**Objective**

Bellabeat is a tech manufacturer of women’s wellness products. They have developed wearable smart devices which record and measure various biometrics of the user. In order to gain further insight into the usage of these types of devices, publicly available data collected from Fitbit users will be analyzed to find key insights that could drive marketing strategy for Bellabeat. The objective is to identify opportunity for growth among Bellabeats users by utilizing trends observed from Fitbit usage data.

**Key Stakeholders**

Urška Sršen: Bellabeat cofounder and chief creative officer

Sando Mur: Bellabeat cofounder and member of the executive team

Bellabeat Marketing Analytics team

**Questions**

What are some observable trends in smart device usage?

How can these trends be applied to Bellabeat customers?

How can Bellabeat marketing strategy be oriented to these trends?

**Data**

FitBit Fitness Tracker Data (CC0: Public Domain, dataset made available through Mobius)

The data contains various personal data of 30 Fitbit users who consented to the submission of fitness tracker data. Data includes physical activity, heart rate, and sleep monitoring. All data collected was anonymous.

Limitations to the data include the following:

* Small sample size. Data from only 30 Fitbit users was collected.
* No demographic information accompanies the data. We do not know the users age, gender, geographic location, economic status, country of citizenship, race, or ethnicity. The possibility exists, therefore, that this data is biased towards specific demographics.
* Data was only collected for a few months. It would not be unreasonable to think that the time of year may factor into some of the data collected. Similarly to the small sample size, a longer window of collection may have provided more reliable data on usage.
* Lastly, the data is also somewhat outdated. It was collected in the Spring of 2016. Smart device usage and public opinion may have changed in the intervening years.

**Data Preparation**

I focused mainly on data collected on the users’ daily sleep and activity level.

Data collected for Daily Activity included: ID, ActivityDate, TotalSteps, TotalDistance, TrackerDistance, VeryActiveDistance, ModeratelyActiveDistance, LightlyActiveDistance, VeryActiveMinutes, FairlyActiveMinutes, LightlyActiveMinutes, SedentaryMinutes, and Calories. Data was organized by user ID and the Date recorded.

Data regarding sleep included ID, SleepDay, TotalSleepRecorded,TotalMinutesAsleep,TotalTimeInBed.

Data was first uploaded to Google Sheets for preliminary cleaning. Any duplicate data and white space was trimmed. Additionally, as we were looking at daily data, time stamps were removed as irrelevant. I also inserted a new column to each data set identifying the day of the week the data was collected. Lastly, using the concatenate function, I created a SpecialID column using the data recorded and ID of the user.

Next, the data was uploaded to Microsoft Azure Data Studio for further cleaning and exploration via SQL.

The **Sleep Data** revealed only 24 users recorded data out of the 30 who participated in the survey:

Select **Count** (Distinct Id)

From Bellabeat\_CaseStudy.dbo.Sleep

Group By (Id)

Additionally, 15 rows showed less than 2 hours of sleep recorded for that day. This to me indicates the device was recording naps but not night time sleep. This data was not removed from the data set but does suggest that at least a portion of users do not record their sleep data with a smart device.

Select TotalMinutesAsleep

From Bellabeat\_CaseStudy.dbo.Sleep

Where TotalMinutesAsleep < 120

Average amount of time asleep were determined for each user and each day of the week:

Select Id, **Avg** (TotalMinutesAsleep) As AvgUserAsleep

From Bellabeat\_CaseStudy.dbo.Sleep

Group By Id

&

Select DayofWeek, **Avg**(TotalMinutesAsleep)as AvgMinutesAsleep

From Bellabeat\_CaseStudy.dbo.Sleep

Group By DayofWeek

No data was found containing 0 hours of sleep recorded:

Delete From Bellabeat\_CaseStudy.dbo.Sleep

Where TotalMinutesAsleep = 0

From the **Activity** data set; rows indicating 0 total distance recorded for the day were dropped. This to me indicated days where the device was not actually worn:

Delete From Bellabeat\_CaseStudy.dbo.Activity

Where TotalDistance = 0

Another column was added to the dataset for the total number of minutes the device was worn for a particular day:

Alter TABLE Bellabeat\_CaseStudy.dbo.Activity

Add TotalMinutes Int

Update Bellabeat\_CaseStudy.dbo.Activity

Set TotalMinutes = ( VeryActiveMinutes + FairlyActiveMinutes + LightlyActiveMinutes + SedentaryMinutes)

Unused columns were dropped from the dataset:

Alter Table Bellabeat\_CaseStudy.dbo.Activity

Drop Column TrackerDistance, VeryActiveDistance, ModeratelyActiveDistance, LightActiveDistance

