By- Amish Garg

GOOGLE CAPSTONE PROJECT

Bellabeat Fitbit Tracker

- ## Task

To analyze Bellabeat's fitness tracker data to gain insight and derive trends from consumer usage of Fitbit app in order to recommend strategies for Bellabeat's marketing.

- ## About

Urška Sršen and Sando Mur founded Bellabeat, a high-tech company that manufactures health-focused smart products. Sršen used her background as an artist to develop beautifully designed technology that informs and inspires women around the world. Collecting data on activity, sleep, stress, and reproductive health has allowed Bellabeat to empower women with knowledge about their own health and habits. Since it was founded in 2013, Bellabeat has grown rapidly and quickly positioned itself as a tech-driven wellness company for women. By 2016, Bellabeat had opened offices around the world and launched multiple products. Bellabeat products became available through a growing number of online retailers in addition to their own e-commerce channel on their website. The company has invested in traditional advertising media, such as radio, out-of-home billboards, print, and television, but focuses on digital marketing extensively. Bellabeat invests year-round in Google Search, maintaining active Facebook and Instagram pages, and consistently engages consumers on Twitter. Additionally, Bellabeat runs video ads on YouTube and display ads on the Google Display Network to support campaigns around key marketing dates.

- ## Business Objectives

1. What are some trends in smart device usage?
2. How could these trends apply to Bellabeat customers?
3. How could these trends help influence Bellabeat marketing strategy?

- ## <u>Stakeholders</u>

Urška Sršen: Bellabeat's cofounder and Chief Creative Officer.
Sando Mur: Mathematician and Bellabeat's cofounder.
Bellabeat marketing analytics team: A team of data analysts responsible for collecting, analysing, and reporting data that helps guide Bellabeat's marketing strategy.
 Customers: Using Bellabeat Fitbit app.

- ## **<u>Limitations</u>**
1) There are only 31 unique customers.
2) Datasets are 8 years old.
3) Cannot confirm the reliability of the data.

- ## **<u>About the Data</u>**
1) Source of the data is Kaggle: Fitbit Fitness Tracker Data.
2) I have used activity, weight, sleep datasets.
3) Data is generated by respondents from a survey via Amazon Mechanical Turk between 12 March 2016 to 12 May 2016.

- ## **<u>Process</u>**

1) I have used Python to solve the above stated business problem.
2) Firstly, I opened Python through Jupyter Notebook and imported all the required libraries alongside my data.
3) Then, I spent a lot of time in exploring and understanding my data.
4) Then, I start by checking my data for all the inconsistencies that can hinder my analysis.
5) After checking, I clean my data by removing all the null values.
6) Later, I perform some useful codes by transforming and manipulating my data and also by doing some necessary calculations which will help me with data visualisations.
7) Also, I merged the datasets using primary key.
8) Lastly, I created various Data Visualisations in order to identify trends between variables.

- **Visuals**

No of Calories Burnt Vs Total Steps
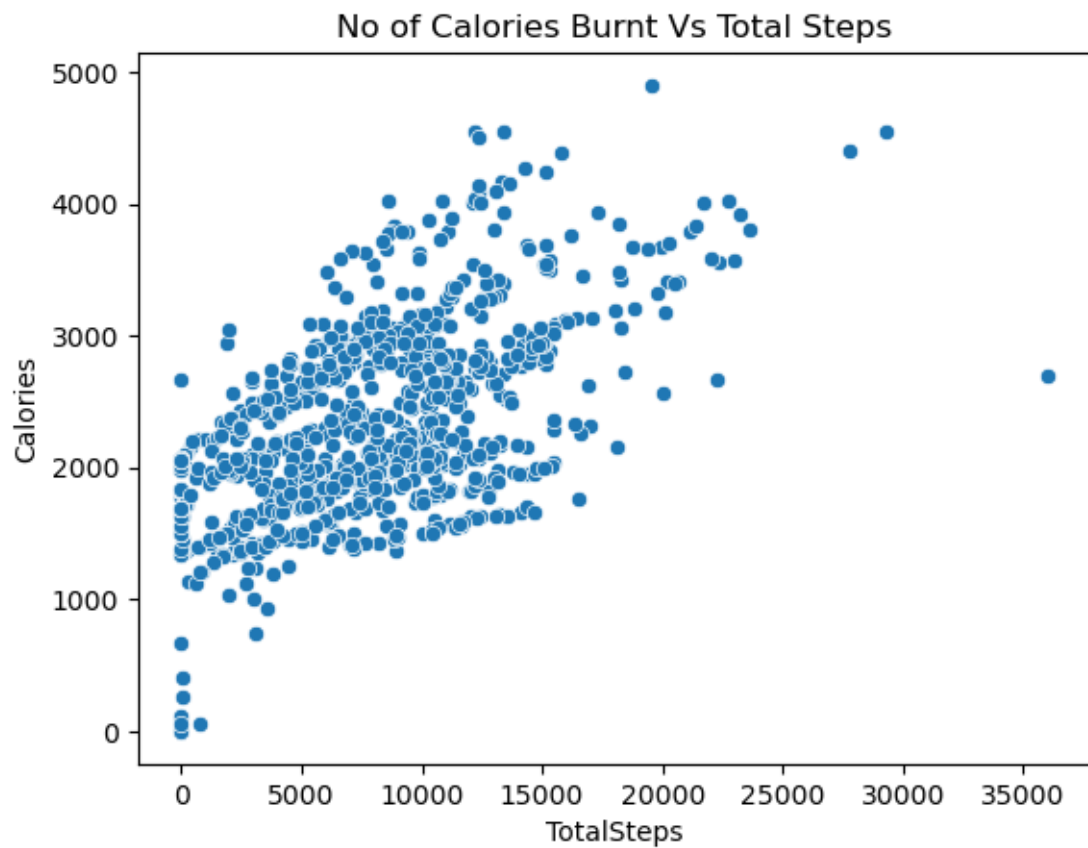
- **Inference**

