Bellabeat Case Study

Introduction:

This case study was performed for Bellabeat, a company who creates health focused tech products for women. Smart device data from a competitor with similar products was analyzed to discover trending. The purpose of this study was to deliver these insights to the marketing team to help them develop new strategies for campaigns.

Process:

The datasource was available through the user Mobius on Kaggle. The dataset was derived from thirty-three volunteer participants who gave permission for usage of their smart device data. The dataset was in the form of 18 cvs files and downloaded to SQL for analysis. Once the analysis was complete, visualizations were created through Tableau. You can find the code used for analysis and cleaning log files attached below, in this portfolio. Questions for the analysis were formed organically from cleaning, parsing and getting familiar with the dataset.

Summary:

Most of what I discovered through the data, confirmed my hypotheses of what the data might be about, but there were some surprises and insight learned along the way. The story the data tells is described below. First, through the questions I asked, then my findings to those questions and finally, the key takeaway actions.

Limitations:

Please note that limitations on this case study are as follows: The dataset is from a very small sample size of only 33 people with unknown population demographics. This could lead to bias in the study. The study is from only one source. The study is not current; it took place in 2016.

Questions:

1. Which features on the device are being used the most and how many users are using all three of these features: movement (distance/steps tracker), sleep and weight?

2. How are the intensities of user activities distributed in the movement feature?

3. What is the total and average distance recorded for each user and how do they compare to the rest of the users?

4. How many days did the participants use their device during the one month time period?

5. What time of the day and what days of the week are the users most active?

6. How many participants used the “Logged Activities” function?

7. How many hours of sleep on average are the users getting each night?

Findings:

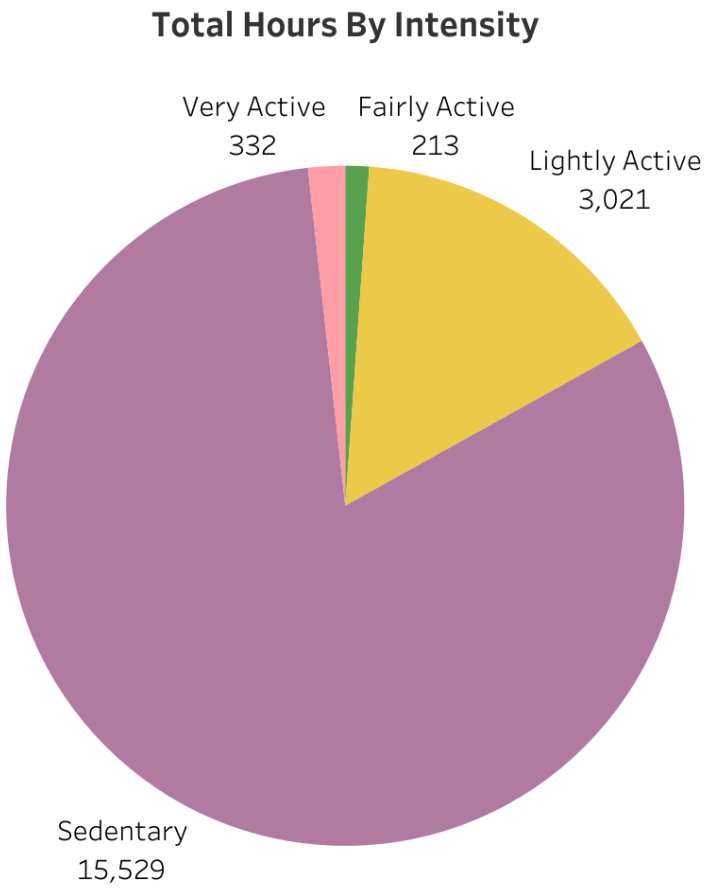
1. The movement feature was used the most. Movement was recorded for all 33 users, followed by the sleep feature (24 users) and then the weight feature (8 users).

