** Capstone Project for Google Data Analytics Professional Certificate

Analysis of Fitness Data for..

bellabeat

By imagining that we are a part of the company Bellabeat, in this project, we will be performing some real-world tasks of a Data Analyst.

- Sreyesh Achary

About

- Bellabeat is a high-tech manufacturer of health-focused products for women.
- Though a successful small company, they have the potential to become a larger player in the global smart device market.
- Urška Sršen, cofounder and Chief Creative Officer of Bellabeat, believes that analyzing smart device fitness data could help unlock growth opportunities for the company.
- The products of the company are:
    - Leaf - a classic wellness tracker
    - Time - a wellness watch to track user activity, sleep, and stress
    - Spring - a water bottle that tracks daily water intake
    - Bellabeat membership - a subscription-based membership program for users, which gives users 24/7 access to fully personalized guidance on nutrition, activity, sleep, health and beauty, and mindfulness based on their lifestyle and goals
    - Bellabeat app that the devices - Leaf, Time & Spring connect to, allowing users to track the respective metrics and provide insights to the users regarding their health.

Objective

- Presenting analysis on the smart device users' daily habits and providing recommendations for a marketing strategy, to the Executive Team of Bellabeat.
- We will be analysing a public data to gain insight into how consumers use non-Bellabeat smart devices.

Dataset

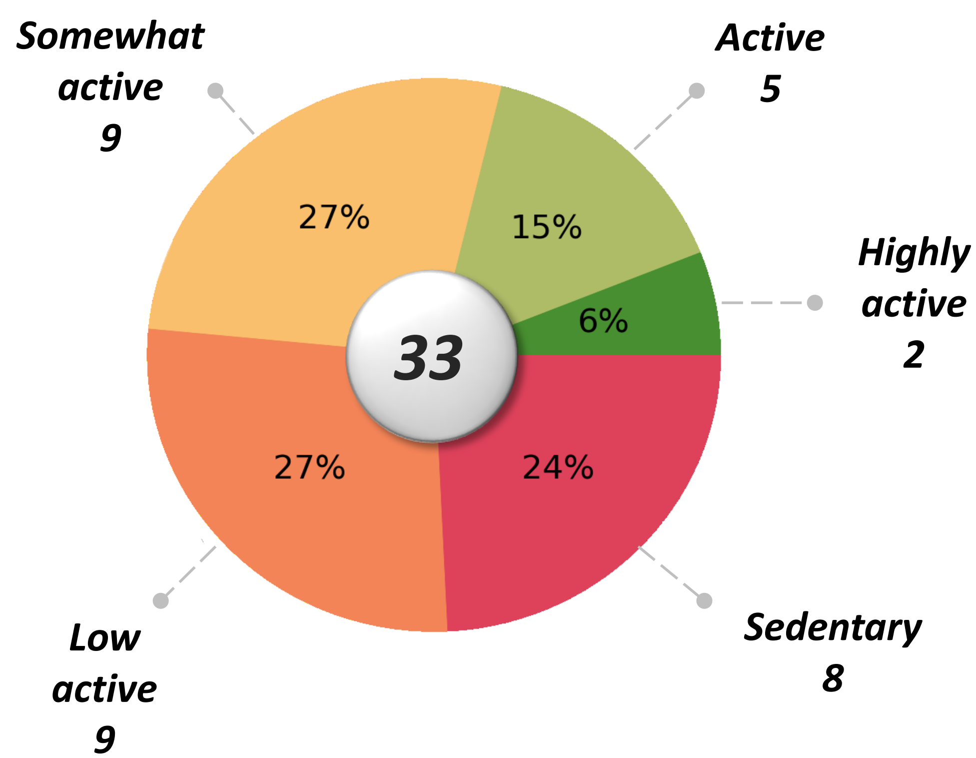
- We used the FitBit Fitness TrackerData by MÖBIUS with data of 33 users from 12-Apr-16 to 12-May-16.
- This dataset group contains 12 files; from these, we used the following 3 datasets for our analysis: Daily activity, Hourly steps and Sleep day

Resources

1. Kaggle for datasets
2. Python for cleaning, analysing and visualising
3. Obsidian for notes
4. MS Excel for rudimentary checks and validation

Let's understand the users who submitted the tracker data..

