Bellabeat Case Study

An analysis of Fitbit users' data to motivate consumer engagement with Bellabeat devices

9/9/2023


bellabeat

Business Task

What are some trends in Fitbit device usage, and how might these trends apply to Bellabeat customers?

How can these trends influence our marketing strategy?

Outline

- Discuss data management, integrity checks, and cleaning process
- Discuss key findings from analysis with visualizations
- Suggest key recommendations

Data Source

Source link: FitBit Fitness Tracker Data | Kaggle

Data management/checking/cleaning

1. Manually update data using python to include times/dates in proper timestamp form. This was necessary for clean and streamlined uploading of all data tables to BigQuery.
2. Check number of unique user ids in each table.

   Most table have data from all 33 participants. This shows that not all participants record all available metrics.

1. Check for records with null Ids (**none!**)

Table Name	Number of unique Ids
dailyActivity_merged	33
dailyCalories_merged	33
dailyIntensities_merged	33
dailySteps_merged	33
heartrate_seconds_mergedupdated	14
hourlyCalories_mergedupdated	33
hourlyIntensities_mergedupdated	33
hourlySteps_mergedupdated	33
minuteCaloriesNarrow_mergedupdated	33
minuteIntensitiesNarrow_mergedupdated	33
minuteMETsNarrow_mergedupdated	33
minuteSleep_mergedupdated	24
minuteStepsNarrow_mergedupdated	33
sleepDay_mergedupdated	24
weightLogInfo_mergedupdated	8

Questions to Inform Business Task

1. What are some basic statistics for metrics including daily steps, distance, calories, and level of activity?
2. At what time(s) during the day are users most active?
3. What are some important trends and non-trends, including calories v. distance travelled/steps taken, calories v. time asleep, etc?

Data Analysis Process

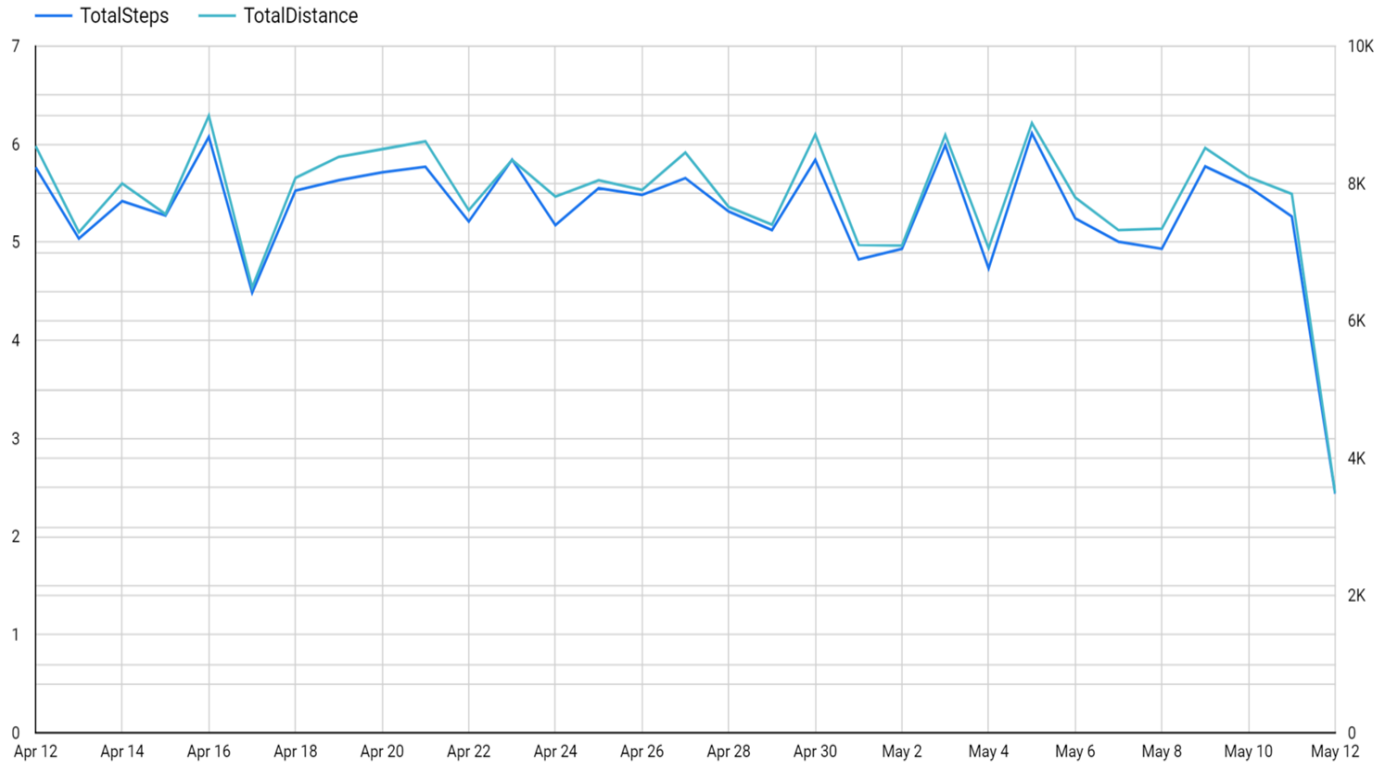
1. Obtain general statistical information (e.g. five-number summaries).
2. Filter/aggregate large data sets for producing plots.
3. Draw defensible conclusions from plots to answer previous questions and motivate business recommendations.