Similarly to the sleep data, average number of minutes worn per user and per day of the week as well as average number of steps per day of the week were found:

Select **Avg**(TotalSteps) as AvgSteps, DayofWeek

From Bellabeat\_CaseStudy.dbo.Activity

Group by DayofWeek

&

Select Id, **Avg**(TotalMinutes) as AvgTimeWorn

From Bellabeat\_CaseStudy.dbo.Activity

Group By Id

&

Select DayofWeek, **Avg**(TotalMinutes) as AvgTimeWorn

From Bellabeat\_CaseStudy.dbo.Activity

Group By DayofWeek

Finally, data was joined between the two data sets utilizing the SpecialID column as the relationship:

Select \*

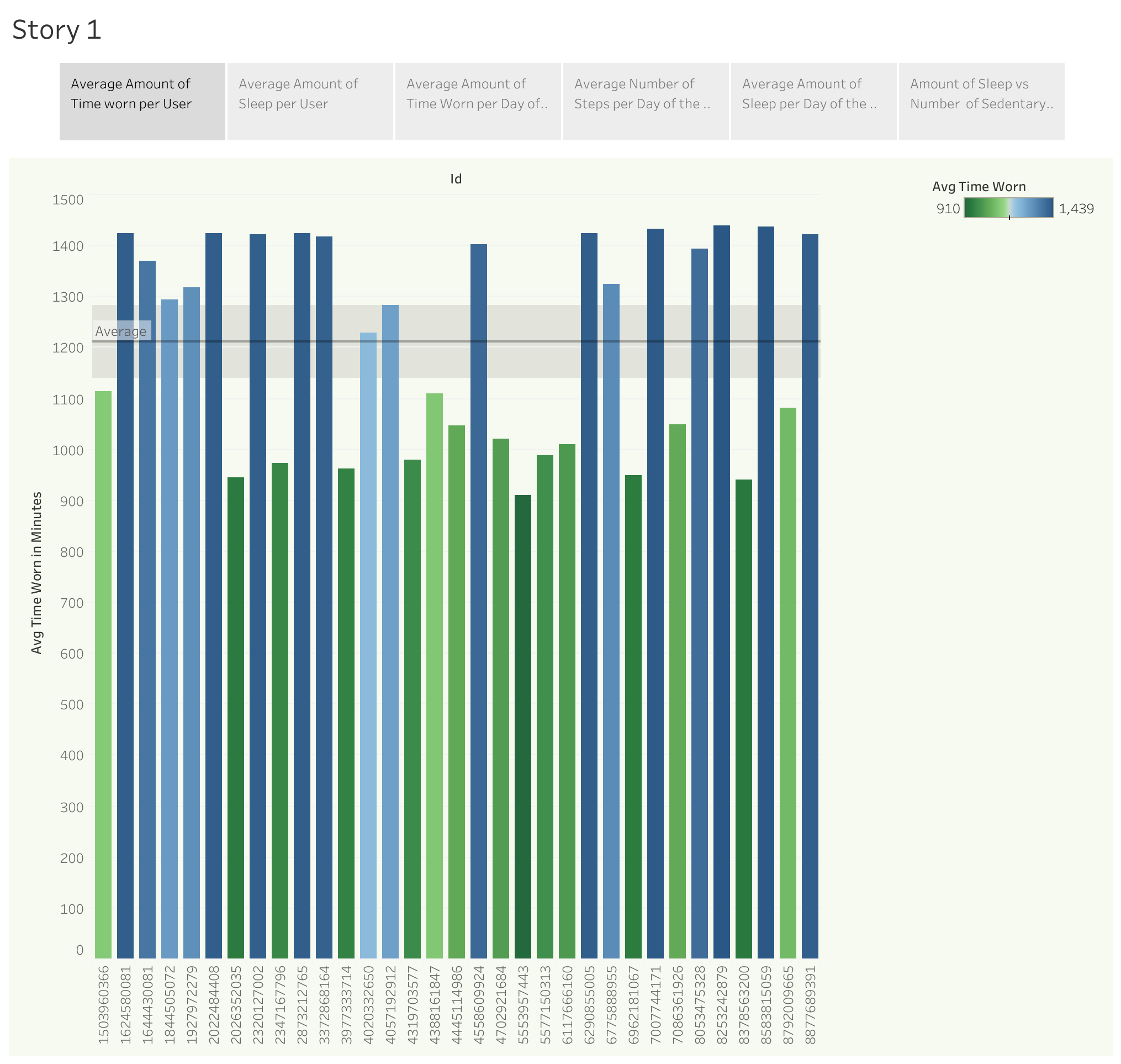
From Bellabeat\_CaseStudy.dbo.Activity

Join Bellabeat\_CaseStudy.dbo.Sleep ON Bellabeat\_CaseStudy.dbo.Sleep.Special\_ID = Bellabeat\_CaseStudy.dbo.Activity.SpecialID

