**Dataset Description**

The dataset comprises about 500 multi-day recordings of wrist-worn accelerometer data annotated with two event types: *onset*, the beginning of sleep, and *wakeup*, the end of sleep. Your task is to detect the occurrence of these two events in the accelerometer series.

While sleep logbooks remain the gold-standard, when working with accelerometer data we refer to sleep as the longest single period of inactivity while the watch is being worn. For this data, we have guided raters with several concrete instructions:

* A single sleep period must be at least 30 minutes in length
* A single sleep period can be interrupted by bouts of activity that do not exceed 30 consecutive minutes
* No sleep windows can be detected unless the watch is deemed to be worn for the duration (elaborated upon, below)
* The longest sleep window during the night is the only one which is recorded
* If no valid sleep window is identifiable, neither an onset nor a wakeup event is recorded for that night.
* Sleep events do not need to straddle the day-line, and therefore there is no hard rule defining how many may occur within a given period. However, no more than one window should be assigned per night. For example, it is valid for an individual to have a sleep window from 01h00–06h00 and 19h00–23h30 in the same calendar day, though assigned to consecutive nights
* There are roughly as many nights recorded for a series as there are 24-hour periods in that series.

Though each series is a continuous recording, there may be periods in the series when the accelerometer device was removed. These period are determined as those where suspiciously little variation in the accelerometer signals occur over an extended period of time, which is unrealistic for typical human participants. Events are not annotated for these periods, and you should attempt to refrain from making event predictions during these periods: an event prediction will be scored as false positive.

Each data series represents this continuous (multi-day/event) recording for a unique experimental subject.

Note that this is a [**Code Competition**](https://www.kaggle.com/competitions/child-mind-institute-detect-sleep-states/overview/code-requirements), in which the actual test set is hidden. In this public version, we give some sample data in the correct format to help you author your solutions. The full test set contains about 200 series.

**Files and Field Descriptions**

* **train\_series.parquet** - Series to be used as training data. Each series is a continuous recording of accelerometer data for a single subject spanning many days.
  + series\_id - Unique identifier for each accelerometer series.
  + step - An integer timestep for each observation within a series.
  + timestamp - A corresponding datetime with ISO 8601 format %Y-%m-%dT%H:%M:%S%z.
  + anglez - As calculated and described by the [GGIR package](https://cran.r-project.org/web/packages/GGIR/vignettes/GGIR.html#4_Inspecting_the_results), z-angle is a metric derived from individual accelerometer components that is commonly used in sleep detection, and refers to the angle of the arm relative to the vertical axis of the body
  + enmo - As calculated and described by the [GGIR package](https://cran.r-project.org/web/packages/GGIR/vignettes/GGIR.html#4_Inspecting_the_results), ENMO is the Euclidean Norm Minus One of all accelerometer signals, with negative values rounded to zero. While no standard measure of acceleration exists in this space, this is one of the several commonly computed features
* **test\_series.parquet** - Series to be used as the test data, containing the same fields as above. You will predict event occurrences for series in this file.
* **train\_events.csv** - Sleep logs for series in the training set recording onset and wake events.
  + series\_id - Unique identifier for each series of accelerometer data in train\_series.parquet.
  + night - An enumeration of potential onset / wakeup event pairs. At most one pair of events can occur for each night.
  + event - The type of event, whether onset or wakeup.
  + step and timestamp - The recorded time of occurence of the event in the accelerometer series.
* **sample\_submission.csv** - A sample submission file in the correct format. See the [**Evaluation**](https://www.kaggle.com/competitions/child-mind-institute-detect-sleep-states/overview/evaluation) page for more details.

kaggle competitions download -c child-mind-institute-detect-sleep-states