TotalActivityId	TotalSleepId	TotalWeightId
33	24	8

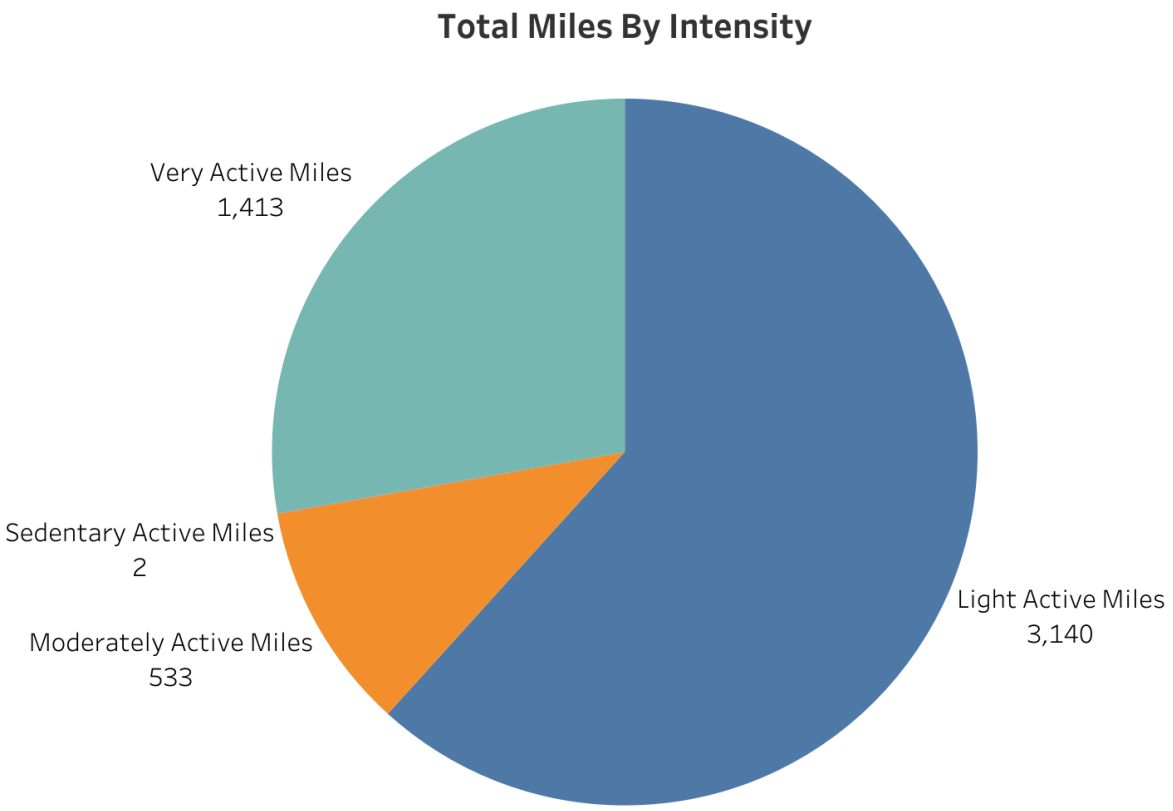
There were only 6 users that used all three of these features.

