Bellabeat Case Study: Data-Driven Insights for Wellness Innovation

Table of content

Introduction

Bellabeat, a leading wellness technology company, empowers women worldwide with beautifully designed smart devices that track health and fitness metrics. To stay ahead in the competitive wellness industry, Bellabeat sought to leverage data analytics to uncover trends, understand customer behavior, and drive strategic decisions. This case study explores the analysis of fitness data, the insights derived, and the actionable recommendations for Bellabeat's growth.

My Role: As a junior data analyst on the marketing analyst team at Bellabeat, my primary responsibility is to analyze smart device usage data to uncover insights into how consumers interact with one of Bellabeat's products. These insights will play a crucial role in shaping the company's marketing strategy. I will present my findings and high-level recommendations to the Bellabeat executive team to inform their decision-making process.

Objective

The goal of this analysis is to:

1. Identify trends in smart device usage among consumers.
2. Understand how these trends can influence Bellabeat's marketing and product strategies.
3. Provide data-driven recommendations to guide Bellabeat's business decisions.

Phase 1: Ask – A Clear Summary of the Business Task

Business Task:
Analyze smart device usage data to understand how consumers interact with non-Bellabeat smart devices. Use these insights to identify trends and apply them to one Bellabeat product, ultimately informing and enhancing Bellabeat's marketing strategy.

Key Questions for Analysis:

1. What are the current trends in smart device usage?
2. How can these trends be applied to Bellabeat's customer base?
3. How can these trends shape and improve Bellabeat's marketing strategy?

Phase 2: Prepare – Description of Data Sources

Data Source:
For this analysis, we used **FitBit fitness tracker data**, a publicly available dataset sourced from **Kaggle** and provided by **Mobius**.

About the Dataset:

- The dataset was collected from 33 eligible FitBit users who consented to share their personal tracker data.
- Data was gathered through a distributed survey on **Amazon Mechanical Turk** between **March 12, 2016, and May 12, 2016**.
- The dataset includes minute-level details on physical activity, heart rate, and sleep monitoring.

Scope of Analysis:
For this case study, I focused on **daily and hourly data** to identify broader trends, rather than delving into minute-level user performance. To facilitate the analysis, I combined relevant tables to create a comprehensive dataset for exploration

---

Tools Used

- **Data Cleaning & Preparation:** Microsoft Excel
- **Data Analysis:** SQL (Microsoft SQL Server Management Studio)
- **Data Visualization:** Power BI (for creating interactive dashboards)

Phase 3: Process - Documentation of any cleaning or manipulation of data

For my analysis, I will use SQL to prepare, process and analyse data available and Power BI to create visualizations to share my findings with the stakeholders.

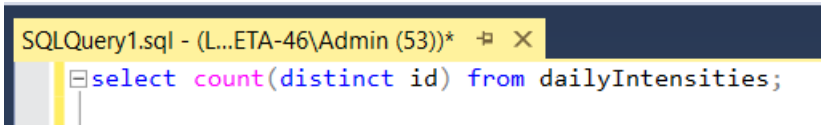
Importing datasets : FitBit Fitness Tracker Data.

Exploring Data:

1. Checking for unique number of users in each table.

select count(distinct id) as total_users from dailyActivity;

	total_users
1	33



	(No column name)
1	33

Select count(distinct id) from dailySteps;
Output:
Count(distint(id)) >> 33

Select count(distinct id) From hourlyintensities;
Output:
Count(distint(id)) >> 33

Select count(distinct id) From hourlycalories;
Output:
Count(distint(id)) >> 33

Select count(distinct id) From hourlysteps;
Output:
Count(distint(id)) >> 33

select count(distinct id) as total_users from sleepDay;

	total_users
1	24

select count(distinct id)as total_users from heartrate_seconds;

	total_users
1	14

select count(distinct id)as total_users from weightLogInfo;

	total_users
1	8

As you can see, dailyActivity table has the all the user activity logged. So, we will be working on this table as well as sleepDay table that most user activity recorded.