1) We can clearly see from the above scatterplot between Total Steps walked by a customer and calories burnt that there is a positive correlation between them.
2) This means that more a customer is going to take steps will result in more calories burnt.
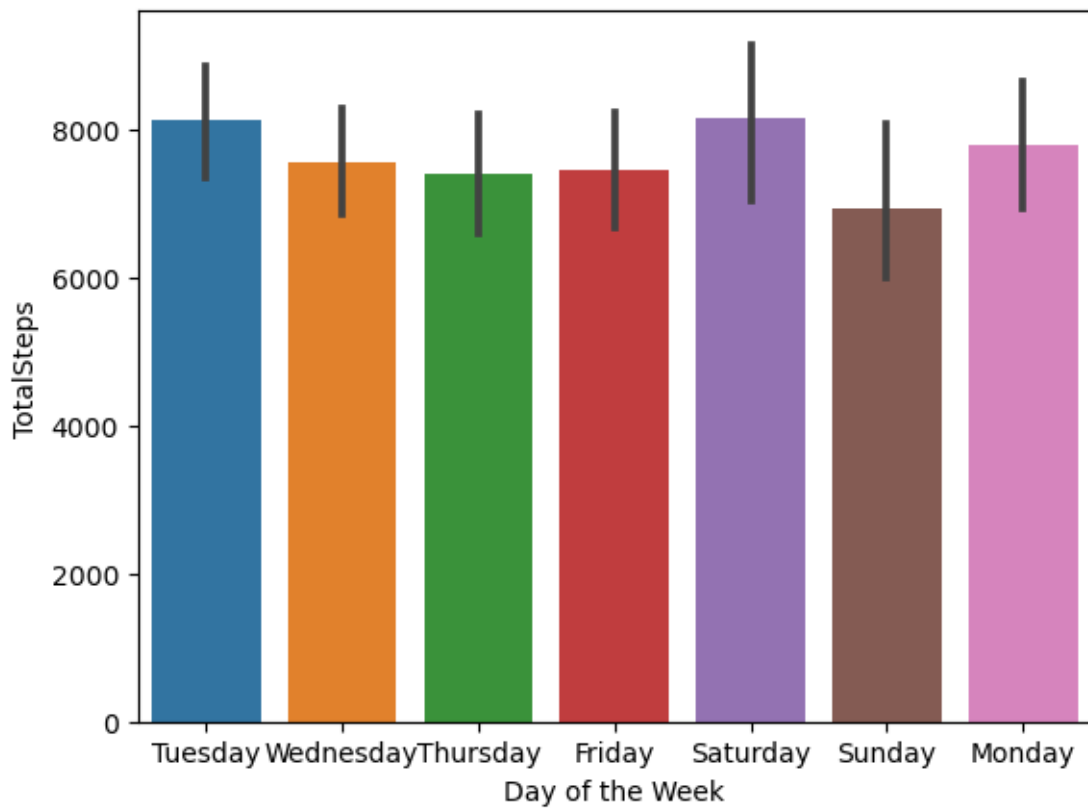
- **<u>Inference</u>**

1) This is a bar graph between each day of the week and total steps walked.
2) We can clearly see that on Sunday, a person has walked less steps compared to other days of the week which can infer to more sedentary activity compared to actual activity.

