

## Cleaning documentation

Date : 5/5/2022

- I saved a new table `activity_daily_important_cols_05_05_2022_V01` that does not contain 3 columns (TrackerDistance, LoggedActivitiesDistance, SedentaryActiveDistance) that decreased the number of columns from 15 to 12 compared to the original table `activity_daily`.
- I Updated `activity_daily_important_cols_05_05_2022_V01` table by rounding numbers to 2 decimals in the following 4 columns `TotalDistance`, `VeryActiveDistance`, `ModeratelyActiveDistance`, `LightActiveDistance`.

Date : 6/5/2022

- I cleaned column names (convert it to snake case) `activity_daily_cleaned__06_05_2022_V01.01.csv` , `calories_daily_cleaned_06_05_2022_v01.csv`

Date : 7/5/2022

- I Removed Fat columns, 97% of entries are NA that decreased number of columns from 8 to 7 columns `weight_daily_cleaned_07_05_2022_v01.csv`
- I cleaned column names (convert it to snake case) `weight_daily_cleaned_07_05_2022_v01.csv`
- I removed 3 duplicate rows that decreased number of rows from 413 to 410 `sleep_daily_cleaned_07_05_2022_v01.csv`
- I added 2 columns `total_hours_asleep` , `total_hours_in_bed` (number of minutes / 60) that increased number of column from 7 to 9 columns `sleep_daily_cleaned_07_05_2022_v01.csv`

Date : 8/5/2022

- I changed column `sleep_day` from Datetime to Date `sleep_daily_cleaned_08_05_2022_v02.csv`
- I joined part of `activity_daily_important_cols_05_05_2022_V01`, and all columns `sleep_daily_cleaned_08_05_2022_v02.csv` new table is `activity_sleep_daily_joined_08_05_2022_v01.csv`
- I cleaned `activity_sleep_daily_joined_08_05_2022_v01.csv` and selected rows that has no NULL values new table is `activity_sleep_daily_joined_not_null_08_05_2022_v02.csv`
- I changed column date from Datetime to Date `weight_daily_cleaned_08_05_2022_v02.csv`
- I joined the new table `activity_sleep_daily_joined_08_05_2022_v02.csv` and `weight_daily_cleaned_08_05_2022_v02.csv` the new table is `activity_sleep_weight_daily_joined_08_05_2022_v02.csv`