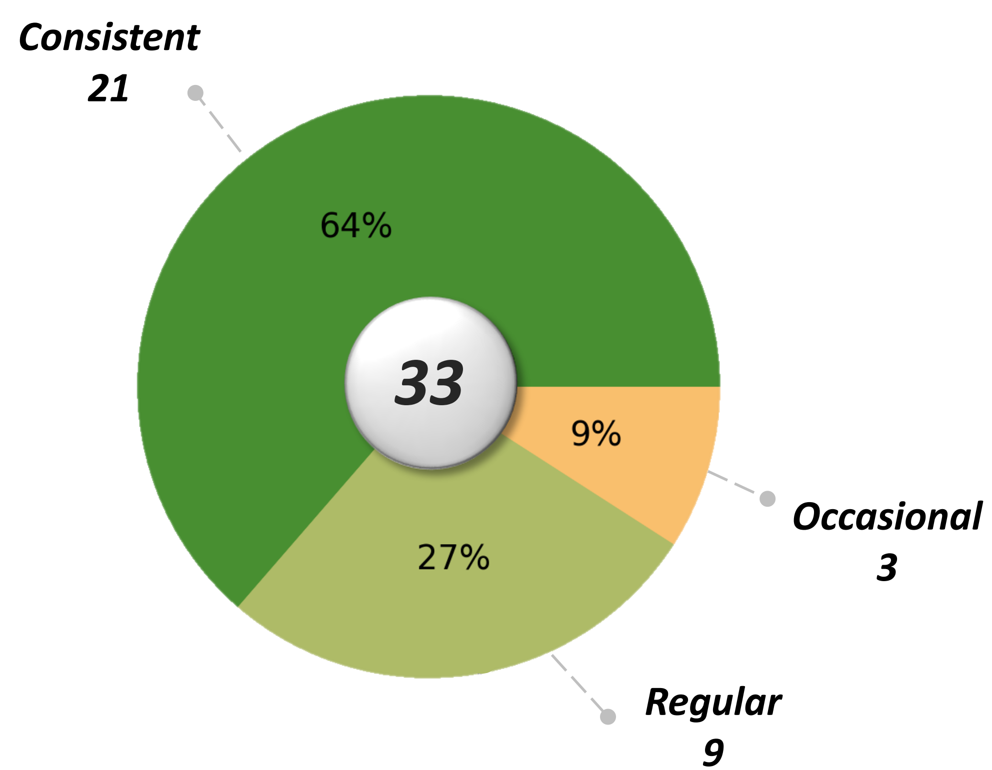
** User Classification by average steps per day



Where average steps per day are :

- more than 12,500 ------------------***Highly active***
- between 10,000 & <12,500 ----***Active***
- between 7,500 & <10,000 ------***Somewhat Active***
- between 5,000 & <7,500 -------***Low Active***
- less than 5,000 -------------------***Sedentary***

