Bitbucket repository. Researchers using these data should ask questions or report problems on the project forum.

Abbreviations and Terminology

Please review the Table of Abbreviations and Special Terminology.

Description of Variables

Summary

This file contains a single line per participant and personal daybreak date, summarizing sleep/wake time, personal daybreak time, sleep-related metrics, etc.

Definition

A personal daybreak day refers to a day defined by personal daybreak. A personal daybreak day starts at last_breaktime and ends at current_daybreak. The current_breaktime then becomes the last_breaktime of the next personal daybreak. The daybreak time is determined by the combination of self-reported (prospective and retrospective) and accelerometer-detected sleep/wake time. Please find the combination rule in the paragraph of current_daybreak.

Metadata Columns

Col: participant_id_numeric

Descriptions: The unique numeric identifier of the participant. It may also be called record_id or REDcap id, which can be used to link with the REDcap data.

Format: Numbers, four-digit numerical numbers (e.g., 9001).

Col: participant_id_text

Descriptions: The unique textual identifier of the participant.

Format: String, sometimes the string will appear to be an email, but it is not an actual email (e.g., “crystaloverdueunfasten”).

Notes: If user_name is not correct, "unknown" or just missing.

Col: day_num

Descriptions: personal daybreak day count. See the definition of personal daybreak day in current_daybreak.

Format: Integer, start date is 0

Col: retrospective_sleep_time

Descriptions: Retrospective self-reported sleep time. Everyday one retrospective sleep survey was scheduled 30 min after plan-to-wakeup time. The survey would be prompted every hour if not answered. The survey included one sleep time question (“What time did you go to sleep last night?”) and one wake time question "What time did you wake up today?". The retrospective questions were asked to confirm the prior night’s actual asleep-awake times.

Format: String, “YYYY-MM-DD HH:MM:SS” (e.g., 2020-03-01 22:00:00).

Col: retrospective_wake_time

Descriptions: Retrospective self-reported wake time. Everyday one retrospective sleep survey was scheduled 30 min after plan-to-wakeup time. The survey would be prompted every hour if not answered. The survey included one sleep time question (“What time did you go to sleep last night?”) and one wake time question "What time did you wake up today?". The retrospective questions were asked to confirm the prior night’s actual asleep-awake times.

Format: String, “YYYY-MM-DD HH:MM:SS” (e.g., 2020-03-02 08:00:00).

Col: prospective_ema_sleep_time

Descriptions: Prospective self-reported sleep time. Prospective sleep surveys were scheduled several times a day, including 30 min after plan-to-wakeup time, reprompt every hour if not answered, re-asked 10 hours after last answered survey, re-asked at the end-of-day survey. The survey included one sleep time question (“What time do you plan to go to sleep next?”) and one wake time question "What time do you plan to wake up next?". The prospective questions were asked to adapt the EMA prompting schedule for each day of the study (i.e., first EMA survey of the day is prompted 30 min after plan-to-wake-up time, and the end-of-day survey is prompted 2 hrs before planned bedtime). In addition, participants also had the option to proactively use the app as needed to change the prospective times (see prospective_manual_sleep_time). Here, prospective_ema_sleep_time is the last self-reported planned bedtime of each day via prospective sleep surveys.

Format: String, “YYYY-MM-DD HH:MM:SS” (e.g., 2020-03-01 22:00:00).

Col: prospective_ema_wake_time

Descriptions: Prospective self-reported wake time. Prospective sleep surveys were scheduled several times a day, including 30 min after plan-to-wakeup time, reprompt every hour if not answered, re-asked 10 hours after last answered survey, re-asked at the end-of-day survey. The survey included one sleep time question (“What time do you plan to go to sleep next?”) and one wake time question "What time do you plan to wake up next?". The prospective questions were asked to adapt the EMA prompting schedule for each day of the study (i.e., first EMA survey of the day is prompted 30 min after plan-to-wake-up time, and the end-of-day survey is prompted 2 hrs before planned bedtime). In addition, participants also had the option to proactively use the app as needed to change the prospective times (see prospective_manual_wake_time). Here, prospective_ema_wake_time is the last self-reported plan-to-wake-up time of each day via prospective sleep surveys.

Format: String, “YYYY-MM-DD HH:MM:SS” (e.g., 2020-03-02 08:00:00).

