```
# This Python 3 environment comes with many helpful analytics libraries installed
# It is defined by the kaggle/python Docker image: https://github.com/kaggle/docker-python
# For example, here's several helpful packages to load

import numpy as np # linear algebra
import pandas as pd # data processing, CSV file I/O (e.g. pd.read_csv)

# Input data files are available in the read-only "../input/" directory
# For example, running this (by clicking run or pressing Shift+Enter) will list all files under the input directory

import os
for dirname, _, filenames in os.walk('/kaggle/input'):
    for filename in filenames:
        print(os.path.join(dirname, filename))

# You can write up to 20GB to the current directory (/kaggle/working/) that gets preserved as output when you create a version using "Save & Run All"
# You can also write temporary files to /kaggle/temp/, but they won't be saved outside of the current session
```

```
/kaggle/input/fitbit/Fitabase Data 4.12.16-5.12.16/minuteCaloriesNarrow_merged.csv
/kaggle/input/fitbit/Fitabase Data 4.12.16-5.12.16/weightLogInfo_merged.csv
/kaggle/input/fitbit/Fitabase Data 4.12.16-5.12.16/sleepDay_merged.csv
/kaggle/input/fitbit/Fitabase Data 4.12.16-5.12.16/dailyIntensities_merged.csv
/kaggle/input/fitbit/Fitabase Data 4.12.16-5.12.16/minuteIntensitiesWide_merged.csv
/kaggle/input/fitbit/Fitabase Data 4.12.16-5.12.16/minuteMETsNarrow_merged.csv
/kaggle/input/fitbit/Fitabase Data 4.12.16-5.12.16/dailyCalories_merged.csv
/kaggle/input/fitbit/Fitabase Data 4.12.16-5.12.16/hourlyCalories_merged.csv
/kaggle/input/fitbit/Fitabase Data 4.12.16-5.12.16/heartrate_seconds_merged.csv
/kaggle/input/fitbit/Fitabase Data 4.12.16-5.12.16/minuteSleep_merged.csv
/kaggle/input/fitbit/Fitabase Data 4.12.16-5.12.16/hourlyIntensities_merged.csv
/kaggle/input/fitbit/Fitabase Data 4.12.16-5.12.16/hourlySteps_merged.csv
/kaggle/input/fitbit/Fitabase Data 4.12.16-5.12.16/minuteStepsNarrow_merged.csv
/kaggle/input/fitbit/Fitabase Data 4.12.16-5.12.16/dailySteps_merged.csv
/kaggle/input/fitbit/Fitabase Data 4.12.16-5.12.16/minuteStepsWide_merged.csv
/kaggle/input/fitbit/Fitabase Data 4.12.16-5.12.16/minuteIntensitiesNarrow_merged.csv
/kaggle/input/fitbit/Fitabase Data 4.12.16-5.12.16/minuteCaloriesWide_merged.csv
/kaggle/input/fitbit/Fitabase Data 4.12.16-5.12.16/dailyActivity_merged.csv
```

Smart Device Usage Insights for Bellabeat

Introduction

Smart devices have become an integral part of our daily lives, providing users with valuable insights into their health and well-being. In this analysis, we delve into the usage patterns of non-Bellabeat smart devices, aiming to uncover trends and behaviors that can inform the strategic decisions of Bellabeat, a leading provider of smart health and wellness products.

Objective: Our goal is to gain meaningful insights into how consumers engage with non-Bellabeat smart devices. By conducting analysis on how users interact with non-Bellabeat technology we aim to identify opportunities for improvement and innovation within the Bellabeat product line.

Key Questions: Analysis is guided by the following questions:

1. **Usage Patterns:** How do consumers interact with their smart devices on a daily basis? What are the typical usage patterns and activities recorded by these devices?
2. **Features and Preferences:** What features and functionalities are popular among users of non-Bellabeat smart devices? Are there specific preferences or behaviors that stand out?

Application to Bellabeat: The insights gained from this analysis will be applied to inform the development and enhancement of Bellabeat's smart health and wellness products. By understanding consumer preferences and behaviors, we can tailor our products to meet the evolving needs of our target audience.

Throughout this analysis, we will leverage data visualization and statistical analysis to present compelling findings.

Analysis: Usage Patters

Section Goal

In this section, we explore the distribution of daily activity levels captured by the Fitbit dataset. Our goal is to visualize how people allocate their time across different activity categories throughout the day. The dataset includes information on Very Active, Fairly Active, Lightly Active, and Sedentary minutes.