TotalActivityId	TotalSleepId	TotalWeightId
6	6	6

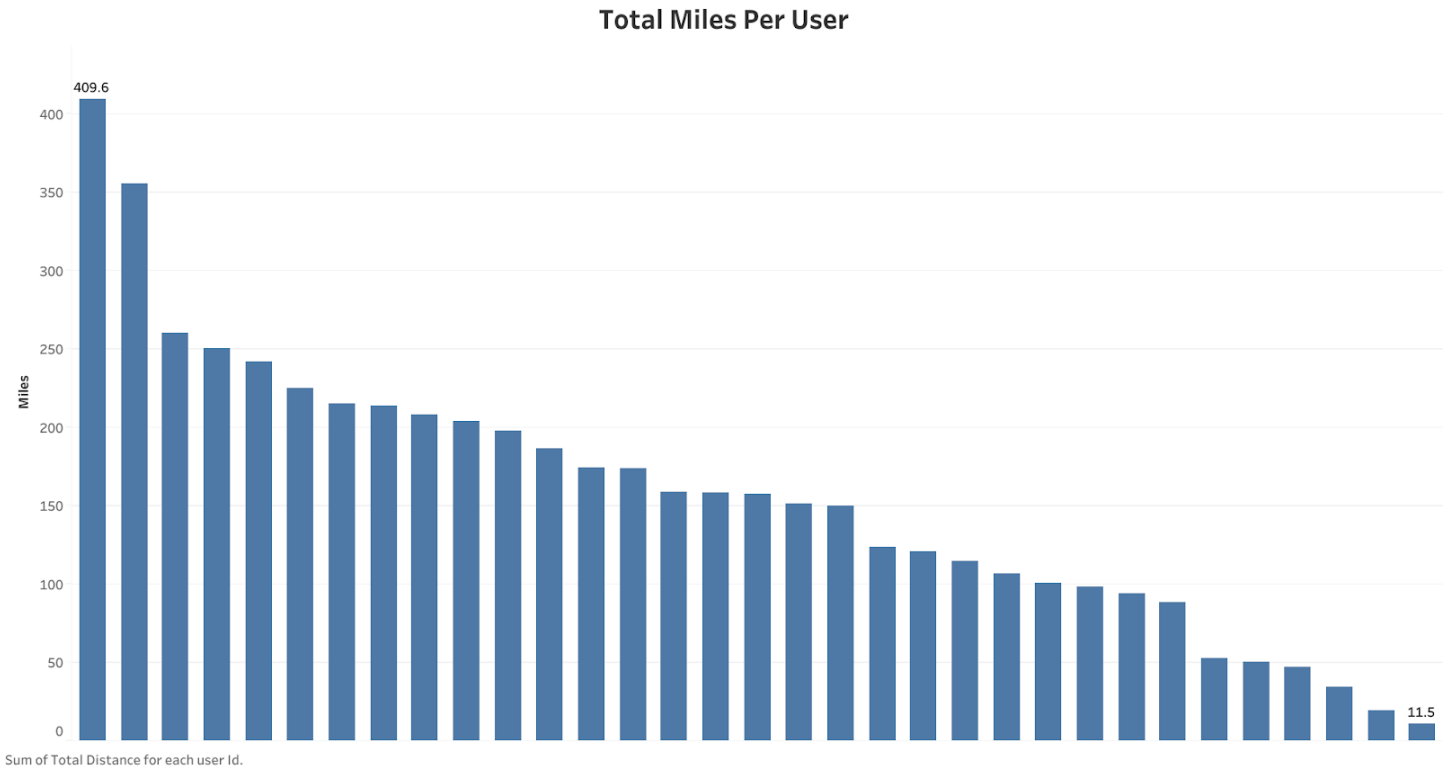
2. The intensity hours data tells us the activity levels of the users. The majority were sedentary and lightly active and a small portion did very active activities. This could imply that on a day to day basis they are wearing the device and it is tracking daily movement such as walking and light activities versus running or other vigorous exercise. They are wearing the device for most of the day, not just during exercise.