User Classification by tracking frequency **

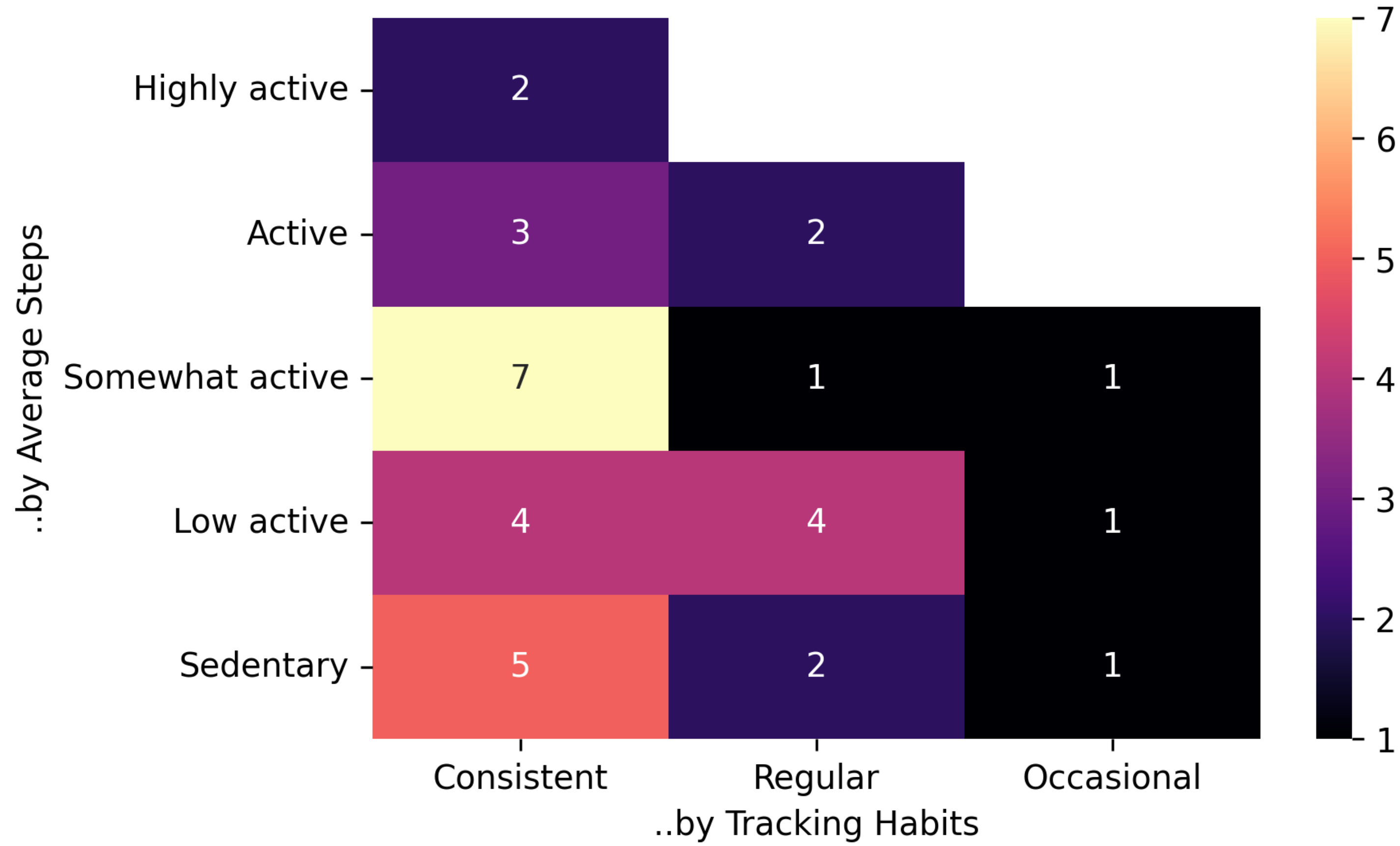


Where activity count in days are :

- equal to 31 ------------------***Consistent***
- between 20 & <31 --------***Regular***
- less than 20 -----------------***Occasional***

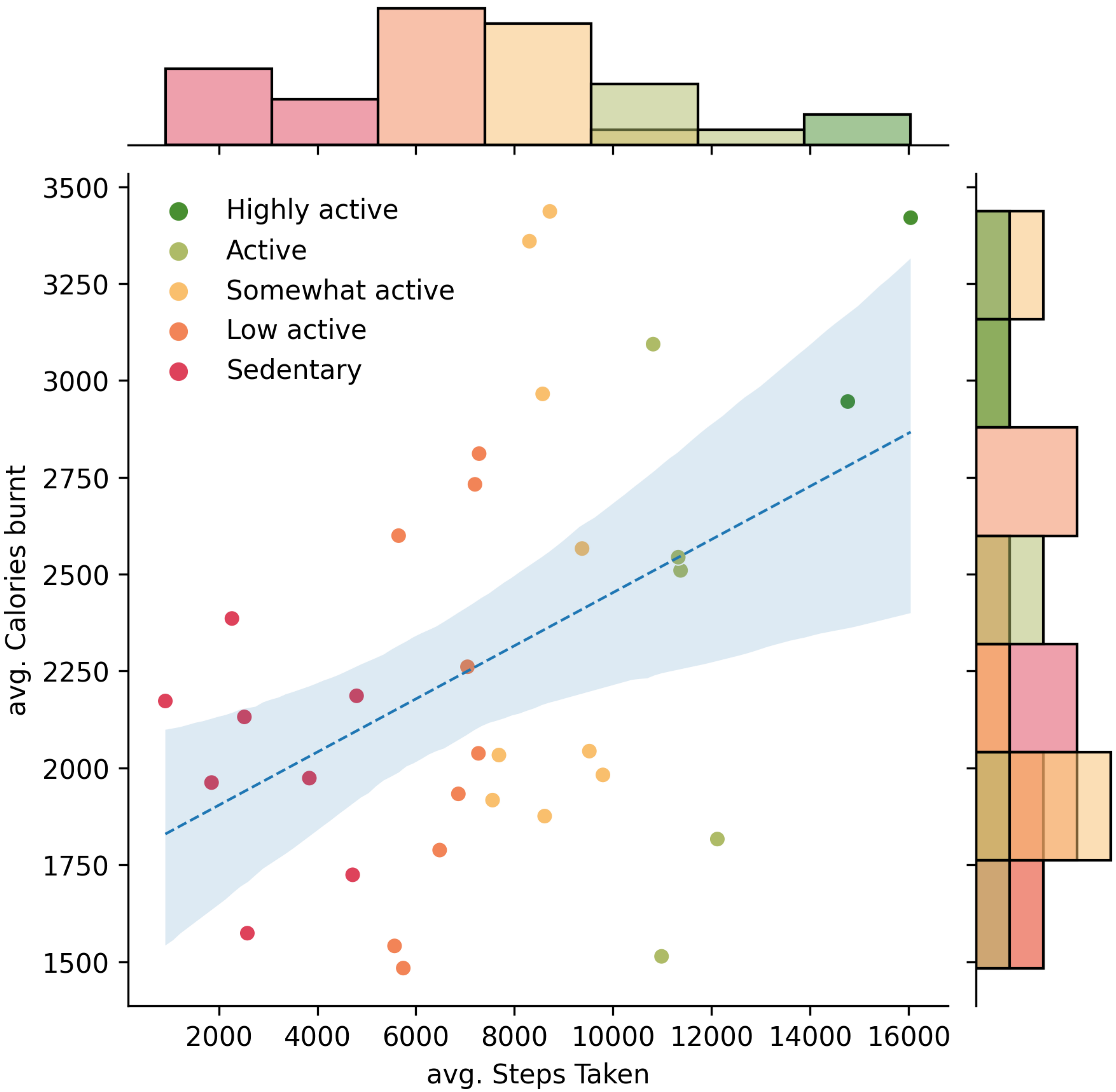
Let's understand the users who submitted the tracker data..