Duplicates Checking for any duplicates records in the tables.

```
SELECT count(duplicates) as total_duplicates FROM
   (SELECT id, sleepday, TotalMinutesAsleep, COUNT(*) as duplicates
   FROM sleepday
   GROUP BY id , sleepday , TotalMinutesAsleep
   HAVING COUNT(*) > 1) as X;
```

Results Messages

	total_duplicates
1	3

#This shows there are 3 duplicate records in sleepday table.

```
Select count(*) from
        (select id, activityDay, Calories,count(*) from dailycalories
        group by id, activityDay, Calories
        having count(*) > 1) X;
```

Output:
Count(*)>> 0

```
Select count(*) from
        (select id, activityDay, count(*) from dailyintensities
        group by id, activityDay
        having count(*) > 1) X;
```

Output:
Count(*)>> 0

```
Select count(*) from
        (select id, activityDay,steptotal,count(*) from dailysteps
        group by id, activityDay,steptotal
        having count(*) > 1) X;
```

Output:
Count(*)>> 0

```
Select count(*) from
        (select id, activityHour, Calories,count(*) from hourlycalories
        group by id, activityHour, Calories
        having count(*) > 1) X;
```

Output:
Count(*)>> 0

```
Select count(*) from
        (select id, activityHour, count(*) from hourlyintensities
        group by id, activityHour
        having count(*) > 1) X;
```

Output:
Count(*)>> 0

```
Select count(*) from
        (select id, activityHour,steptotal,count(*) from hourlysteps
```

```
        group by id, activityHour,steptotal
        having count(*) > 1) X;
```

Output:
Count(*)>> 0

Remove duplicates from sleepDay and create a new table sleepDay_New:

```
SELECT DISTINCT
   id, sleepday, totalsleeprecords, totalminutesasleep, totaltimeinbed
INTO sleepDay_New FROM sleepday;
```

Check whether the new sleepDay_New table has any duplicates:

```
SELECT count(duplicates) as total_duplicates FROM
   (SELECT id, sleepday, TotalMinutesAsleep, COUNT(*) as duplicates
   FROM sleepday_new
   GROUP BY id , sleepday , TotalMinutesAsleep
   HAVING COUNT(*) > 1) as X;
```

	total_duplicates
1	0

Joining tables

Creating table Combined_Sleep_Activity from dailyactivity table and columns from sleepday_new table joined by id and date.
Extracting date from ActivityDate and extracting the day of the week from ActivityDate and adding it as a column ActivityDay. Explicitly converting the datatypes of TotalSteps and Calories to integer from nvarchar to perform aggregation. Also, adding a column for user ActivityLevel based on number of steps taken in a **day and CalorieLevel based on calories burnt each day.**

```
select a.id, CAST(ActivityDate AS DATE) as Activity_Date, DATENAME(WEEKDAY,
ActivityDate) as ActivityDay,
        TotalDistance,VeryActiveDistance,ModeratelyActiveDistance, LightActiveDistance,
SedentaryActiveDistance,
        VeryActiveMinutes,FairlyActiveMinutes as
ModeratelyActiveMinutes,LightlyActiveMinutes,SedentaryMinutes,
        CAST(TotalSteps as int) as TotalSteps,
        CASE
                WHEN TotalSteps < 500  THEN 'Sedentary'
      WHEN TotalSteps BETWEEN 500 AND 4999  THEN 'Light Active'
      WHEN TotalSteps BETWEEN 5000 AND 14999 THEN 'Moderately Active'
      WHEN TotalSteps >= 15000  THEN 'Very Active'
        END as ActivityLevel, CAST(Calories as int) as Calories,
        CASE
      WHEN Calories < 1800 THEN 'Low'
      WHEN Calories BETWEEN 1800 AND 2500 THEN 'Medium'
```

```
        WHEN Calories > 2500 THEN 'High'
        ELSE 'Unclassified'
    END AS CalorieLevel,
          s.totalsleeprecords, s.totalminutesasleep,s.totaltimeinbed
INTO Combined_Sleep_Activity
from dailyactivity a
left join sleepDay_New s on a.id=s.id
             and a.ActivityDate = s.sleepday;
```

Checking for null values in the combined table:

```
select count(id) as count_null from Combined_Sleep_Activity
where totalsleeprecord is null
```

	count_null
1	530

```
select count(id) as count_null from Combined_Sleep_Activity
where totalminutesasleep is null
```

	count_null
1	530

```
select count(id) as count_null from Combined_Sleep_Activity
where totaltimeinbed is null
```

	count_null
1	530

Replacing the null/missing values with 0 as no sleep activity was recorded by the user.

```
UPDATE Combined_Sleep_Activity
SET totalsleeprecords = 0,
    totalminutesasleep = 0,
    totaltimeinbed = 0
WHERE totalsleeprecords IS NULL
      OR totalminutesasleep IS NULL
      OR totaltimeinbed IS NULL;
```

Verifying whether all null/missing values have been replaced

```
select count(id) as count_null from Combined_Sleep_Activity
where totalsleeprecord is null
```

	count_null
1	0

```
select count(id) as count_null from Combined_Sleep_Activity
where totalminutesasleep is null
```