Total Fat Vs Total Steps
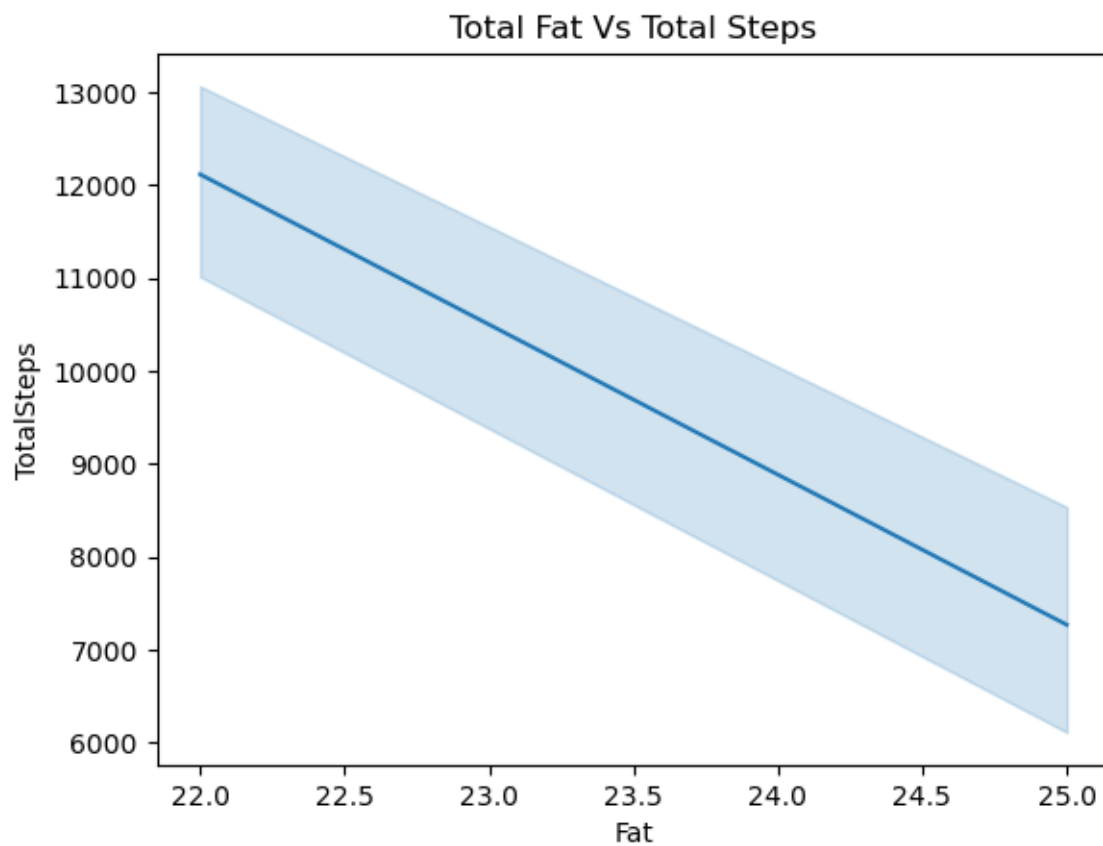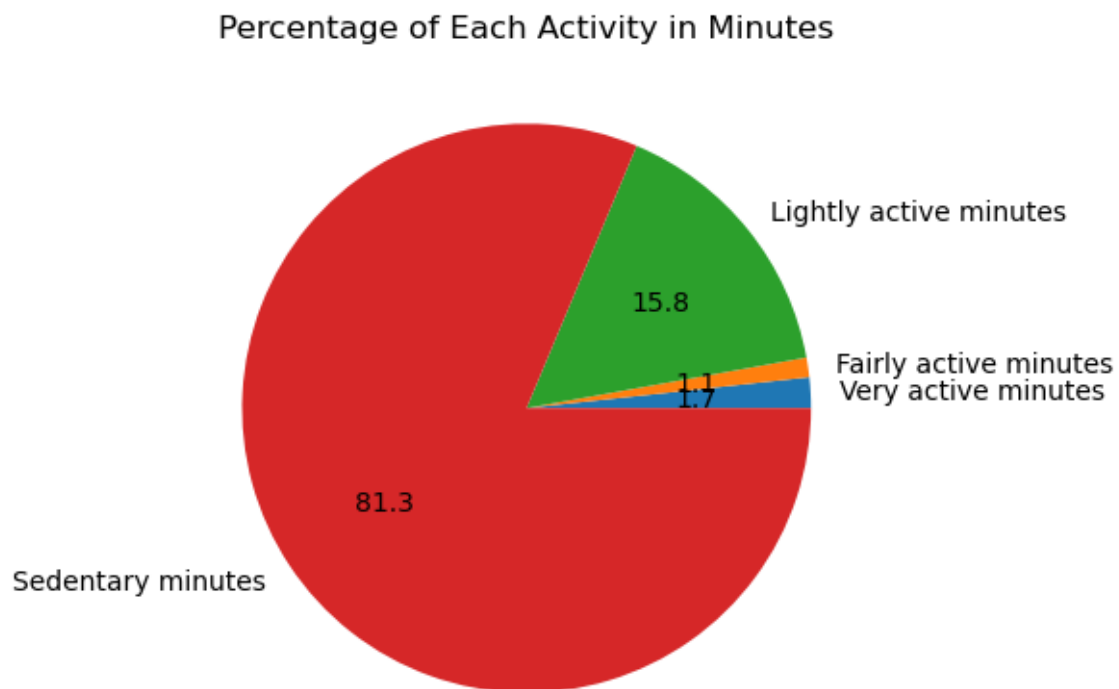
- **Inference**