** Summary



- ✓ Out of the 33 participants that submitted the data, 8 are sedentary – quite high (relatively)
- ✓ Majority of the users are somewhat active and track consistently
- ✓ No occasional tracker among Highly active and active users
- ✓ More than half of the sedentary users tracked their activity every single day, quite odd.
- ✓ The categorisation of users based on the average number of steps taken is based on the research established by Catrine Tudor-Locke of the University of North Carolina - Charlotte.

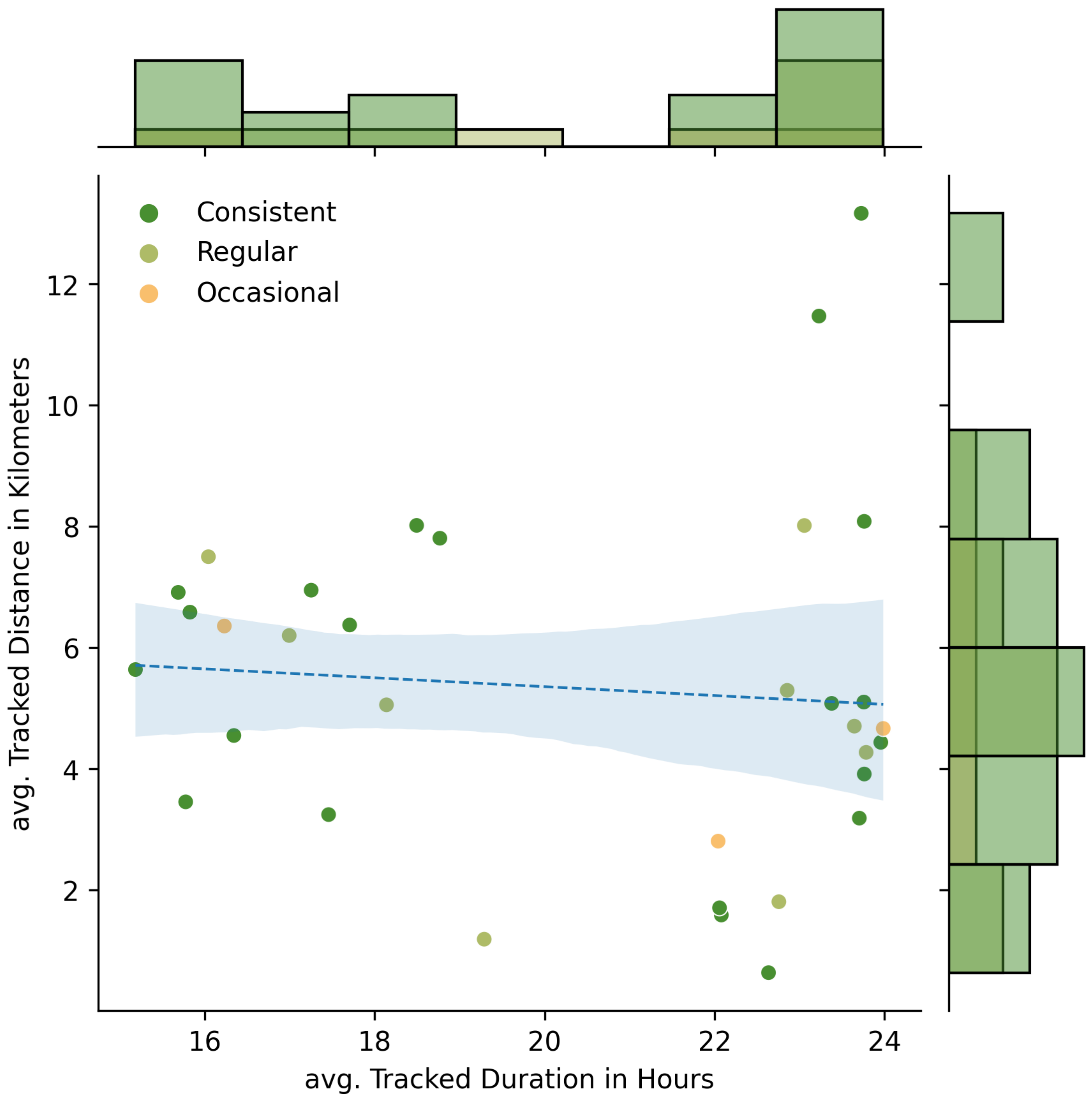
..now, we visualise the user traits – daily steps & calories burnt