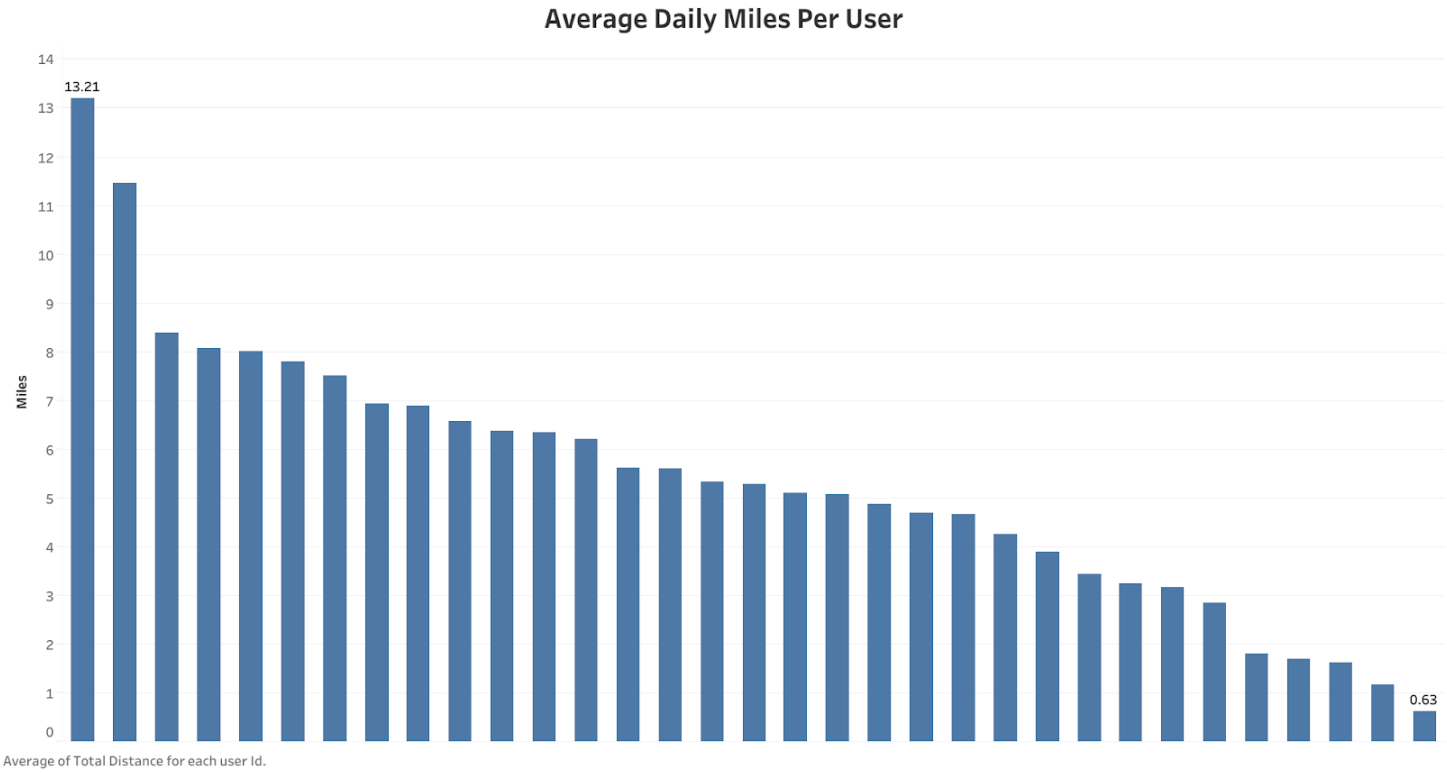
The intensity in miles data tells us that more than half of the users recorded light distance miles. These are most likely activities like walking or daily movements. A little less than half the users recorded miles in the very active and moderately active distance miles; most likely for more strenuous activities like running or aerobic exercises. So, the device is being used almost evenly for both purposes.



3. Similar to intensity, the majority of the users recorded miles for the 1 month time period totaling to around 50-250 miles per user.