General Stats

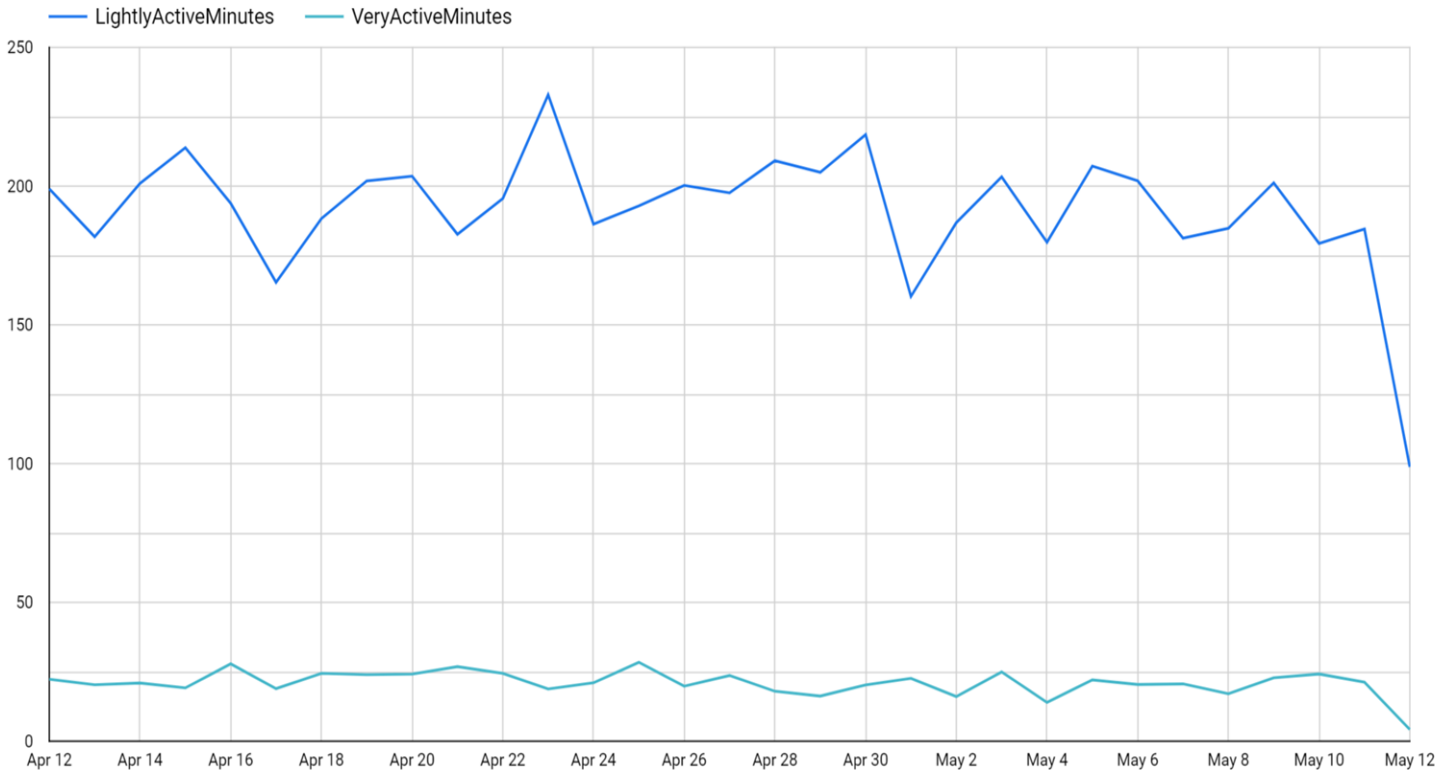
BigQuery used to obtain 1st, 2nd, 3rd quartiles and means for daily steps, calories, distance, and activity levels. Values are from total data records, and is not grouped by participant Id or day.