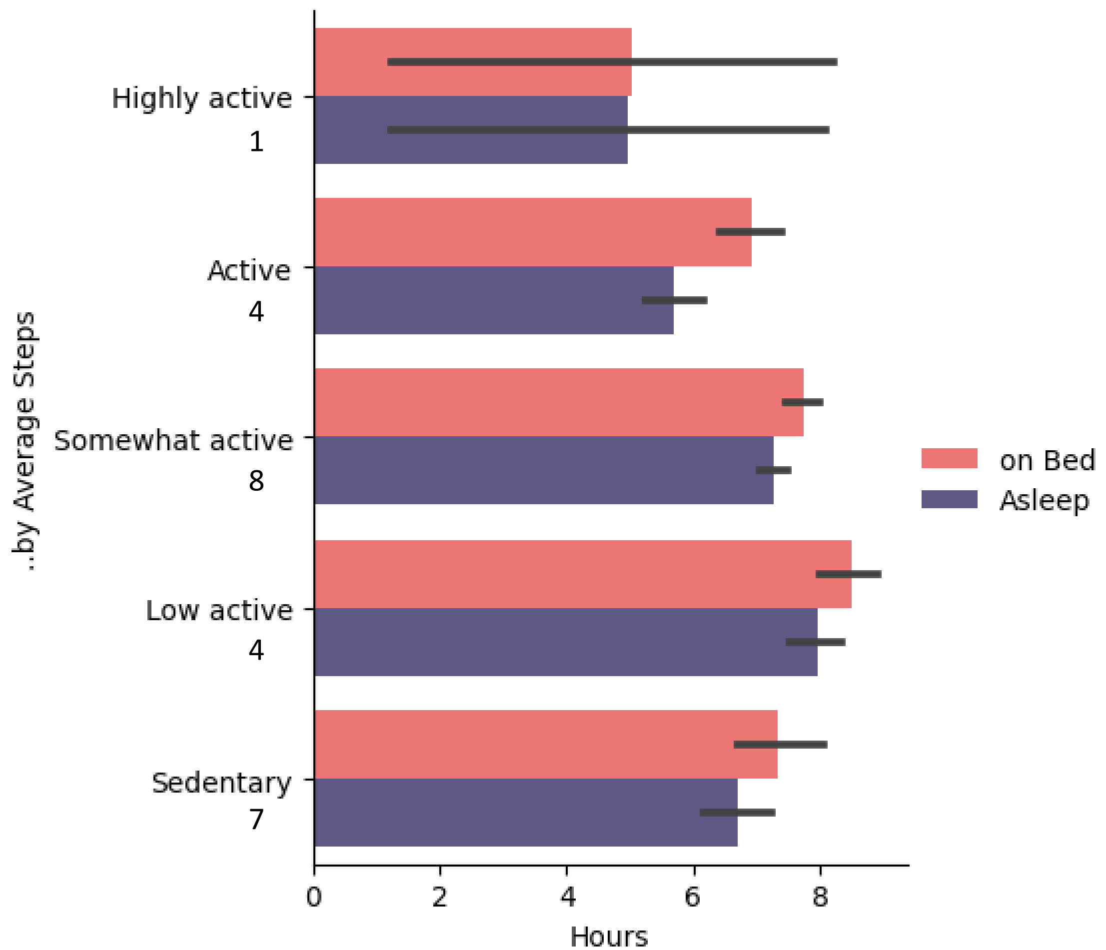
- Each dot in here, is a user
- Where they lie on the plot, depends on the average daily steps taken per day [X axis] and the average calories burnt per day [Y axis] by the user
- The recommended daily numbers for both metrics (for reference) are – 10,000 steps and ~2000 calories
- Here, it is surprising to see some 'Sedentary' users with <= 2000 avg. steps, burning more calories than 'Somewhat active' users
- Both the Highly active users tend to burn relatively high amount of calories
- So, from a Bird's Eye View we infer that - users belonging to the same category can have different body types (calories burnt) and may differ a lot from one another.