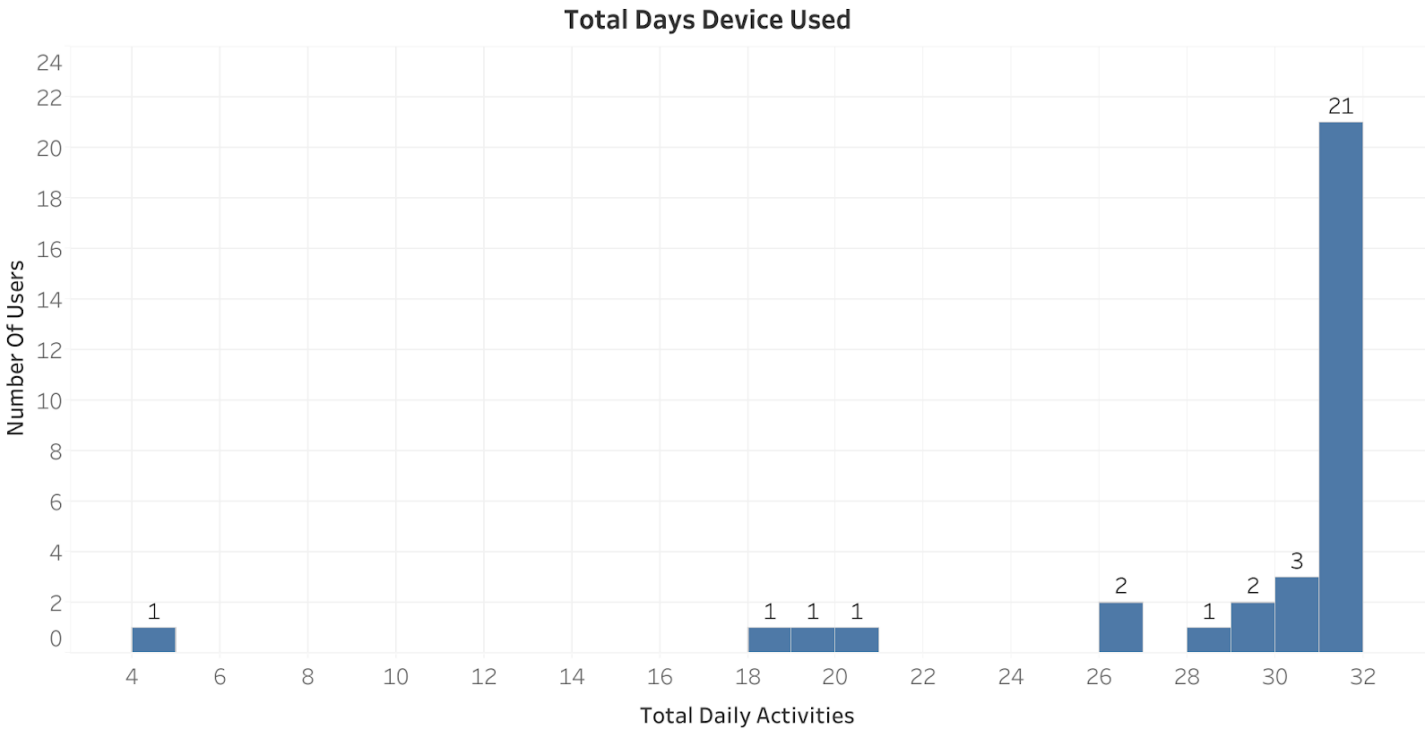
The average miles per user was around 4-7 miles a day, for the 1 month time period. Again, this confirms the hypothesis that the majority of the participants used their devices to track movement every day.