Visualization: Activity Levels Throughout the Day

To gain insights into the distribution of activity levels, we create a bar chart that illustrates the proportion of time spent in each activity category. The chart provides a visual representation of how individuals engage in various levels of activity during a typical day.

Let's proceed with generating the chart using Plotly Express

In [2]:

```python
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import plotly.express as px
import plotly.graph_objects as go

data = pd.read_csv('/kaggle/input/fitbit/Fitabase Data 4.12.16-5.12.16/dailyActivity_merged.csv')

# Convert 'ActivityDate' to datetime format
data['ActivityDate'] = pd.to_datetime(data['ActivityDate'], format='%m/%d/%Y')

# Extract day name
data['Day'] = data['ActivityDate'].dt.day_name()

# Create the plot using Plotly Express with color-friendly palette
fig = px.bar(data, x='Day', y=['VeryActiveMinutes', 'FairlyActiveMinutes', 'LightlyActiveMinutes'],
        color_discrete_map={
            'VeryActiveMinutes': '#1f78b4',  # Blue
            'FairlyActiveMinutes': '#33a02c',  # Green
            'LightlyActiveMinutes': '#e31a1c'  # Red
        },
        labels={'value': 'Minutes'},
        title='Daily Activity Breakdown',
        barmode='group',
        category_orders={'Day': ['Monday', 'Tuesday', 'Wednesday', 'Thursday', 'Friday', 'Saturday', 'Sunday']}
        )
fig.update_xaxes(tickangle=-45)
fig.show()
```

Key Insights:

1. Weekly Patterns

- **Very Active Minutes:** The data suggests that individuals tend to have higher levels of activity during the beginning of the week but slowly decreases over time where we can see varied levels of activity during the weekend. This varied level of activity could possibly be due to the work week tiring most users to be very active throughout the week
- **Fairly Active Minutes:** Fairly active minutes exhibit a similar pattern to very active minutes, with lower values on weekdays and a varying throughout the weekend
- **Lightly Active Minutes:** Lightly active minutes relatively vary throughout the week. The weekdays see a slight increase then consistent decrease from Monday To Friday. On weekends we see that Sundays exihbit the lowest amount of active minutes on average on any day of the week ### 2. Peak Activity Days
- **Tuesday and Wednesday:** These days stand out as having the highest overall activity levels. This trend supports the possible explanation that users exhibt a increase in activity during the beginning of the week.

3. Weekend Patterns

- **Saturday and Sunday:** Saturdays show a slight increase to the decreasing activity throughout the work week. As mentioned previously, On Sundays, users exhibit the lowest amount of activity on average than any other day of the week

---

Exploring Sedentary Behavior Throughout the Week

Sedentary Minutes Analysis

- **Average Percentage of Sedentary Minutes:**
  - Its important to observe how much users are sedentary (At rest) throughout the week in order to visualize a percentage of how long on average users are not interacting with their smart watches.
  - Lets investigate sedentary behavior by calculating the average percentage of sedentary minutes for each day of the week.