..now, we visualise the user traits – tracked duration & distance

- ✓ Each dot in here is a user
- ✓ Where they lie on the plot, depends on the average tracked duration per day [X axis] and the average tracked distance per day [Y axis] by the user
- ✓ Lot of users tracked their activity throughout the day; everyone in the group tracks more than 15 hours on average at least.
- ✓ In case of average Tracked duration, we can see that most users lie in either of the extremes, with very few outliers in the middle.
- ✓ In case of average Tracked distance, we can see that most of the users lie between the 4 to 7 km range, with some outliers at both extremes.
- ✓ So, from a Bird's Eye View we infer that - tracking habits doesn't affect the tracked duration or distance.

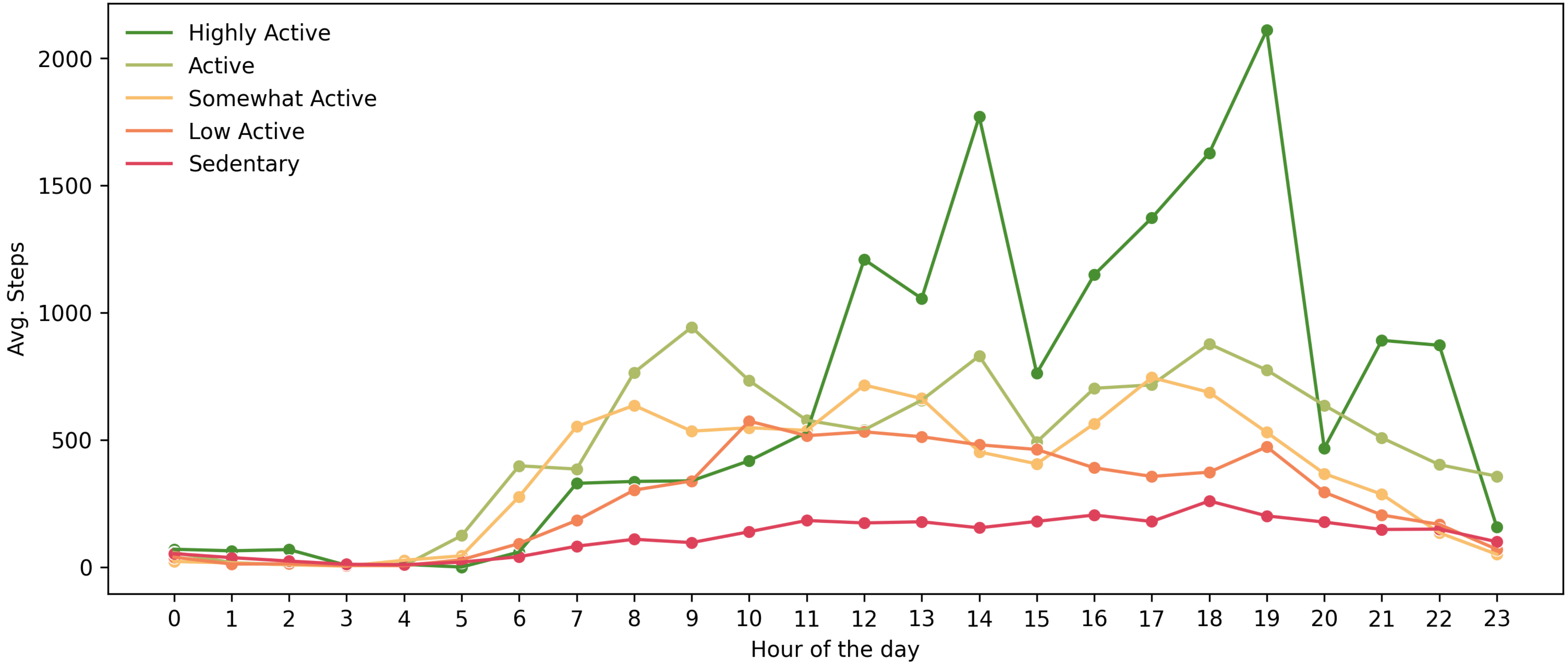


..sleep records were available for some users; let's visualise that



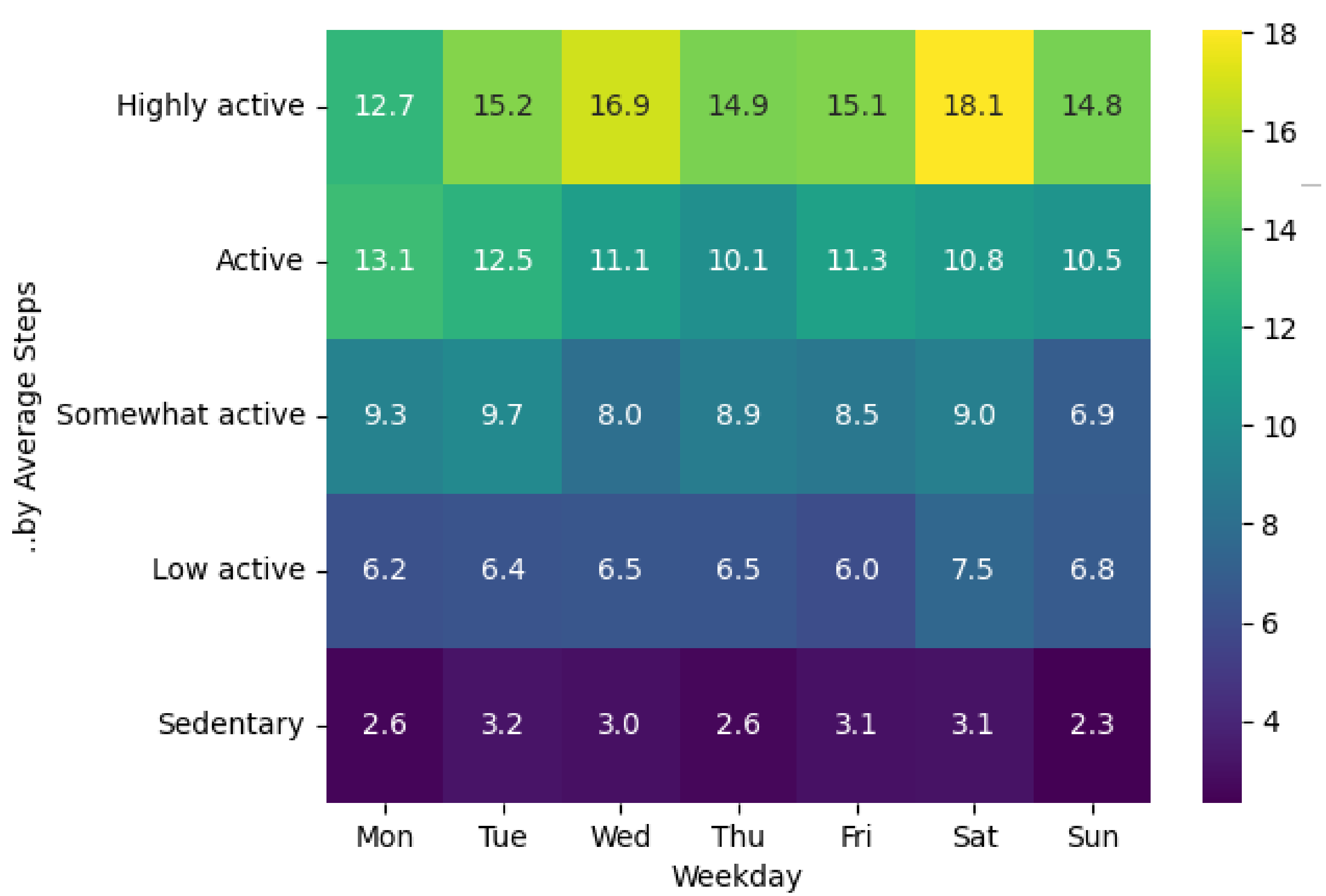
- ✓ This data is of 24 users (out of 33); the number beneath each category is the count of users in it
- ✓ This visual is a representation of date wise per day average of sleep duration by user Category
- ✓ The extent to which each bar extends towards the right represents the average value of sleep duration [X axis]. The black lines are 'error bars' which represent the range of deviation in data – i.e. for certain dates, the values may be lower than average or above average; longer the line, higher the deviation in values.
- ✓ The user group averages at 6.9 hours of sleep per day, against the recommended 8 hours.
- ✓ We observe here, that in general no category has an average of the recommended sleep duration.
- ✓ From a Bird's Eye View, we infer that – these users can potentially be sacrificing sleep for being more active. Active users, followed by Sedentary users take more time to fall asleep compared to other categories; overall sleep latency is ~39 minutes.