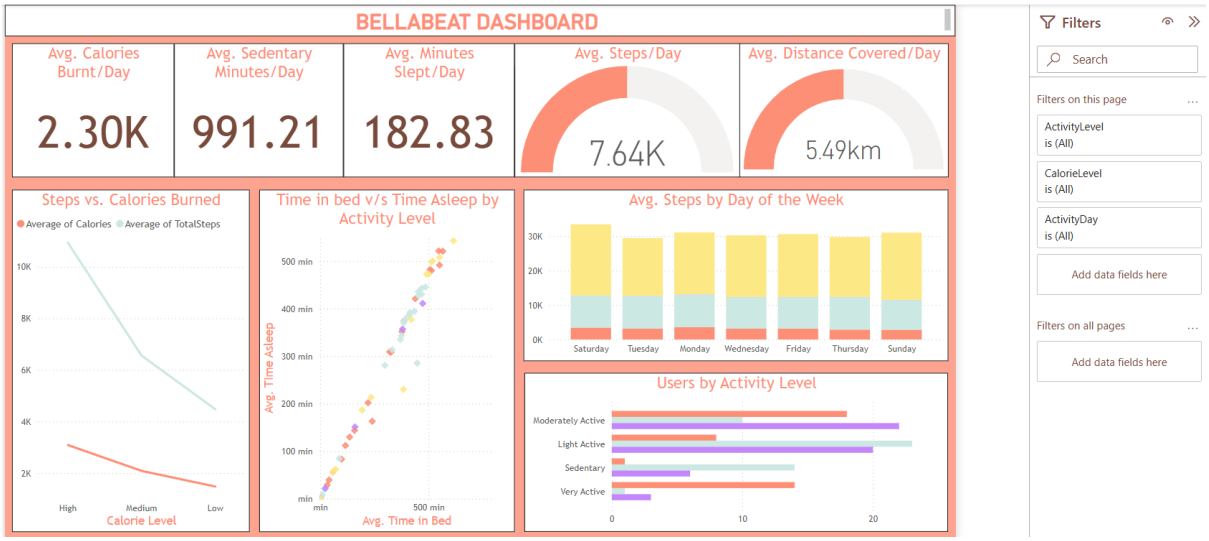
```
select count(id) as count_null from Combined_Sleep_Activity
where totaltimeinbed is null
```

Exported a new table DailyActivity_Sleep from MySQL and created a dashboard using Power BI.

Phase 4: Analyze - A summary of your analysis

AND

Phase 5: Share - Supporting visualizations and key findings



Analysis & Insights

Average steps taken: 7.64K per day with an average distance of 5.49 km.

Average calories burned: 2.30K per day

Most of the participants are lightly active.

Sleep efficiency is low, with users spending 50 extra minutes in bed.

Weekends show a 20% increase in activity levels compared to weekdays.

Phase 6: Act - High-level recommendations based on the analysis.

Established in 2014, Bellabeat is a company that pioneered one of the first wearable devices tailored specifically for women. Over the years, it has expanded its offerings to include a range of digital products aimed at monitoring and enhancing women's health.
Upon analyzing the FitBit Fitness data, which includes metrics such as activity levels, step count, calories burned, exercise intensities, and sleep patterns, several trends emerged:

- There is a positive correlation between the number of steps taken and the number of calories burned.
- A positive correlation also exists between the minutes slept and the total time spent in bed.

However, a notable concern is the average sedentary time, which exceeds 16 hours per day, highlighting a significant issue.
It's important to note that the dataset has a limited sample size and may be biased, as it lacks demographic information about the users.

My Recommendations:

1. **Sedentary Behavior Reminders:**
   Introduce notifications or reminders to encourage users to reduce prolonged periods of inactivity. For example, the app could prompt users to stand up, stretch, or take a short walk after a certain period of sitting. This can help combat the average sedentary time of over 16 hours per day and promote a more active lifestyle.
2. **Sleep Hygiene Tips:**
   Incorporate a feature that provides users with actionable tips to improve sleep efficiency, such as maintaining a consistent sleep schedule, avoiding screens before bed, or creating a relaxing bedtime routine. These tips can be delivered through notifications or within the app to help users achieve the recommended 8 hours of sleep.
3. **Weekend Fitness Challenges:**
   Launch engaging weekend fitness challenges, such as step-count competitions or activity-based goals, to motivate users to stay active during weekends. These challenges can foster a sense of community, increase user engagement, and encourage consistent physical activity throughout the week.

Conclusion

This case study demonstrates the power of data analytics in driving business decisions for Bellabeat. By leveraging insights from fitness data and visualizing them through Power BI, we've identified key opportunities to enhance user engagement, improve product offerings, and strengthen marketing strategies. With these actionable recommendations, Bellabeat can continue to innovate and lead in the wellness technology industry.

Thank you for reading/reviewing my Bellabeat Case Study!