In [3]:

```python
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import plotly.express as px
import plotly.graph_objects as go

data = pd.read_csv('/kaggle/input/fitbit/Fitabase Data 4.12.16-5.12.16/dailyActivity_merged.csv')
# Convert 'ActivityDate' to datetime format
data['ActivityDate'] = pd.to_datetime(data['ActivityDate'])

# Extract the day of the week
data['DayOfWeek'] = data['ActivityDate'].dt.day_name()

# Calculate total sedentary minutes for each day
sedentary_minutes_per_day = data.groupby('DayOfWeek')['SedentaryMinutes'].sum()

# Calculate the total minutes for each day
activity_columns = ['VeryActiveMinutes', 'FairlyActiveMinutes', 'LightlyActiveMinutes']
total_minutes_per_day = data.groupby('DayOfWeek')[activity_columns].sum().sum(axis=1)

# Calculate the percentage of sedentary minutes out of the total minutes for each day
sedentary_percentage_per_day = (sedentary_minutes_per_day / total_minutes_per_day) * 100

# Plotting a pie chart for each day
colors = plt.cm.Paired(range(len(sedentary_percentage_per_day)))

plt.figure(figsize=(10, 6))
plt.pie(sedentary_percentage_per_day, labels=sedentary_percentage_per_day.index, autopct='%1.1f%%', startangle=90, colors=colors)
plt.title('Percentage of Sedentary Minutes for Each Day of the Week')
plt.show()
```
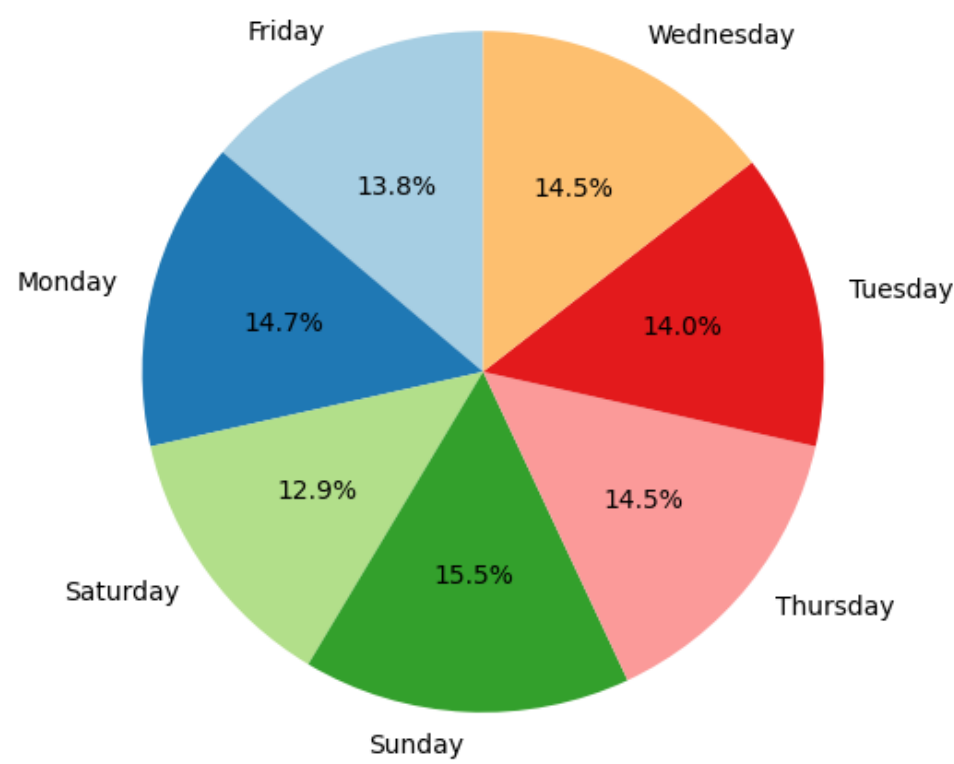
Percentage of Sedentary Minutes for Each Day of the Week

KeyFindings:

- **Monday to Friday Trends:**
  - On average, individuals tend to exhibit a consistent percetage throughout the day where they are at rest
  - Many of the weekdays show a higher and consistent percentage, possibly reflecting a more desk-bound start to the workweek.
- **Weekend Patterns:**
  - Fridays and Saturday display a decrease in sedentary minutes, potentially indicating a more active approach toward the end of the week.
  - Sundays demonstrate the largest increase in sedentary minutes as this could reflect the need to rest before the beginning of the work week

---

Usage Patters Explored

Examining how customers use their smart devices on a daily basis, our analysis shows a pattern of weekly user activity that is consistent. The data highlights customers' willingness to staying active by showcasing common usage patterns. What's interesting to note is that our results show specific examples of sedentary behavior, which are especially apparent at times that are anticipated to align with the workweek.

The patterns that have been noticed, which range from increased activity at the start of the workweek to decreasing activity as the weekend draws closer, highlight how dynamic user interactions are. The use of smart devices seems to fit into consumers' lives, fitting in with weekend leisure activities as well as job-related responsibilities. Designing features and functionalities that meet the varied demands of users throughout can be informed by an understanding of these minor patterns.

Analysis: Features and Prefrences

Section Goal

In this section, we delve deeper into user behavior by exploring the features that users engage with the most throughout the day. By analyzing how users use and log different health information throughout the week, we aim to identify patterns in user interactions with specific functionalities, shedding light on which features are particularly popular and actively utilized throughout the week

## Visualization: WeightLog Use Throughout the Day

To gain insights into the most popular features among users, lets create a pie chart that illustrates the proportion of users who utilize the Weight Log feature of the Fitbit Smart Watch and those who dont in order to see what percentage of the userbase utilizes this feature