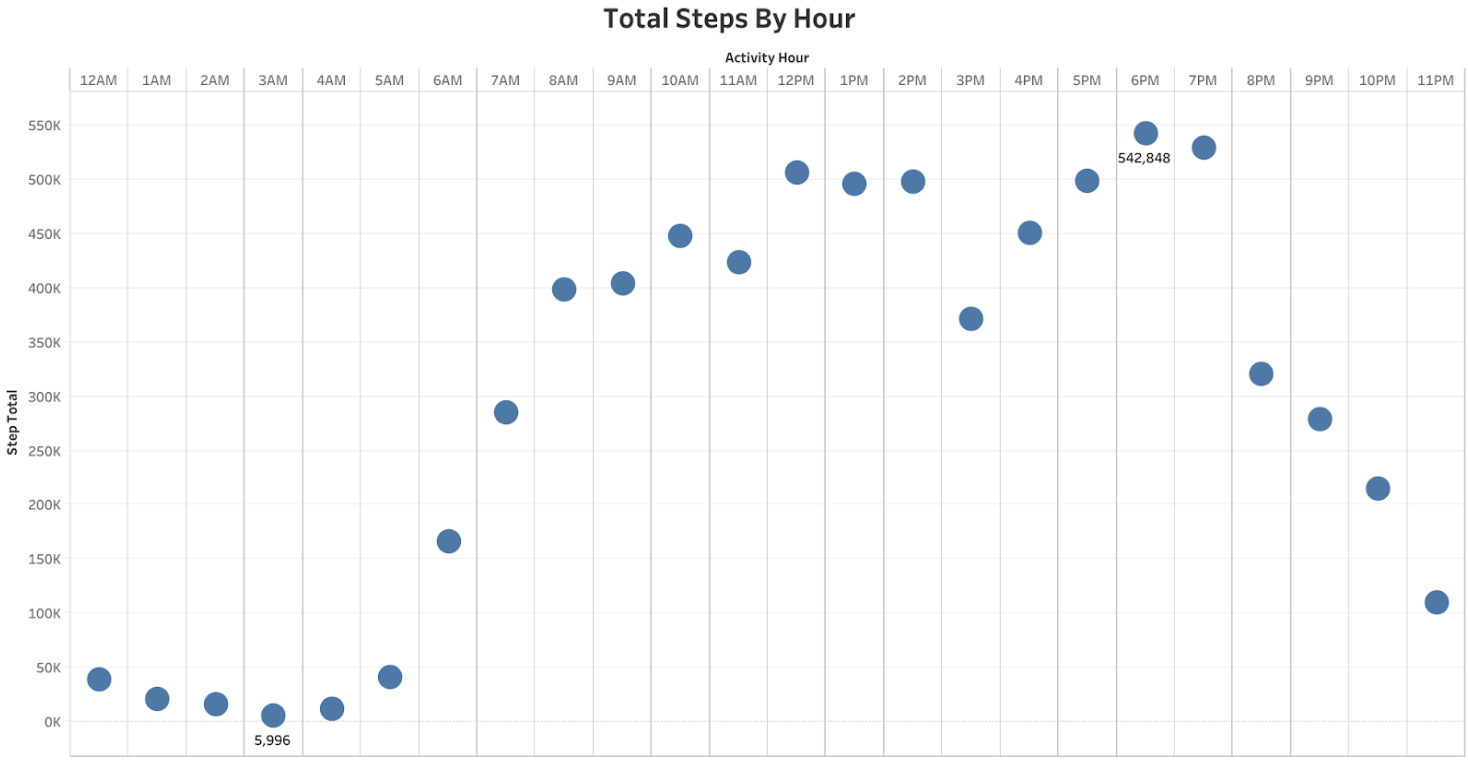
We confirm the average distance findings further, by pulling the average distance for all users, as well as the max and min:

AverageDistance	ShortestDistance	LongestDistance
5.4897021219882944	0.0	28.03000069