Col: prospective_manual_sleep_time

Descriptions: Prospective manually-changed sleep time. The TIME smartphone app allows participants to proactively change the prospective times as needed (see prospective_ema_sleep_time). Here, the value is the last manually-changed planned bedtime of each day on the TIME app. If there is no manual change after the end-of-day survey, this value would be nan.

Format: String, “YYYY-MM-DD HH:MM:SS” (e.g., 2020-03-01 22:00:00).

Col: prospective_manual_wake_time

Descriptions: Prospective manually-changed wake time. The TIME smartphone app allows participants to proactively change the prospective times as needed (see prospective_ema_wake_time). Here, the value is the last manually-changed plan-to-wake-up time of each day on the TIME app. If there is no manual change after the end-of-day survey, this value would be nan.

Format: String, "YYYY-MM-DD HH:MM:SS" (e.g., 2020-03-02 08:00:00).

Col: prospective_sleep_time

Descriptions: Prospective sleep time. The value is prospective_ema_sleep_time or prospective_manual_sleep_time if it is not nan.

Format: String, "YYYY-MM-DD HH:MM:SS" (e.g., 2020-03-01 22:00:00).

Col: prospective_wake_time

Descriptions: Prospective wake time. The value is prospective_ema_wake_time or prospective_manual_wake_time if it is not nan.

Format: String, "YYYY-MM-DD HH:MM:SS" (e.g., 2020-03-02 08:00:00).

Col: swan_sleep_time

Descriptions: SWaN-detected sleep time (using watch accelerometer data only). The value is the sleep onset time of each day based on SWaN predictions. Here, we use minute-level SWaN predictions to detect the major sleep of each day. The major sleep of a day is defined as the longest consecutive sleep of the day. Sleep predictions that are within 30 minutes of the major sleep period are included in order to concatenate sleep periods splitted by short-term mid-night wake-ups. Note that the accuracy of this value is subject to many factors, including accelerometer data availability, participants' watch nonwear behavior, SWaN algorithm accuracy, and major sleep detection algorithm accuracy, and therefore needs to be used in caution.

Format: String, "YYYY-MM-DD HH:MM:SS" (e.g., 2020-03-01 22:00:00).

Col: swan_wake_time

Descriptions: SWaN-detected wake time (using watch accelerometer data only). The value is the wake-up time of each day based on SWaN predictions. Here, we use minute-level SWaN predictions to detect the major sleep of each day. The major sleep of a day is defined as the longest consecutive sleep of the day. Sleep predictions that are within 30 minutes of the major sleep period are included in order to concatenate sleep periods splitted by short-term mid-night wake-ups. Note that the accuracy of this value is subject to many factors, including accelerometer data availability, participants' watch nonwear behavior, SWaN algorithm accuracy, and major sleep detection algorithm accuracy, and therefore needs to be used in caution.

Format: String, "YYYY-MM-DD HH:MM:SS" (e.g., 2020-03-02 08:00:00).

Col: last_breaktime

Descriptions: last breaktime of the personal daybreak day. This value is computed as the middle point of the major sleep of each day. The major sleep is defined by the retrospective sleep/wake time if available

(if retrospective is not available, using prospective sleep/wake time; if prospective is not available, impute the missing value by the time that is 24 hrs after the last available breaktime).

Format: String, “YYYY-MM-DD HH:MM:SS” (e.g., 2020-03-01 00:00:00).

Col: current_breaktime

Descriptions: current breaktime of the personal daybreak day. This value is computed as the middle point of the major sleep of each day. The major sleep is defined by the retrospective sleep/wake time if available (if retrospective is not available, using prospective sleep/wake time; if prospective is not available, impute the missing value by the time that is 24 hrs after the last available breaktime).

Format: String, “YYYY-MM-DD HH:MM:SS” (e.g., 2020-03-01 00:00:00).

Col: num_hours

Descriptions: number of hours between last_breaktime and current_breaktime.

Format: Float (e.g., 23.5)

Col: waking_date

Descriptions: the date that has the majority of waking time of each personal daybreak day.

Format: String, “YYYY-MM-DD” (e.g., when “last_breaktime” is “2021-06-22 02:15:00” and “current_breaktime” is “2021-06-23 04:00:00”, the “waking_date” is “2021-06-22”)

Col: current_sleep_time

Descriptions: the sleep time of the current personal daybreak day. The major sleep is defined by the retrospective sleep/wake time if available. If retrospective is not available, using prospective sleep/wake time; if prospective is not available, impute the missing value by the time that is 24 hrs after the last available breaktime.

Format: String, “YYYY-MM-DD HH:MM:SS” (e.g., 2020-03-01 22:00:00).