In [4]:

```python
df = pd.read_csv('/kaggle/input/fitbit/Fitabase Data 4.12.16-5.12.16/weightLogInfo_merged.csv')
# Calculate the percentage of people who manually reported and who didn't
percentage_manual_reports = df['IsManualReport'].value_counts(normalize=True) * 100
```

```
# Plot the pie chart
plt.figure(figsize=(8, 8))
labels = ['Manually Reported', 'Not Manually Reported']
plt.pie(percentage_manual_reports, labels=labels, autopct='%1.1f%%', startangle=90, colors=plt.cm.Paired.colors)
plt.title('Percentage of People Reporting Weight Log (Manual vs. Not Manual)')
plt.show()
```

Percentage of People Reporting Weight Log (Manual vs. Not Manual)
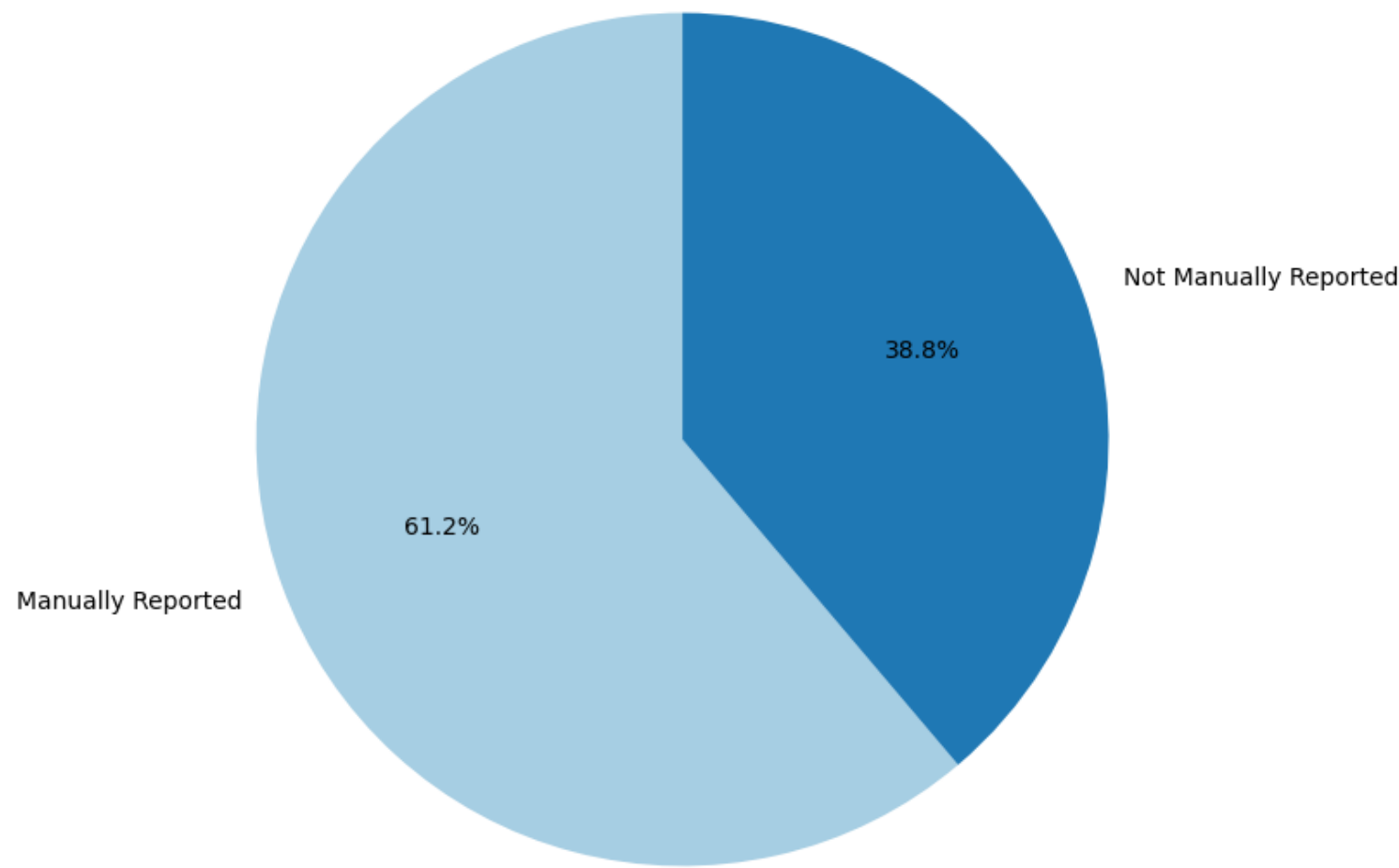
Key Findings:

1. **High Utilization:** The visual representation illustrates that a significant proportion of users are consistently engaging with the weight logging feature. This indicates a strong user commitment to monitoring and recording their weight regularly.
2. **User Preference for Manual Reporting:** The pie chart breakdown highlights that a considerable percentage of users actively choose to manually report their weight logs. This suggests a user preference for maintaining a hands-on approach to their health data, possibly driven by a desire for accuracy or a more personalized tracking experience.

