Questions

1. Where is your data is located?
   public data in Kaggle
   1.1 MÖBIUS
   FitBit Fitness Tracker Data
   Pattern recognition with tracker data: : Improve Your Overall Health
   1.2 BEKBOLAT KURALBAYEV · UPDATED 3 YEARS AGO
   Exported data from Xiaomi Mi Band fitness tracker
2. How is the data organized?
   several files in csv format
3. Are there issues with bias or credibility in this data?
4. Does your data ROCCC? Reliable, original source, comprehensive, current, cited

   Reliable-can check sources for bias, both complete and accurate

   Original source - the first source is given by the second party, the second source is given by the first party. So the second source is more reliable.

   Comprehensive - both

   Current - not current, because before 2020 people didn't worry about their health and the view on activity by people has changed after COVID. So the best years for estimation are 2021-2022 then the restrictions connected by COVID became limited.

5. How you addressing licensing, privacy, security, and accessibility?

   fitbit
   CC0: Public Domain and Creative Commons Attribution 4.0 International on https://zenodo.org/records/53894#.YMoUpnVKiP9
   The person who associated a work with this deed has dedicated the work to the public. You can copy, modify, distribute and perform the work, even for commercial purposes, all without asking permission.

   xiaomi
   ATTRIBUTION 4.0 INTERNATIONAL
   You are free to share — copy and redistribute the material in any medium or format for any purpose, even commercially. Adapt — remix, transform, and build upon the material for any purpose, even commercially.
   Attribution — You must give appropriate credit , ***provide a link to the license, and indicate if changes were made .*** *You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.*

6. How did you verify the data's integrity?
   The information about physical activity, heart rate, and sleep monitoring is connected to the Id of the person and the date.

7. How does it help you answer your question?
   Data about personal tracker data will help to analyze the lifestyle of people who use health trackers and explore their habits. Knowledge about the habits of health tacker users will guide the company's marketing strategy.

8. Are there any problems with data?
   8.1 We use only public data from other non-Bellabet users so the data is obtained from a second party source and we can't guarantee 100% of reliability.

   8.2 The main problem with the dataset is that it's not current.
   If I was a data analyst I would recommend to add data up to 2019 year, add some other datasets: Health App Revenue and Usage Statistics (2024), Number of data trackers on period and women health apps as of June 2022, by operative system

   Key tasks
   Ensure that you have all of the data you need for your analysis and that you have credible useful data

   - ☐ collect data and store it appropriately
   - ☐ identify how it's organized

   2 directories with csv files (Fitabase Data 4.12.16-5.12.16 and
   Fitabase Data 3.12.16-4.11.16
   total id 35, 33 days of statistics
   Daily Activity (Id, ActivityDate, TotalSteps, TotalDistance, TrackerDistance, LoggedActivitiesDistance, VeryActiveDistance, ModeratlyActiveDistance, LightActiveDistance, SedentaryActiveDistance, VeryActiveMinutes, FairlyActiveMinutes, LightlyActiveMinutes, SedetaryActiveMinutes, Calories)
   Heart Rate (Id, time, value)
   Hourly Calories (Id, ActivityHour, Calories)
   Hourly Intensities (Id, ActivityHour, TotalIntensity, AverageIntensity)
   Hourly Steps (Id, ActivityHour, StepTotal)
   Minute Steps (Id, ActivityMinute, Steps)
   Minute Sleep (Id, date, value, logId)
   Minute METs (Id, ActivityMinute, METs)
   Minute Intensity Narrow (Id, ActivityMinute, Intensity)
   Weight lod info (Id, Date, WeightKg, WeihgtPounds, BMI, IsManualReport, LogId)

   Fitabase Data 4.12.16-5.12.16
   total id 33, 31 days of statistics
   Daily Activity (Id, ActivityDate, TotalSteps, TotalDistance, TrackerDistance, LoggedActivitiesDistance, VeryActiveDistance, ModeratlyActiveDistance, LightActiveDistance, SedentaryActiveDistance, VeryActiveMinutes, FairlyActiveMinutes, LightlyActiveMinutes, SedetaryActiveMinutes, Calories)

Heart Rate (Id, time, value)
Hourly Calories (Id, ActivityHour, Calories)
Hourly Intensities (Id, ActivityHour, TotalIntensity, AverageIntensity)
Hourly Steps (Id, ActivityHour, StepTotal)
Minute Steps Narrow (Id, ActivityMinute, Steps) and Wide
Minute Sleep (Id, date, value, logId)
Minute METs (Id, ActivityMinute, METs)
Minute Intensity Narrow (Id, ActivityMinute, Intensity) and WIde
Minute Calories Narrow (Id, ActivityMinute, Calories) and Wide
Weight lod info (Id, Date, WeightKg, WeihgtPounds, BMI, IsManualReport, LogId)
Daily Calories (Id, Activity, Calories)
Daily Intensities (Id, ActivityDay, SedentaryMinutes, VeryActiveDistance, ModeratlyActiveDistance, LightActiveDistance, SedentaryActiveDistance, VeryActiveMinutes, FairlyActiveMinutes, LightlyActiveMinutes, SedetaryActiveMinutes)
Daily Steps (Id, ActivityDay, StepTotal)
Sleep Day (Id, SleepDay, TotalSleepRecords, TotalMinutesAsleep, TotalTimeInBed)

I add Xiaomi Me Band dataset
total id 1, 385 days of statistics
Activity (date, lastSyncTime, steps, distance, runDistance, calories)
Sleep (date, lastSyncTime, deepSleepTime, shallowSleepTime,WakeTime, start, stop)
Heartrate (date, lastSyncTime, heartRate, Timestamp)
Heartrate Auto - every 2 minutes (date, time, heartRate)
Activity Stage (date, start, stop, distance, calories,steps)
Activity minute (date, time, steps) every minute

Common information in 3 datasets

- steps (sum and every minute/hour)
- distance (sum and every minute/hour)
- calories (sum by day and every minute/hour)
- sleep total minute)
- heart rate (every 1-2 minute)

- [ ] soft and filter data
- [ ] determine the credibility of the data

Deliverable
A description of all data sources used

This Kaggle data set contains personal fitness tracker from thirty fitbit users. Thirty eligible Fitbit users consented to the submission of personal tracker

data, including minute-level output for physical activity, heart rate, and sleep monitoring. It includes information about daily activity, steps, and heart rate that can be used to explore users' habits.

Personal data collected by BEKBOLAT KURALBAYEV during 1 year.