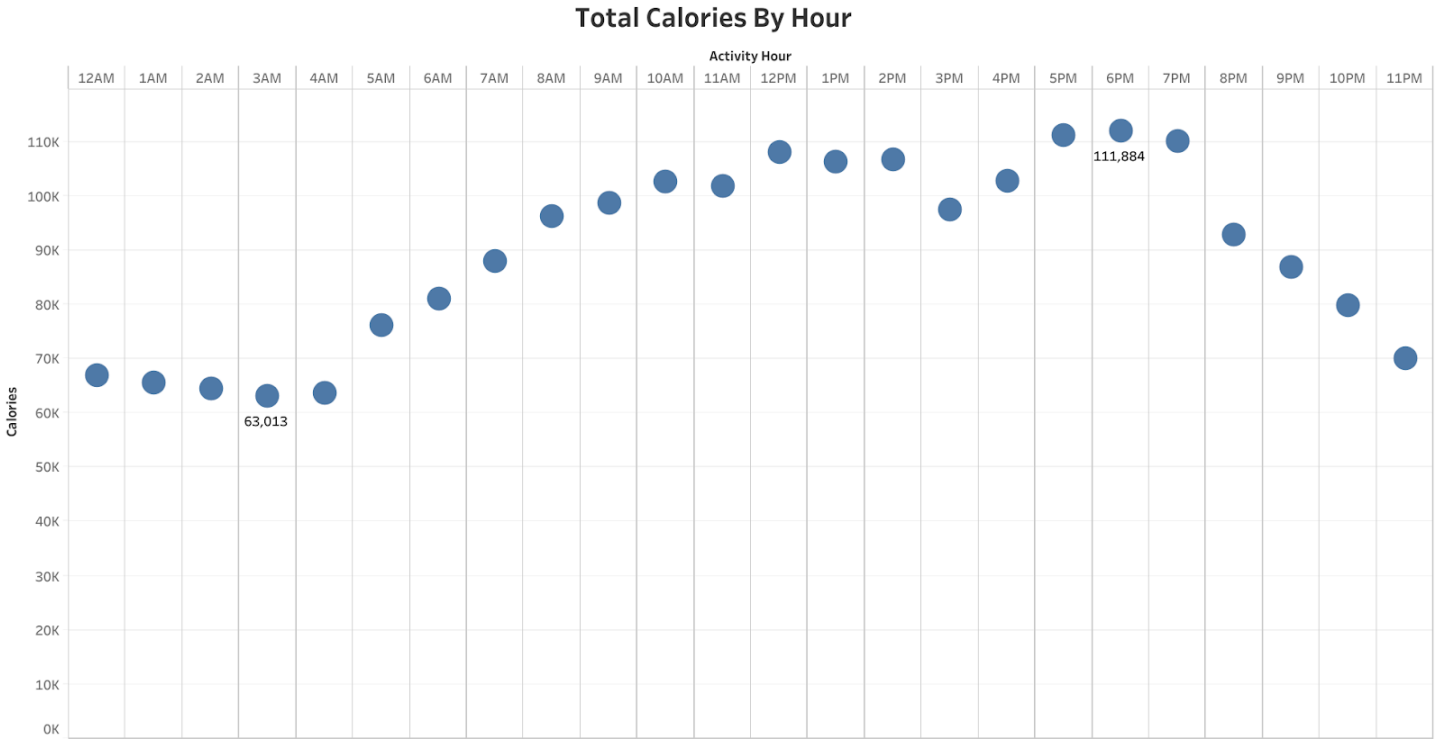
4. In addition, for the time period of April 12 through May 12, the amount of days the participants used the device was close to or equal to 31 days for almost all of the users. Once again, this solidifies the idea that most of the users are wearing this device every day.