Visulazation: Correlation Between Weight of Users and Usage of Watch Features

As mentioned above we can see that the weight log feature is a very popular among users. To expand upon that, we can determine if there is a correlation between weight of different users and how often they use the weight logging feature

User Categorization:

- **Weight Categories:**
  - Users were categorized into the following groups based on weight:
    - **Underweight**
    - **Normal**
    - **Overweight**
    - **Obese**
- **Active Minutes Categories:**
  - Users were categorized based on their very active minutes into the following groups:
    - **Low**
    - **Moderate**
    - **High**
    - **Very High**

In [5]:

```
import pandas as pd
import matplotlib.pyplot as plt

weight_logging_data = pd.read_csv('/kaggle/input/fitbit/Fitabase Data 4.12.16-5.12.16/weightLogInfo_merged.csv')

# Convert 'Date' to datetime
weight_logging_data['Date'] = pd.to_datetime(weight_logging_data['Date'], format='%m/%d/%Y %I:%M:%S %p')
# Categorize users based on weight
bins_weight = [0, 60, 80, 100, float('inf')]
labels_weight = ['Underweight', 'Normal', 'Overweight', 'Obese']
weight_logging_data['WeightCategory'] = pd.cut(weight_logging_data['WeightKg'], bins=bins_weight, labels=labels_weight)

# Count the number of weight logs in each category
weight_logs_count = weight_logging_data['WeightCategory'].value_counts()

# Plotting a pie chart
plt.figure(figsize=(8, 8))
plt.pie(weight_logs_count, labels=weight_logs_count.index, autopct='%1.1f%%', startangle=90)
plt.title('Weight Category Distribution and Logging Frequency')
plt.show()
```
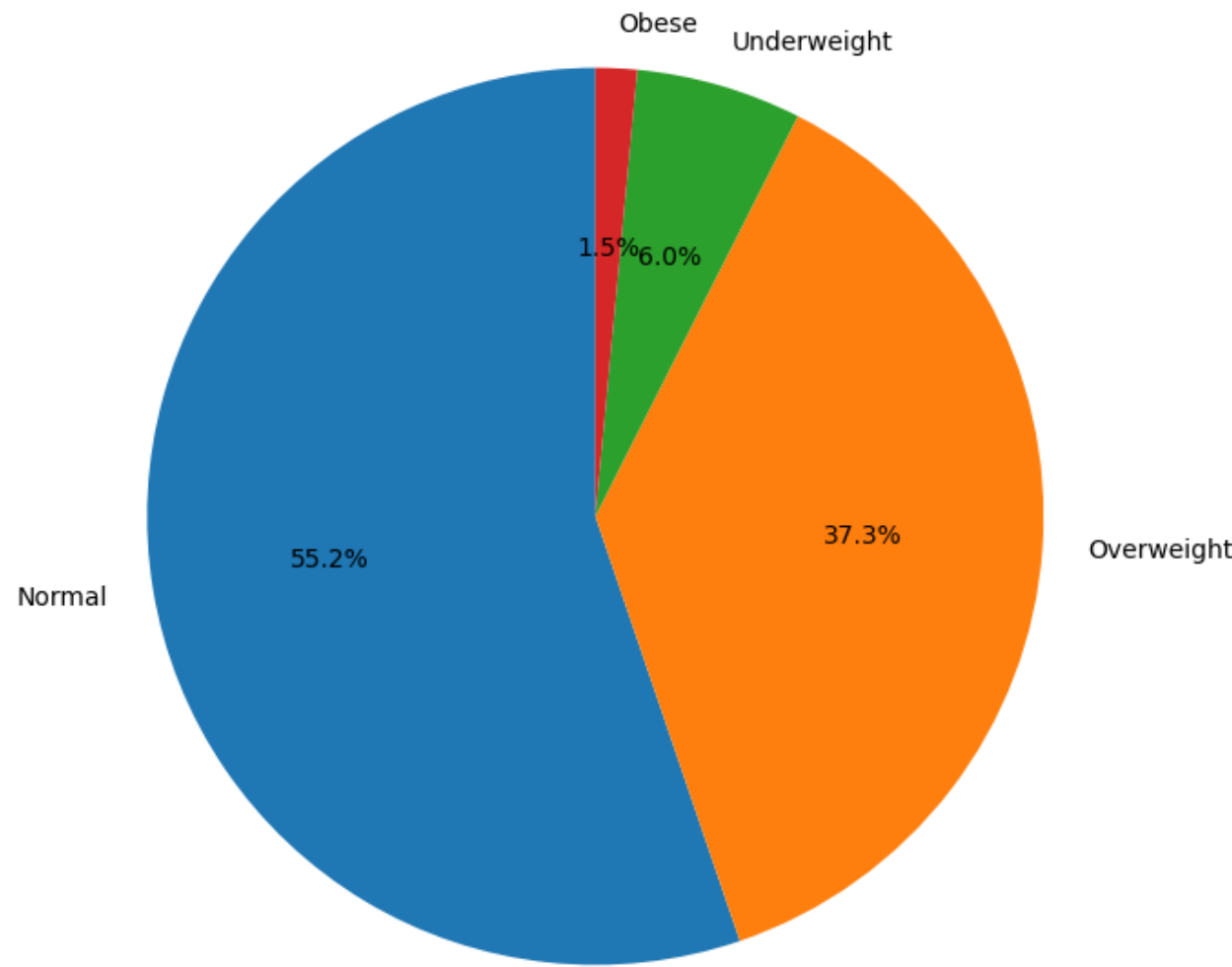
Weight Category Distribution and Logging Frequency