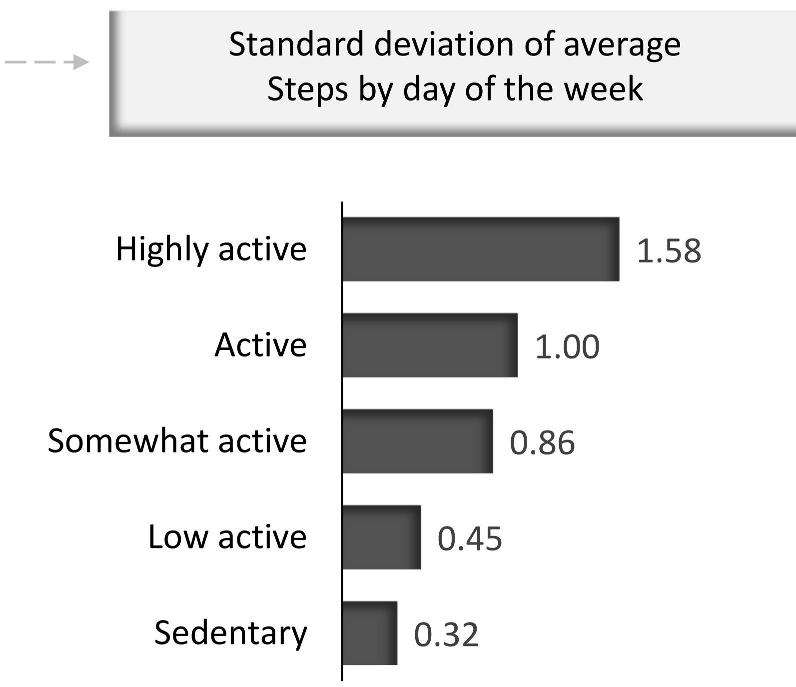
..let’s explore the average steps by hour of the day



- ✓ This line graph represents the average number of steps [Y axis] registered by a user in each hour [X axis] of the day; this visual is to understand preferred time for activity. e.g. 14th hour of the day means 2:00 pm to 3:00 pm
- ✓ We see here that - 5:00 pm to 8:00 pm is the most popular time bracket, when users register the most number of steps; overall, across all categories.
- ✓ This user group of 33, is not generally active before 9:00 am
- ✓ The hour with max value of each category is more than category below, aligning with the hierarchy of avg. steps.

..let's explore the average daily steps by day of the week





- ✓ This heatmap, represents the average daily steps of a user across the day of the week [X axis] by User Category based on average steps; this visual is to understand preferred day of the week for activity.
- ✓ We see here that – Saturdays and Tuesday were with the most number of average steps per day in the week, across all categories, with 8.15K and 8.12K steps respectively. Also, users are not that active on Sundays.
- ✓ We see a descending trend of standard deviation in average steps by day of the week across the categories – denoting that Highly active and Active users were more active on certain days of the week

Conclusion

- ✓ We can infer that the consumers come in varying lifestyles and body types and hence segregation of users based on the activity they prefer is quite important and using steps alone as the means to measure customer activity may not be an apt approach.
- ✓ Active user groups prefer to be more active on a certain days of the week and certain time frames during the day
- ✓ Lot of consumers lead a relatively sedentary life (overall avg. is ~7600 steps); they should be made aware of it. Also, most of the users lack quality sleep (overall avg. is ~6.9 Hours against the recommended 8 Hours) and have a high sleep latency (overall avg. of 39 minutes to fall asleep, against the recommended 20 minutes)
- ✓ Very few customers track their weights

Recommendations based on the data we analysed..

1. Bellabeat app can be designed in a way so that users can enter their height, weight, age while logging in or signing up – this info can be used to provide recommendations to the user to improve their overall health. User can be offered to opt in to regularly be notified to enter/track their weight if they have fitness goals.
2. Users should be able to create a plan, set workout times or choose the set of activities they perform every; with an arrangement to notify them of their schedule to keep them on track
3. The app should contain recommendations for a healthy lifestyle with information related to BMI, Diet, Sleep, Daily Activity, etc. and should notify the user on their activity dashboard in an instance of an unhealthy lifestyle
4. Hardware improvements
   i. Add functions in devices to track Heartbeat, oxygen level tracking & sleeping patterns
   ii. Finetune the devices to understand most common activities like walking, running, cycling, yoga, etc.