Scenario

Bellabeat, a high-tech manufacturer of health-focused products for women, has asked me to focus on one of their products and analyze smart device data to gain insight into how consumers are using their smart devices. Urška Sršen, cofounder and Chief Creative Officer of Bellabeat, believes that analyzing smart device fitness data could help unlock new growth opportunities for the company. Any insights discovered during this analysis process could help guide marketing strategy for the company, so my analysis will be presented to the Bellabeat executive team along with any of my high-level recommendations for Bellabeat’s marketing strategy.

Questions

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?

Report Summary

1. Clear summary of the business task

2. Description of all data source used

3. Documentation of cleaning/editions on data

4. Summary of the analysis

5. Supporting visualization and key finding

6. Top high-level content recommendations based on analysis

Key Tasks

1. Identify the business task; analyze how people are using smart devices.

2. Consider the key stakeholders; give high-level recommendation for how these trends can improve **Bellabeat marketing strategy** in order to benefit said key stakeholders.

I am choosing to focus on Bellabeat’s Leaf wellness tracker for this project. I say this because I have data for two of the three tracked categories for it, those being activity and sleep, as well as the reason that Bellabeat Leaf tracks the same data as Bellabeat Time.

Issues

-Data is shown on various files instead of being congregated

-Variability in daily, hourly, and minutely in data monitoring and tracking

-No stress data is available in the dataset

Analysis

Data provided has been taken out of a 2-month span from 03/12/16 to 05/12/16. It was provided by Amazon Mechanical Turk and the submission of personal tracker data was consented to by thirty eligible users. Find it [here](https://www.kaggle.com/arashnic/fitbit).

Steps Taken Pre-Analysis

* Pick the data I will be using. In this study I will be using;
  1. dailyActivity\_merged.csv
  2. sleepDay\_Merged.csv
  3. hourlySteps\_merged.csv

I am using these datasets for their large sample size.

* My chosen workspace will be R for data cleaning, data analyzing, data visualization, and any insights. I will be doing most, if not all of my written report on Microsoft Word.

The first step was to load libraries in. For this project, I used ggplot2, tidyverse, here, skimr, janitor, lubridate, ggrepel, and dplyr.

My next step was to import my files and rename them to something easy to remember. I went with daily\_activity, hourly\_steps, and daily\_sleep.

I then previewed my data and checked the summaries of each column. Once this was confirmed, I checked to see how many unique users were in each dataset (done using n\_unique()).

Checking for duplicates was next on the to-do list. After using sum(duplicated()) to take a look, I removed the duplicates and checked for them again to confirm there were no more.

All columns had their format changed to lower case to ensure they were all in the correct syntax and format for merging.

Date/time were then formatted to be the same for all datasets.

I checked the datasets again before touching on the hourly\_steps sheet.

I merged daily\_activity and daily\_sleep to check if there was correlation between the two, as well as to prepare for the next step.

In order to classify the users, I decided to go with the daily number of steps they took per day. This can be broken down as follows;

* Sedentary: Less than 5000 steps per day.
* Lightly active: Between 5000 and 7499 steps per day.
* Fairly active: Between 7500 and 9999 steps per day.
* Very active: More than 10000 steps per day.

I made this categorization standard based on the following article: <https://www.verywellfit.com/how-many-pedometer-steps-per-day-are-enough-3432827>

I also decided not to make another category for “Highly active” and just use >10000 steps as my benchmark because the difference was less than in other categories (2500 steps compared to 5000) as well as to better display on the final charts.

First thing was to calculate the daily steps average by each user.

Next was to categorize users based on their daily average steps.

After that was creating a data frame with the percentages of each user type to better visualize the numbers.

After this was creating a graph. I went with a pie chart as it is very easily understood and I figured I could find a tutorial on how to make one more easily (I’m new to R, does it show?). \*For User Type Distribution graph, the total adds up to 101%; this is because of RStudio auto-rounding to the nearest whole number, the unrounded numbers only make up a 100% total.

Now that we have one graph done, we want to know how often users are utilizing their devices. All of this work goes to waste if people are not using their Bellabeat products, does it not? We will break the datasets down into three categories;

* high use – users who use their device between 21 and 31 days per month
* moderate use – users who use their device between 10 and 20 days per month
* low use – users who use their device between 1 and 10 days per month

First thing is to crate a new data frame; this will be done by grouping Id, calculating the number of days the device is used, and creating a new column with the same categorization as previously talked about.

Next is creating a percentage data frame to help with the visualization of results within our next graph, as well as ordering usage levels.

Now that everything is together, we can create our next graph. Another pie chart to aid in quick understanding of the data, nothing too fancy.

Conclusion and Recommendations

Based on my analysis, I can say that Bellabeat has a problem with keeping people using their product daily. I have included my recommendations and reasoning behind them below.

|  |  |
| --- | --- |
| Recommendation | Background |
| 1. Daily reminders/notifications/emails if item is not being worn/used | Based on what one of the pie charts show, half of all users do not use their device every day, meaning they are more likely to stop using the device altogether due to it not being a habitual part of their life. If Bellabeat can manage to involve these customers and get them to use their product daily, it will be more beneficial in terms of cost and brand image compared to having to continually focus more efforts on gaining new customers due to low overall retention. |
| 1. Change product design/marketing stance | To tie in with the first point, Bellabeat may have to change product design a little bit, to make it so users want to continuously keep their device on them at all times. As for the marketing side, aiming for the niche of consumers who are already leading active lives would help to boost data numbers for clientele if seeking to appease shareholders with numbers. Aiming for a wider spread of customers could generate more front-end revenue from sales, but that would require more in-depth research into the consumer base, as well as figuring out how to make products appeal to as many people as possible. Associating the item with the ultra wealthy may do this, or possibly brand deals with influential social media figures, in order to convince the public that it is a product they need, much like how Apple marketed the iPhone in the earlier stages. |
| 1. Aim for increased activity to improve public opinion/client satisfaction | The other pie chart reveals that about 21 percent of users are “sedentary”, meaning they are not really moving much at all during the time they are wearing the device. Another 21 percent are only “lightly active”, meaning that they get under 7500 steps a day in during the time they are wearing their device. Both these categories may be bumped up in activity level with daily notifications on steps taken for the day, and the possible use of alerts if they are not reaching the desired activity threshold. It may be useful to explain to them the benefits of that goal, although doing so will require more research into the target’s life and mind, as intrinsic motivation (desire from within) is the only thing that will truly get them to change and we cannot force that, only help it develop. |

To end, if the customer lifetime value can be worked out by Bellabeat, the marketing team will have their overall goal to shoot for. It may be focusing on customer retention, customer acquisition, or a blend of both in a balanced manner. This is not some type of nefarious activity to be ashamed of; this is simply knowledge that will be needed in order to make smarter, more accurate investments in the company’s relationships. Spending too much money on wasteful things will just damage Bellabeat in the long run, so it is something that must be figured out first before pushing forward with anything else.