Key Findings:

1. **Weight Category Distribution:**

   - **The majority of users fall into the 'Normal' weight category.**
   - **'Overweight' and 'Obese' categories also have notable representation.**
   - **There is a smaller proportion of users classified as 'Underweight.'**

2. **Logging Frequency:**

   - **Users in the 'Normal' weight category have the highest frequency of weight logging.**
   - **'Overweight' and 'Obese' categories show moderate logging frequency.**
   - **'Underweight' category exhibits the lowest logging frequency.**

Features and Preferences Exploration:

In the context of understanding user behaviors and preferences among non-Bellabeat smart device users, the engagement with weight logging emerges as a notable pattern. This high level of utilization indicates a user base that places importance on health monitoring, specifically in the realm of weight management.

The preference for manual reporting further signifies a user demographic that values a hands-on and participatory approach to their health tracking. Such insights are crucial for device manufacturers seeking to understand and cater to the preferences of their user base, potentially influencing the design and functionality of future iterations of their smart devices.

Conclusion and Course of Action

In the exploration of non-Bellabeat smart device usage data, our analysis has illuminated significant insights into consumer behaviors and preferences. The key findings unveil dynamic patterns of user activity, with a notable interplay between sedentary behavior and active engagement throughout the week.

Recap of Insights Derived:

1. **Usage Dynamics Throughout the Week:**

   - **Users exhibit varying levels of sedentary behavior, notably higher during the early workweek and decreasing towards the weekend.**
   - **Weekends showcase unique usage patterns, indicating a shift towards leisurely or recreational activities while at points diplaying the lowest days of activity throughout the week**

1. **Usage Pattern And User Prefrences:**

   - **Among non-Bellabeat smart device users, a significant pattern is observed with enthusiastic engagement in weight logging, indicating a user base highly focused on health monitoring and specifically weight management.**
   - **Users prefer a hands-on and actively involved approach which is indicated by the constant use of manual loging features**

Course of Action:

Applying Insights to Bellabeat Product Presentation:

In aligning these insights with a Bellabeat product for presentation, Bellabeat smart watches emerges as a promising candidate. Leveraging the observed user behaviors, the following course of action is recommended:

1. **Tailoring Features for Dynamic Weekdays:**

   - **Enhance the product's features to align with user behaviors during weekdays, offering functionalities that cater to users during desk-oriented or work-related periods.**

2. **Promoting Weekend-Focused Features:**

   - **Introduce features that resonate with the observed shift towards leisure and recreation during weekends, encouraging users to engage with the device in ways that complement their weekend activities.**

3. **User Engagement Campaigns:**

   - **Develop targeted user engagement campaigns that highlight the product's adaptability to users' diverse lifestyles, emphasizing the balance between work and leisure.**

By strategically incorporating these insights into the Time Wellness Watch, Bellabeat can showcase a product that not only meets the functional needs of users but also resonates with their dynamic usage patterns throughout the week. This approach aims to enhance user satisfaction and foster a deeper connection between consumers and the Bellabeat brand.