1) This is line between total steps taken by a customer and amount of fat in their body.
2) We can clearly see a negative relation between them.
3) On 22 Fat, a customer has walked highest number of steps i.e., 12000 and at 25 Fat, a customer has walked least number of steps.
4) Hence, as the number of steps increases, the number of fat decreases.
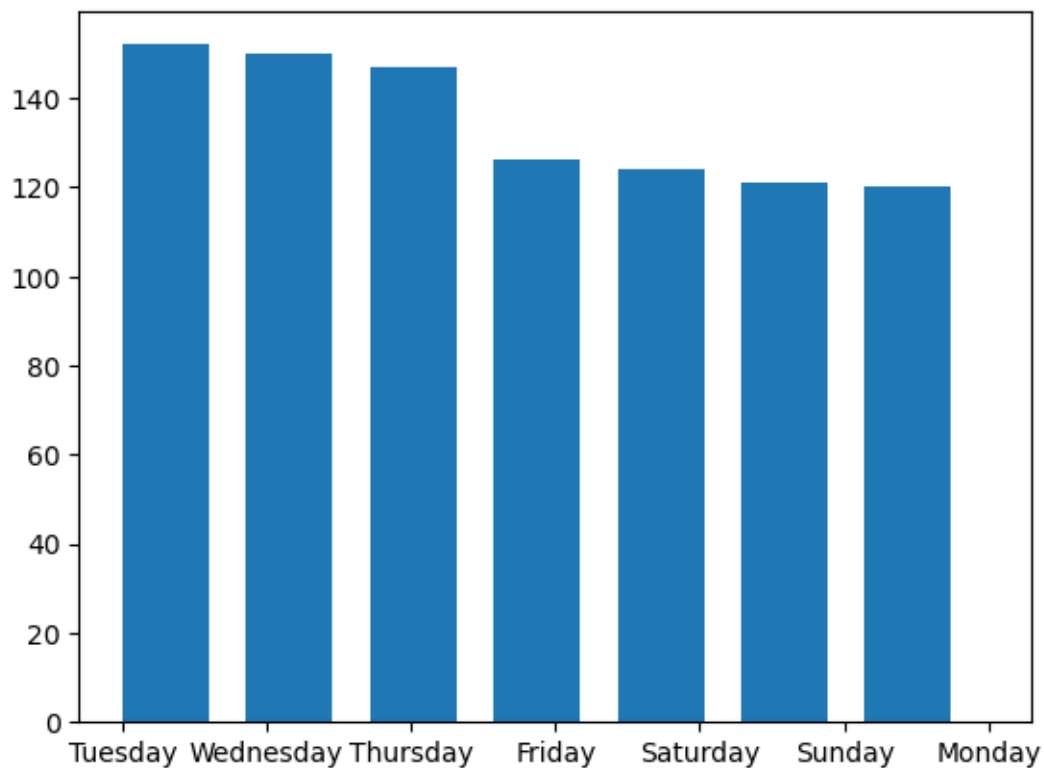
- **Inferences**

   1) This is a histogram of how many times users have logged into the bellabeat app on each day of the week.
   2) We can see that more users have logged into the app during the mid-week from Tuesday to Thursday.
   3) Number of logins in the app reduces during the weekends.