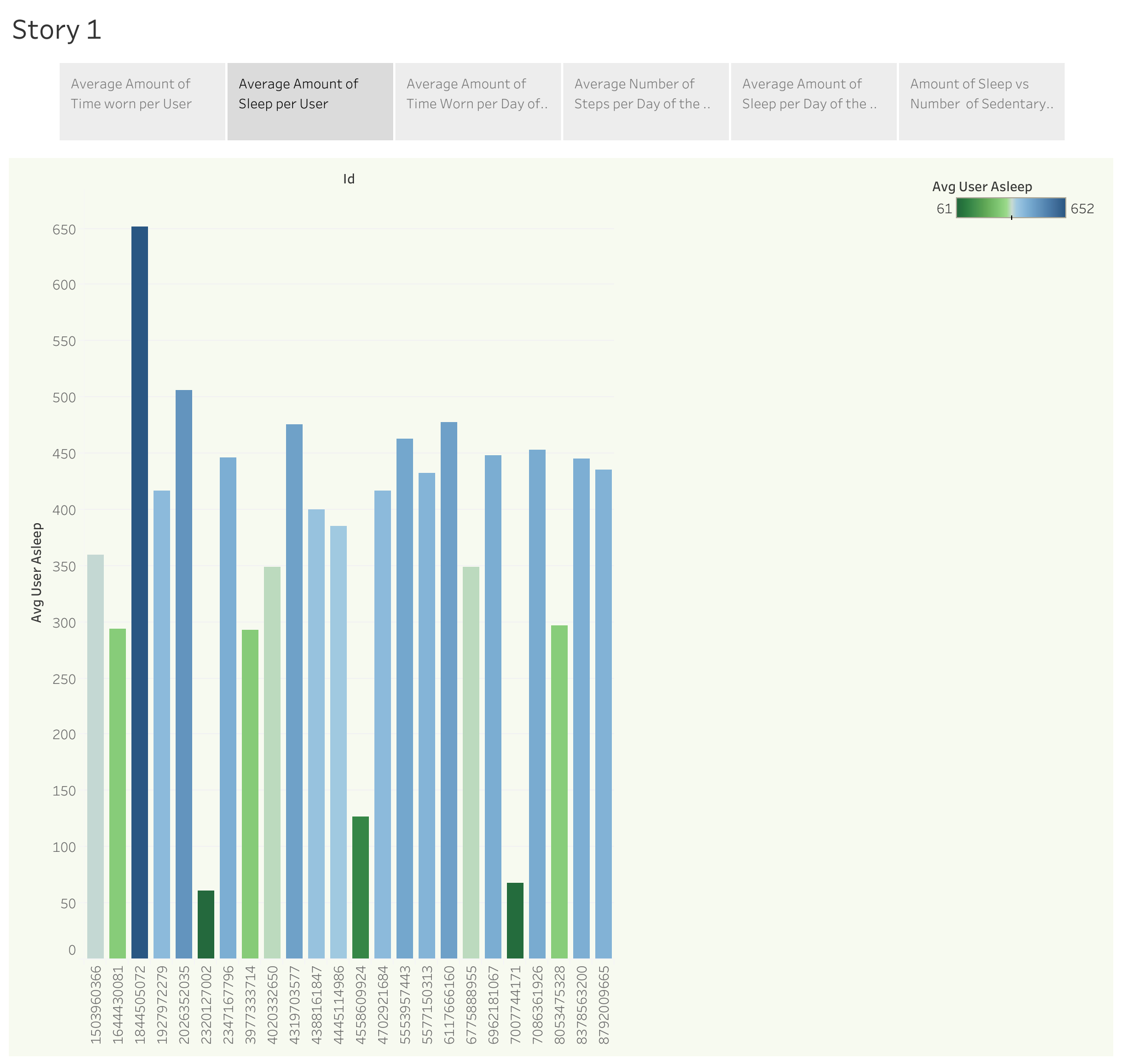
**Visualizing and Analyzing Data**

Data visualizations were created utilizing Tableau Public software.

The first trend identified was the amount of time users wore their devices. They fell into roughly two camps; those that wore their device all day and night, and those that only wore their device for part of the day. The green users below wore their devices for between 950-1100 minutes or about 15 ½ to 18 hours.



Only 24 users recorded data out of the 30 participants with a very wide range of amount of sleep recorded.



No significant correlation was found between the average amount of time worn for each day of the week.

However, users were most active, as measured by number of steps taken on Tuesdays and Saturdays. Sundays and Fridays were the least active days.