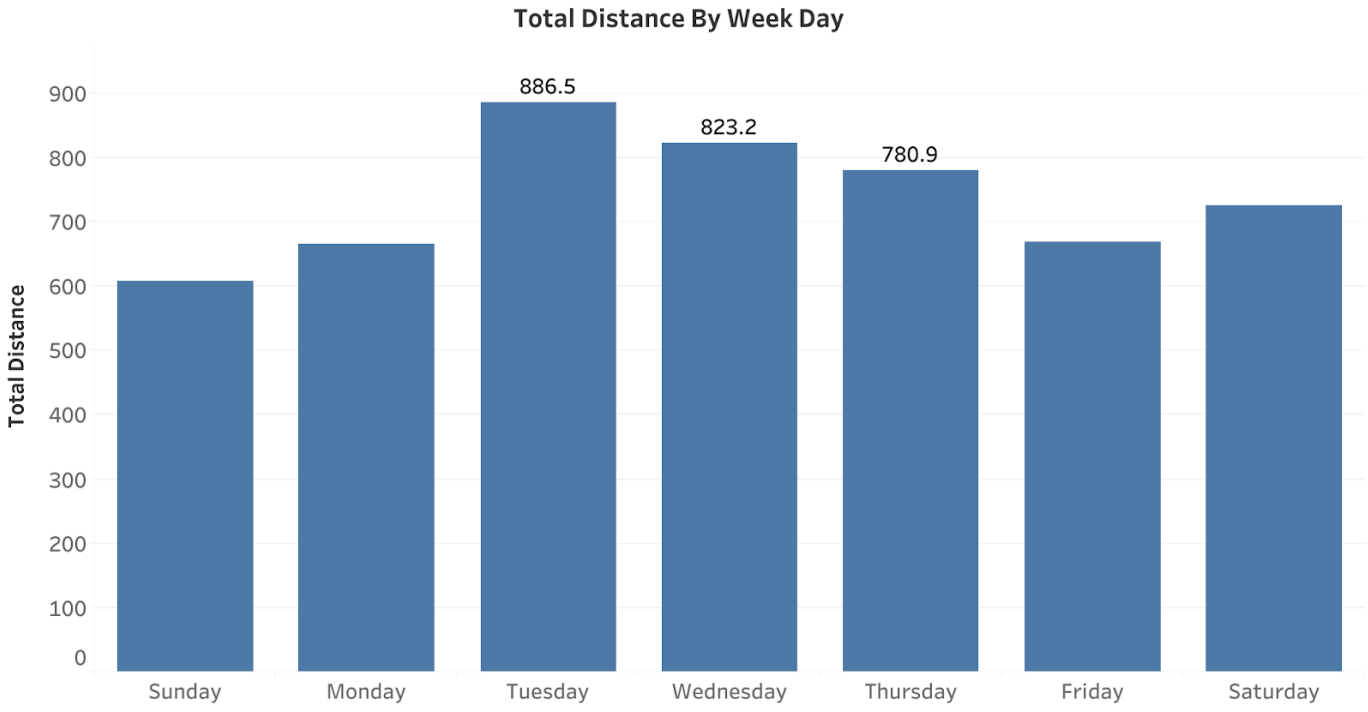
5. The time periods that the users were most active were between 8am and 7pm, with peak activity hours between 12-2pm (lunch time) and 5-7pm (after work).



To support these findings and further illustrate the way activity is distributed throughout the day, we graph the total calories burned per hour. The output looks similar to the total steps graph, with the exact same peak and minimum hours. The highest calorie burning time is at 6pm and the lowest is at 3am.



Most activities were done in the middle of the week (Tuesday, Wednesday, Thursday).



6. The “logged distance” function is barely being used at all. Only 4 of the 33 user participants used this feature.

NumberOfLoggedActivities
4

Out of those 4, only 3 used it to track a significant distance. This discovery should be investigated further.

Id	LoggedActivities
7007744171	55.085506676999991
8378563200	34.600891589
6962181067	10.034689188
6775888955	1.959596038

7. On average, the users are getting about 7 hours of sleep a night

Average	Least	Most
6.99	0.96	13.26

Actions:

Since the distance and steps tracker were the most used, the marketing team should put more of their focus on that feature. In the future, if they decide to develop more wearable device products, they could base the pricing around the use of these features. They could make a pricier model that has all 3 features (movement, sleep and weight) and a less expensive model that just tracks movement.

Another idea to consider is that not all the participants used the device for sleep tracking. More information is needed to figure out why this is the case. Perhaps, it's not comfortable to wear while sleeping or maybe the participants just aren't interested in those metrics. Alongside that, the weight tracking feature was barely used at all. Only 8 participants used it. The reason for this should be investigated further as well.

Based on the data, it is clear that the device is getting used all day, every day; not just for workouts. Some suggestions for the marketing team would be to also focus on this aspect. They could make the device more versatile to wear during work and daily living. They could advertise it as more comfortable, less intrusive and emphasize all the options and variety of styles for each unique lifestyle or situation.

We know when the users are most and least active. The marketing team could use this information to leverage their marketing campaign. They could put an emphasis on how the devices seamlessly transition from easing into activity in the morning to breaking a sweat with a good workout after lunch and then being productive and getting things done after a workday. They could do something similar with the weekly data.