Metrics	1st Quartile	2nd Quartile	3rd Quartile	Average
Steps	3789	7399	10725	7637.9
Total Distance (all activity levels included)	2.62	5.24	7.71	5.5
Calories	1827	2133	2793	2303.6
Lightly Active Minutes	127	199	264	192.8
Fairly Active Minutes	0	6	19	13.6
Very Active Minutes	0	4	32	21.2

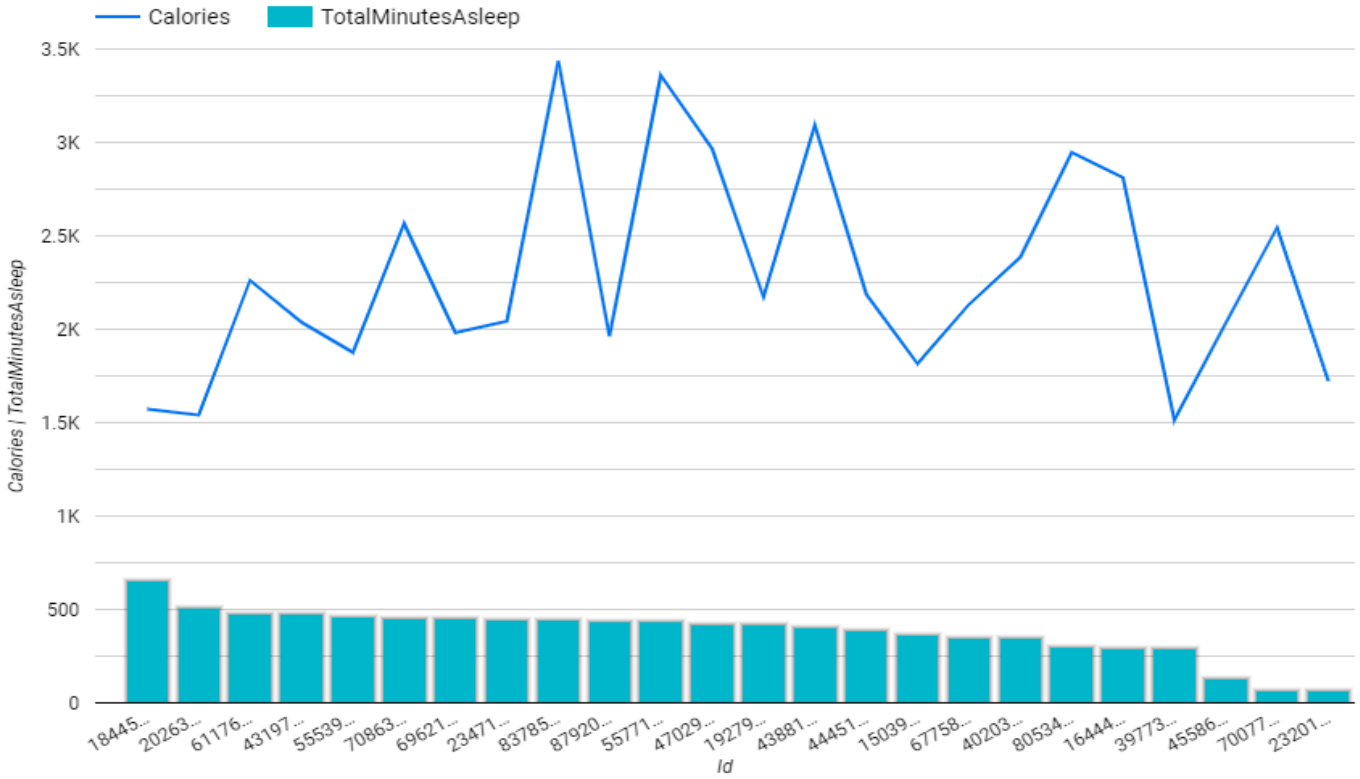
Average daily steps/distance



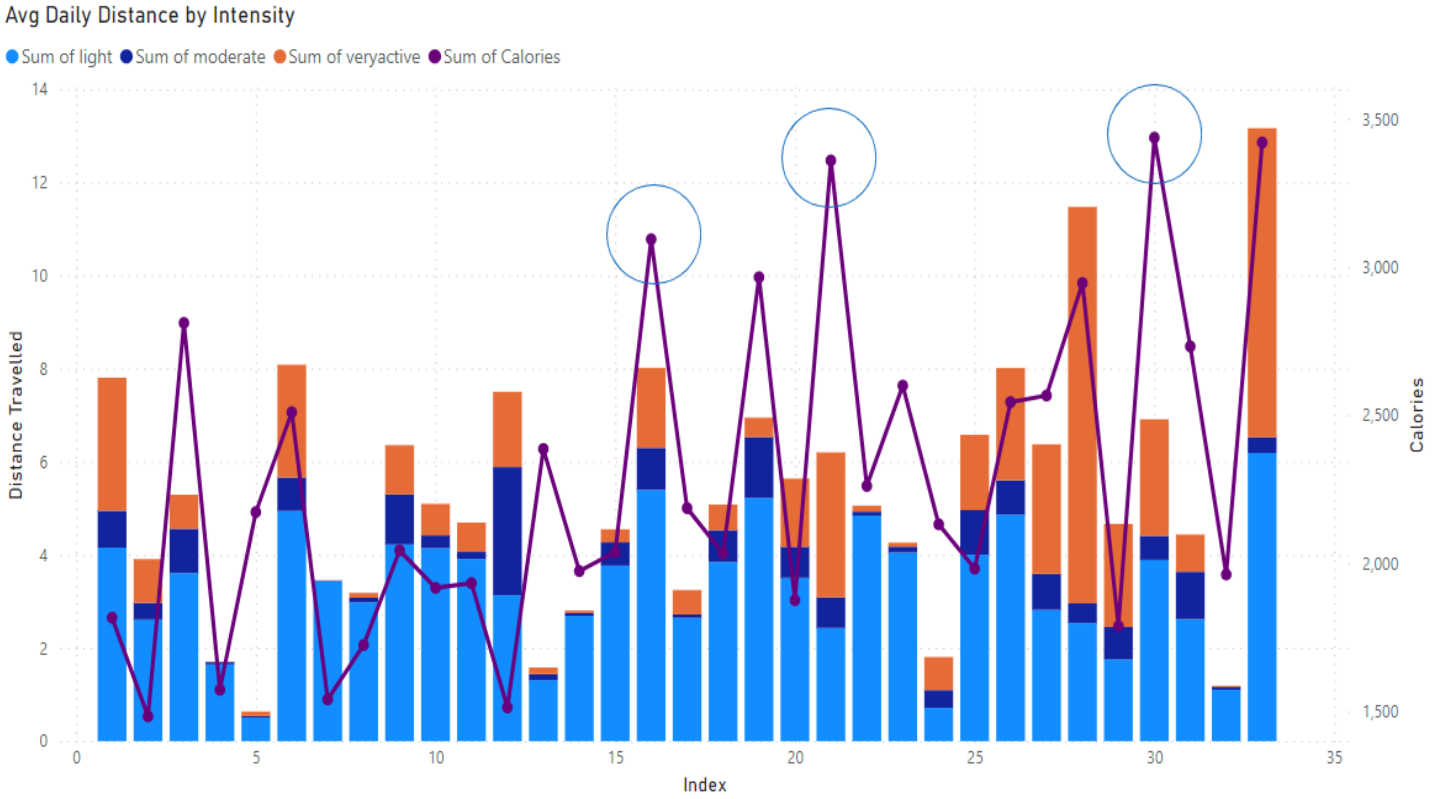
Average daily activity intensity