Percentage of Each Activity in Minutes

Lightly active minutes

15.8

Fairly active minutes
Very active minutes

1.1
1.7

81.3

Sedentary minutes

- **Inferences**

  1) This is a pie chart of the distribution of each type of activity that a user spends their time doing.
  2) We can clearly see that the most amount of time spent by a user is by doing sedentary activity i.e., 81%.

Weight in Kg Vs Total Distance
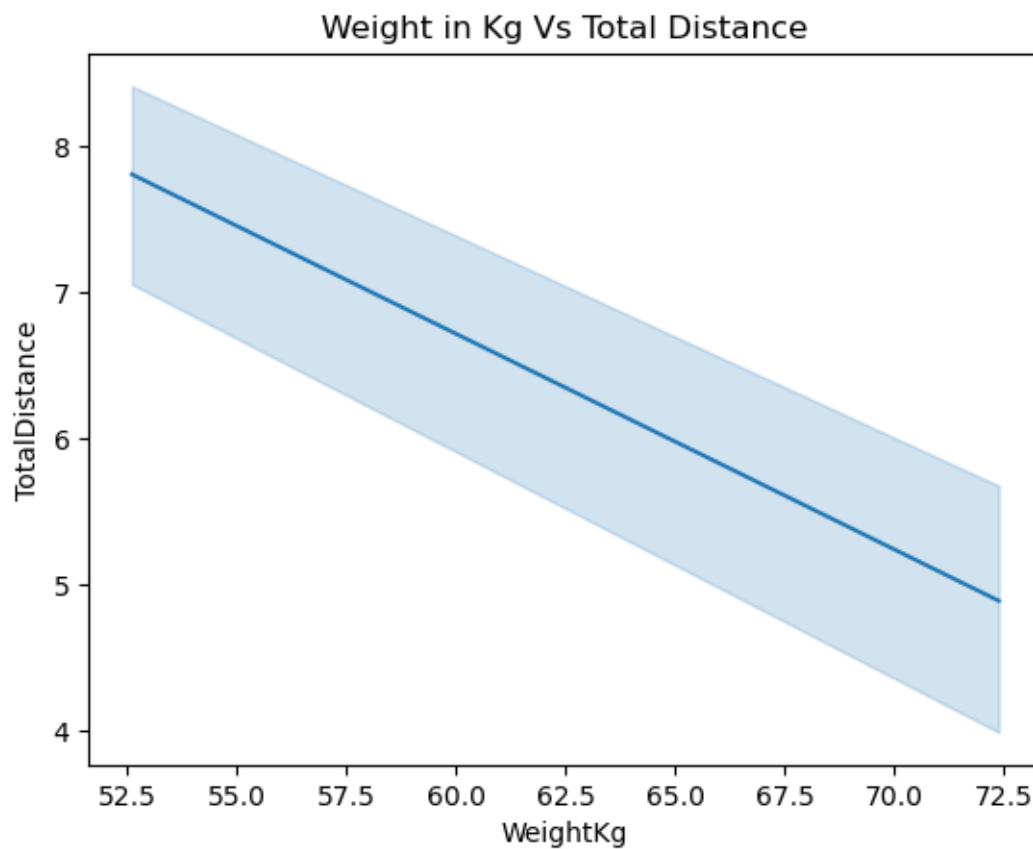
- **Inference**

  1) Again, this is a line plot between total distance covered and weight.
  2) We can see there is a negative relationship between total distance covered by a user and their weight.
  3) As on 52.5 kgs which is the least weight, the distance covered is 8 kms and on 72.5 kgs weight, distance covered is around 4.5 kms.

```
df.describe()
```