Col: current_wake_time

Descriptions: the wake time of the current personal daybreak day. The major sleep is defined by the retrospective sleep/wake time if available. If retrospective is not available, using prospective sleep/wake time; if prospective is not available, impute the missing value by the time that is 24 hrs after the last available breaktime.

Format: String, “YYYY-MM-DD HH:MM:SS” (e.g., 2020-03-02 08:00:00).

Col: current_sleep_min

Descriptions: duration of the self-reported sleep. Time difference between current_sleep_time and current_wake_time.

Format: Float

Col: spt

Descriptions: duration of sleep period time window (min). This is the main sleep window extending from accelerometer-detected sleep onset to accelerometer-detected waking time during the guider window, which is defined by current_sleep_time and current_wake_time. This window is intended to be used as a "guider window" similar to the method described in the R package GGIR uses with the argument guider = sleep log.

Format: Float (e.g., 237 min)

Col: waso

Descriptions: amount of wake time after sleep onset. This is the total minutes of wake that occur within the SPT window (i.e., between sleep onset and waking up time).

Format: Float (e.g., 82 min)

Col: wcaso

Descriptions: number of wake occurrences after sleep onset. This is the total number of wake occurrences that occur within the SPT window (i.e., between sleep onset and waking up time).

Format: Integer (e.g., 2 times)

Col: tst

Descriptions: total sleep time. Calculated as: SPT duration – WASO = TST.

Format: Float (e.g., 237-82 = 155 min)

Col: se

Descriptions: sleep efficiency. This is the proportion of time in bed that is spent sleeping. Approximated as: TST / SPT.

Format: Float (e.g., 155/237 = 65.4%)

Col: swan_sleep_onset

Descriptions: the onset time of SPT window.

Format: String, "YYYY-MM-DD HH:MM:SS" (e.g., 2020-03-01 22:00:00).

Col: swan_sleep_end

Descriptions: the end time of SPT window.

Format: String, "YYYY-MM-DD HH:MM:SS" (e.g., 2020-03-02 08:00:00).

Col: swan_missing_min_during_retro_sleep

Descriptions: number of minutes of accelerometer data being missing during the retrospective self-reported "guider window" sleep time (i.e., "current_sleep_time", "current_wake_time").

The value should be greater than or equal to zero, while the negative value indicates special cases:

- <-1>: retrospective wake/sleep time is nan
- <-2>: swan info of wake day missing
- <-3>: swan info of sleep day missing
- <-4>: swan info of wake and sleep day missing

Recommendation of usage: filter the dataset and exclude days with all negative values, as they indicate retrospective sleep information or accelerometry data being missing.

Format: Integer

Col: swan_nonwear_min_during_retro_sleep

Descriptions: number of minutes of being detected as watch-nonwear by SWaN during the retrospective self-reported "guider window" sleep time (i.e., "current_sleep_time", "current_wake_time").

The value should be greater than or equal to zero, while the negative value indicates special cases:

- <-1>: retrospective wake/sleep time is nan
- <-2>: swan info of wake day missing
- <-3>: swan info of sleep day missing
- <-4>: swan info of wake and sleep day missing

Recommendation of usage: filter the dataset and exclude days with all negative values, as they indicate retrospective sleep information or accelerometry data being missing.

Format: Integer

Col: swan_missing_percentage_during_retro_sleep

Descriptions: percentage of accelerometer data being missing during the retrospective self-reported "guider window" sleep time (i.e., "current_sleep_time", "current_wake_time"). Calculated as: swan_missing_min_during_retro_sleep divided by current_sleep_min

Format: Float

Col: swan_nonwear_percentage_during_retro_sleep

Descriptions: percentage of being detected as watch-nonwear by SWaN during the retrospective self-reported "guider window" sleep time (i.e., "current_sleep_time", "current_wake_time"). Calculated as: swan_nonwear_min_during_retro_sleep divided by current_sleep_min

Format: Float

Col: swan_missing_nonwear_percentage_during_retro_sleep

Descriptions: percentage of accelerometer data being missing + being detected as watch-nonwear by SWaN during the retrospective self-reported "guider window" sleep time (i.e., "current_sleep_time", "current_wake_time"). Calculated as: swan_missing_percentage_during_retro_sleep + swan_nonwear_percentage_during_retro_sleep

Format: Float

Recommendation of usage: filter the dataset and exclude days with percentage threshold according to different analysis needs.