Ask:

What is the problem you are trying to solve?

The main problem that we are trying to solve is how people are using non-Bellabeat smart devices and how we can take that insight to apply to Bellabeats marketing strategy for their products.

How can your insights drive business decisions?

Depending on the insights we obtain from the analysis could indicate trends and patterns within the data on how consumers use smart devices. With this information we could then make business decisions based on what the average consumer prefers or what some competition has that the consumer likes.

Prepare:

Where is your data stored?

The data is originally from Kaggle, the data was then downloaded and stored locally on the computer. Data was later imported into SQL Server and Tableau.

How is the data organized? Is it in long or wide format?

The data is organized in long format for all data that is being used. Each file also separated with different details such as calories burned, time spent sleeping and weight progression.

Are there issues with bias or credibility in this data? Does your data ROCCC?

Roccc stands for reliable, original, comprehensive, current and cited.
Reliable - The data is taken from a third party source so information on how this was obtained is unknown.
Original - The data was taken by Amazon Mechanical Turk
Comprehensive - The data is comprehensive and well organized
Current - The data is not current as it ranges from 03/12/2016 to 05/12/2016
Cited - The only citation is from a survey via Amazon Mechanical Turk

How are you addressing Licensing, privacy, security, and accessibility?

The data comes from a public data set that anyone can download so licensing will not be an issue. No names or personal information is within the data so there is not a privacy concern either.

Deliverable: What data sources are used?

The data sources that we will be using is from the fitabase data from Amazon Mechanical Turk. Specifically we are using dailyActivity_merged, heartrate_seconds_merged_sleepDay_merged and weightLogInfo_merged

Process:

What tools are you choosing and why?

The tools that I am using will be SQL and Tableau. We will clean some of the data within SQL and check for inconsistencies then import to R to do further analysis.

Have you ensured your data's integrity?

The data is the same throughout the data set.

What steps have you taken to ensure that your data is clean?

The first step was to find the column types that were correct which they were, there was a small issue with uploading that is explained below in the data cleaning step. The next was to find if there were any duplicate values, since the data is in long format there are duplicate values. Lastly, I took a look to see if there were any null values, the only spot that had null values was in a column that we will not be using.

How can you verify that your data is clean and ready to analyze?

Taking small samples was one way to verify that the data was clean but also running functions in SQL and seeing the schema in BigQuery to make sure the types were correct.

Have you documented your cleaning process so you can review and share those results?

Yes the doc for my cleaning steps can be found here

Analyze:

How should you organize your data to perform analysis on it?

The data was organized into 3 different files, these files were then uploaded into Bigquery an SQL IDE. Once there, if a query was done that provided results we were looking for, that query was saved as a new table. This will later be saved as a new csv file for visualization purposes.

Has your data been properly formatted?

All columns and data were in the proper formats, checks were done in order to make sure that there were not null values or complete duplicate values.

What surprises did you discover in the data?

Some surprises that I discovered from the data was that users spend a very large amount of the time sedentary but even more so users are very active more then they are fairly active. Something else that was surprising was that usually on the weekends users spend 30-60 minutes in bed from sleeping.

What trends or relationships did you find in the data?

- Some trends I saw was that users are more active during the weekend and sleep more during the weekend.
- Also that the average amount of sleep is around 7 hours
- Users tend to start the week off with higher numbers of activity but as the week goes on it reduces but picks back up during the weekend.

How will these insights help answer your business questions?

The insights will help answer the business question on how users are using non-Bellabeat products and how we can associate that to bellabeat products. We can see that most of the users tend to not be very active in fact most of the time they are sedentary. This will allow us to make some thoughts to market the products.

Share:

Were you able to answer the business questions?

Yes, I believe that the business questions were answered. There are some information that would be better to have and I think that this would need further analysis as well.

What story does your data tell?

One thing that it shows is the average usage of the smart devices, also the days that users are the most active. We see that most users are not using the devices for very active workouts, most users are sedentary or lightly active. This shows that users may not be that active while wearing the smart devices. We also see the sleep patterns of the average user. I think one thing that it shows is that maybe the market is steering more towards designing the smart devices for those users.

How do your findings relate to your original question?

My findings were able to relate to the original question by showing how users use non bellabeat smart devices. This will be discussed further in the presentation.

Who is your audience? What is the best way to communicate with them?

The main audience will be Urska Srsen, Sando Mur, and the bellabeat marketing analytics team.

Act:

What is your final conclusion based on your analysis?

My final conclusion is that a majority of users perform either sedentary or light activity with their smart devices. We also see that most users only get more than 7 hours of sleep on average with the most being on the weekend.

How could your team and business apply your insights?

I believe that the insights can show ways to better improve the Bellabeat products and associate products. Since most users are generally lightly - Sedentary we can apply these findings to the market strategy.

What next Steps would you or your stakeholders take based on your findings?

One thing that would be good to get some more information on the activities that are being done. This way we can narrow it down more as light activity is vague and could potentially just mean walking in a store. I believe also that the stakeholders can see trends for users sleep and activity which we can then use for are marketing.

Is there additional data you could use to expand on your findings?

- What activities are done for each actual activity grouping.
- More sleep user data.
- Style and fit of smartwear.
- Average age range, and gender of user.

Suggestions:

1. Since most users are not heavy active users and in fact spend most of the time with light or sedentary activities we can focus on building the users health. We can do this by following Bellabeats design now with stylized smartwear but add abilities to make the user try to perform and have a healthier lifestyle.This can include notifications letting the user know when they should be sleeping or going out for activities. As well this can include combining the coaching app as well for real time updates.
2. Partnering with other health lifestyle brands would allow Bellabeat to have users be rewarded for a healthy lifestyle. As the user spends more time being active and following this lifestyle they would receive discounts or coupons for those partner stores. This would help users stay encouraged as well as motivated.
3. While healthy activity is important, so is a healthy sleeping cycle. Either a new bed friendly product could be created or the Ivy technology could be updated with an accessory to encourage use while sleeping as the amount of users for sleep data was lower. This will also help track sleep time and notify users where they need to improve on to get a good night's rest.