	Id	ActivityDate	TotalSteps	TotalDistance	TrackerDistance	LoggedActivitiesDistance	VeryActiveDistance	ModeratelyActiveDistan
count	9.400000e+02	940	940.000000	940.000000	940.000000	940.000000	940.000000	940.0000
mean	4.855407e+09	2016-04-26 06:53:37.021276672	7637.910638	5.489702	5.475351	0.108171	1.502681	0.5675
min	1.503960e+09	2016-04-12 00:00:00	0.000000	0.000000	0.000000	0.000000	0.000000	0.0000
25%	2.320127e+09	2016-04-19 00:00:00	3789.750000	2.620000	2.620000	0.000000	0.000000	0.0000
50%	4.445115e+09	2016-04-26 00:00:00	7405.500000	5.245000	5.245000	0.000000	0.210000	0.2400
75%	6.962181e+09	2016-05-04 00:00:00	10727.000000	7.712500	7.710000	0.000000	2.052500	0.8000
max	8.877689e+09	2016-05-12 00:00:00	36019.000000	28.030001	28.030001	4.942142	21.920000	6.4800
std	2.424805e+09	NaN	5087.150742	3.924606	3.907276	0.619897	2.658941	0.8835

8 rows × 21 columns

```
df.describe()
```

	erratelyActiveDistance	LightActiveDistance	SedentaryActiveDistance	...	FairlyActiveMinutes	LightlyActiveMinutes	SedentaryMinutes	Calories	total_mins
	940.000000	940.000000	940.000000	...	940.000000	940.000000	940.000000	940.000000	940.000000
	0.567543	3.340819	0.001606	...	13.564894	192.812766	991.210638	2303.609574	1218.753191
	0.000000	0.000000	0.000000	...	0.000000	0.000000	0.000000	0.000000	2.000000
	0.000000	1.945000	0.000000	...	0.000000	127.000000	729.750000	1828.500000	989.750000
	0.240000	3.365000	0.000000	...	6.000000	199.000000	1057.500000	2134.000000	1440.000000
	0.800000	4.782500	0.000000	...	19.000000	264.000000	1229.500000	2793.250000	1440.000000
	6.480000	10.710000	0.110000	...	143.000000	518.000000	1440.000000	4900.000000	1440.000000
	0.883580	2.040655	0.007346	...	19.987404	109.174700	301.267437	718.166862	265.931767

ActiveMinutes	LightlyActiveMinutes	SedentaryMinutes	Calories	total_mins	WeightKg	WeightPounds	Fat	BMI
940.000000	940.000000	940.000000	940.000000	940.000000	62.000000	62.000000	62.000000	62.000000
13.564894	192.812766	991.210638	2303.609574	1218.753191	62.500000	137.788914	23.500000	25.050000
0.000000	0.000000	0.000000	0.000000	2.000000	52.599998	115.963147	22.000000	22.650000
0.000000	127.000000	729.750000	1828.500000	989.750000	52.599998	115.963147	22.000000	22.650000
6.000000	199.000000	1057.500000	2134.000000	1440.000000	62.500000	137.788914	23.500000	25.050000
19.000000	264.000000	1229.500000	2793.250000	1440.000000	72.400002	159.614681	25.000000	27.450001
143.000000	518.000000	1440.000000	4900.000000	1440.000000	72.400002	159.614681	25.000000	27.450001
19.987404	109.174700	301.267437	718.166862	265.931767	9.980819	22.003940	1.512245	2.419593

- **<u>Inference</u>**
  1. We can see the average steps taken by the users are 7638 and average distance covered is 5.5 kms.
  2. We can see the average BMI of all users is above 25.

- ## **<u>Conclusion</u>**

1) As per health guidelines, a person should walk 10000 steps or cover 8 kms per day to maintain good health and as per bellabeat's fitness details, we can conclude that none of that conditions are getting satisfied.
2) As per health guidelines, a person should maintain their BMI between 20-25, but here in case of bellabeat users, their BMI is above 25 which tells that most of the users are overweight.
3) Through our visualisations we can conclude that consistency in taking adequate number of steps per day can lead to weight loss and reduction of fat.
4) Also, a greater number of steps taken leads to a greater number of calories being burnt.
5) Hence, we can also conclude that higher number of calories being burnt leads to weight loss and reduction of fat.
6) The biggest area of concern is user activity type. As we can see from the pie chart that approximately 81% of user activity is taken over by sedentary activity which is daily life chores and only 19% is covered by combination of moderate, fairly and very active users.

- ## **<u>Recommendations</u>**

1) Daily health report should be made available in the user's app by 10 pm everyday which world display all the required information about a person's activity. It should also mention key indicators where the user has done will alongside areas to improve.
2) Weekend targets should be given to every individual by Friday so that they can use the Fitbit tracker more during the weekdays as this move will make them transition from sedentary activity to fairly or highly active.
3) Bellabeat app can also add a section of health guidelines where information of general health is made to be available for users to read in their free time which will motivate them to stay on right fitness track.

Thank You