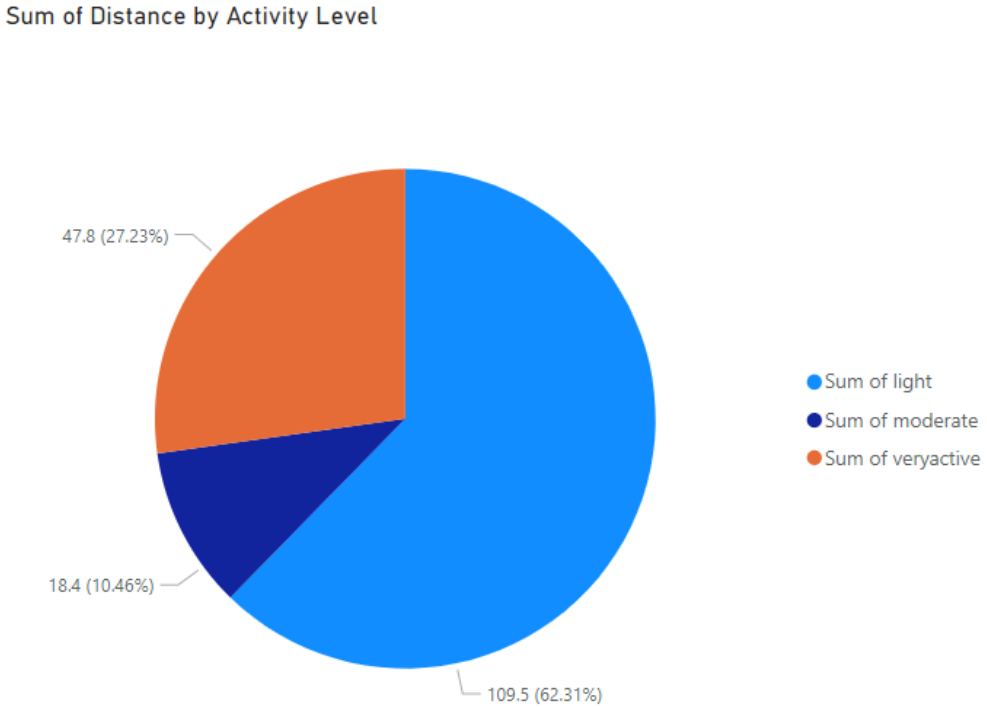
Calories and Sleep



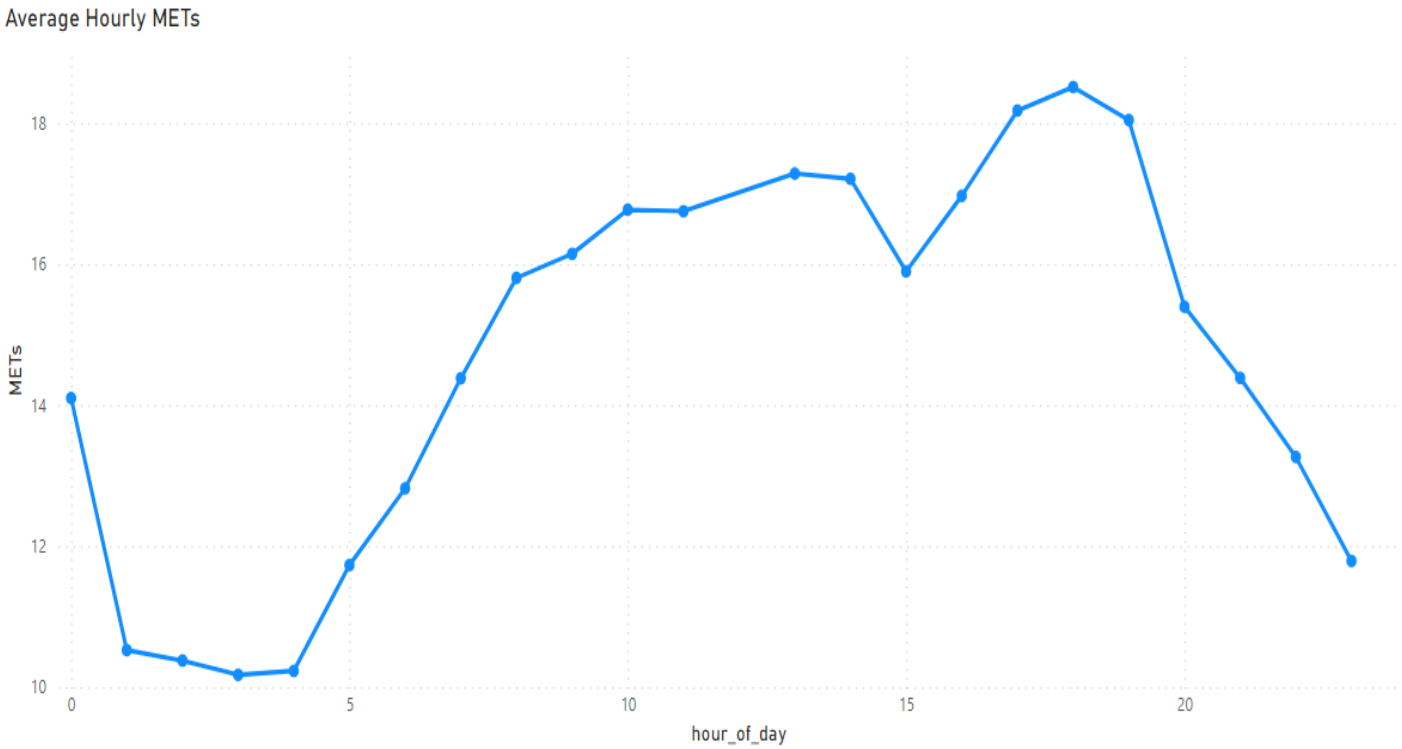
Calories and Active Distance



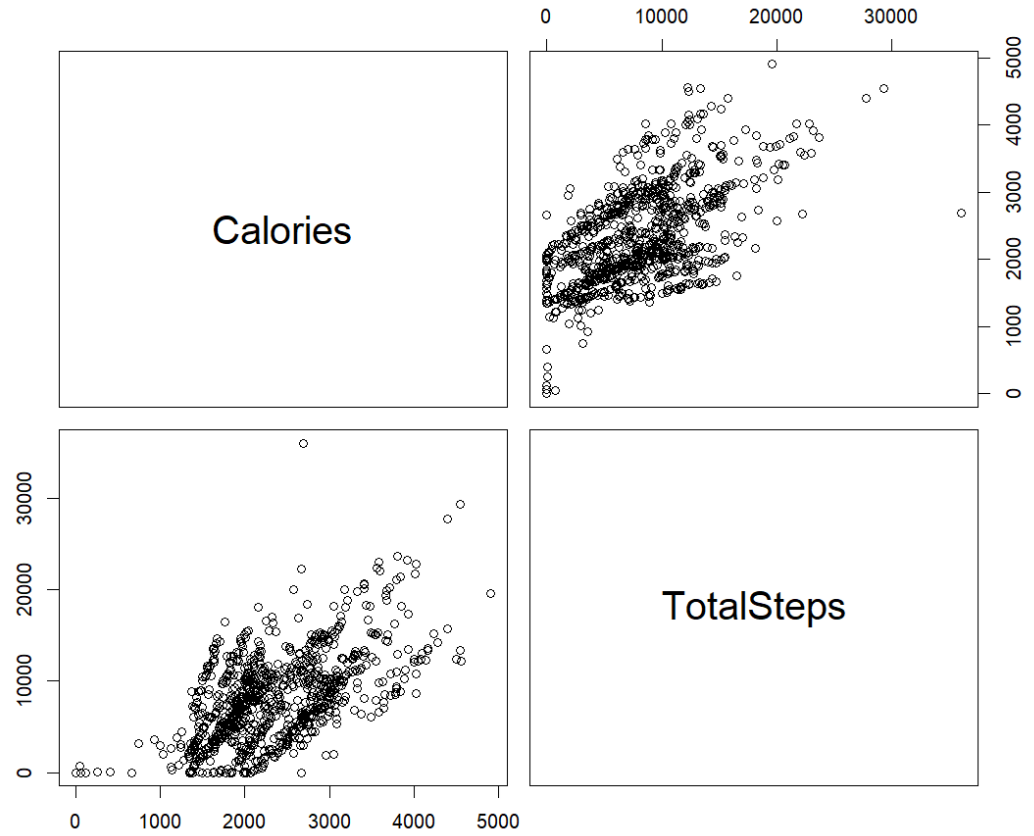
Overall Percentages of Active Distance

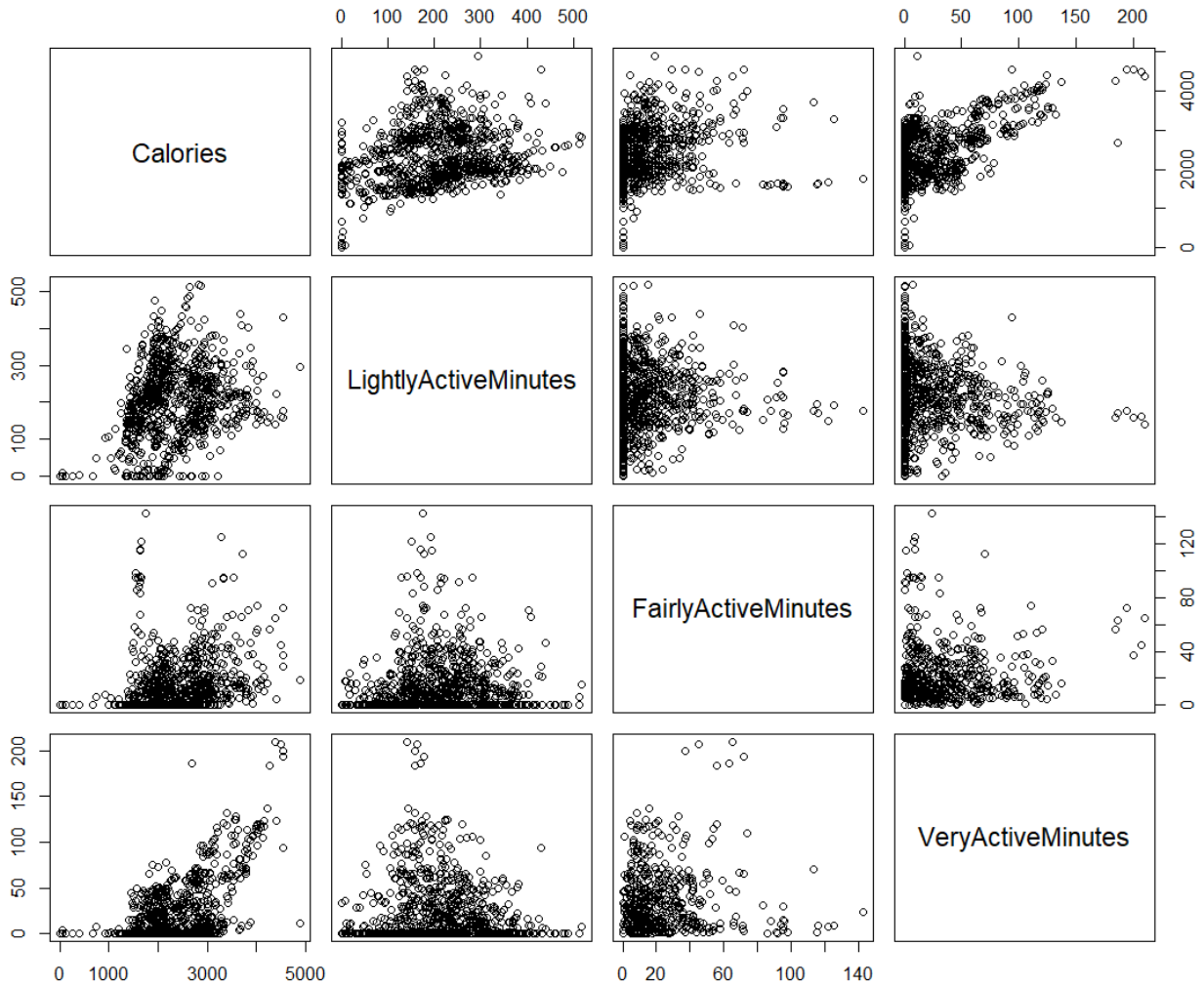


Activity by Time of Day



Important Trends




Calories
LightlyActiveMinutes
FairlyActiveMinutes
VeryActiveMinutes
0 100 200 300 400 500
0 50 100 150 200
0 2000 4000
0 100 300 500
0 40 80 120
0 50 100 150 200
0 1000 3000 5000
0 20 60 100 140

Conclusions and Recommendations

From summary table and plot, it appears that time of rigorous activity for users is small relative to total time of activity. It might be useful to include notification of app that will encourage users to increase their intense activity time once per day based on current level of rigorous time logged.

Users appears to do most of their rigorous exercise around 6 PM. We can add a notifications that reminds users around this time to add a rigorous workout.

Interestingly, total calories burned on average per day does not appear to influence time asleep. Nor does the total distance travelled and percentage of distance travelled under intense exertion appear to relate to total calories burned. This suggests that some users burning high calories might be doing more intense exercise with relatively little distance covered. It would be interesting to add functionality to the app that could give users more motivation for non-running or walking forms (such as weightlifting). (see circled points in combo plot of calories and active distance).

Overall trends indicate a positive correlation between lightly active minutes and calories and a stronger positive correlations between highly active minutes and calories.

Questions

??