Bellabeat Case Study

Harshitha

2025-04-17

Introduction

This report presents a detailed analysis of smart device usage data to help Bellabeat optimize their marketing strategy. The focus is on analyzing Fitbit data to understand user behavior and activity trends.

Business Task

Analyze trends in physical activity, sleep, and heart rate data collected from Fitbit users and identify actionable insights that Bellabeat can use to improve marketing and product decisions.

Data Distribution

The data was collected from 30 FitBit users and includes:

- Daily activity data(steps, distance, calories, etc.)
- Sleep data(duration, time in bed)
- Heart rate data(by second)

Summary Statistics

```
combined_data %>%
  summarize(avg_steps = mean(total_steps, na.rm = TRUE),
            avg_calories = mean(calories, na.rm = TRUE),
            avg_sleep = mean(total_minutes_asleep, na.rm = TRUE),
            avg_heart_rate = mean(avg_heart_rate, na.rm = TRUE))
```
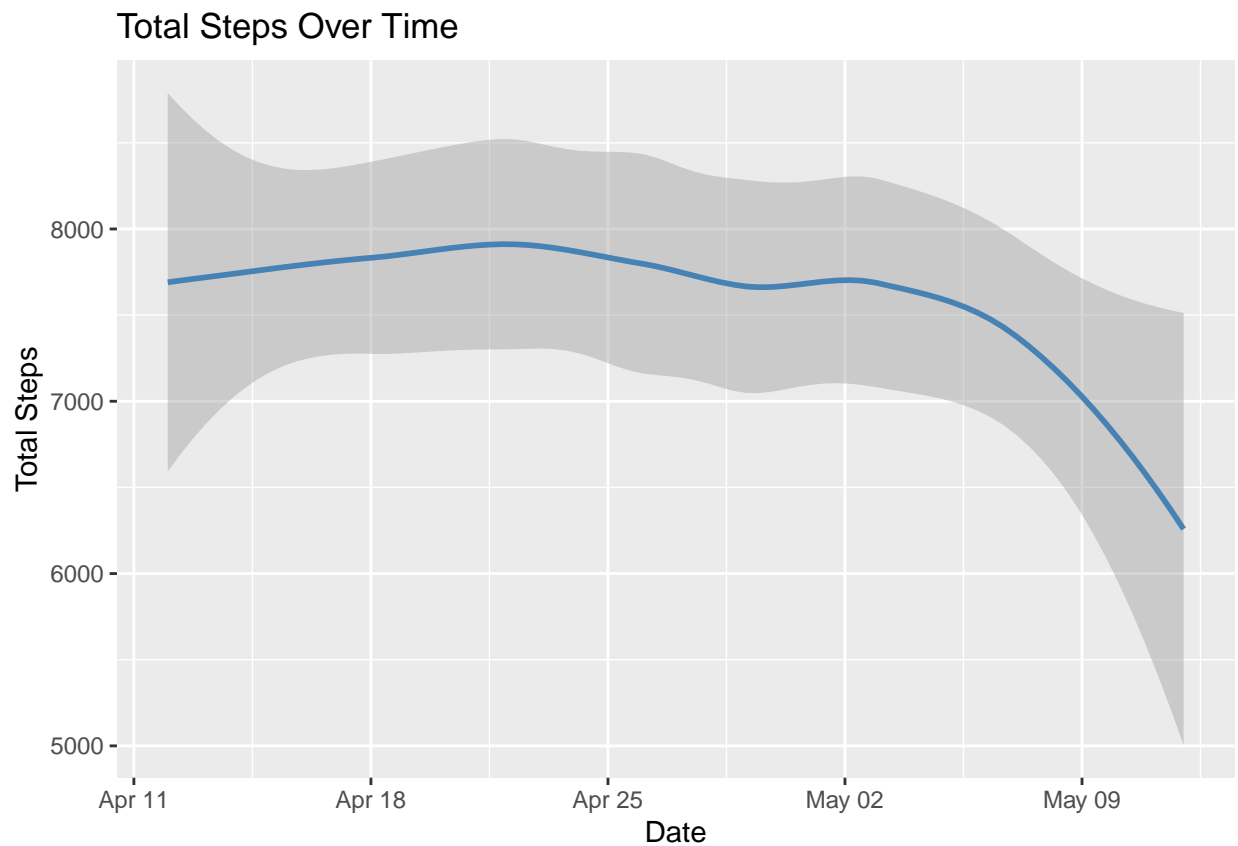
```
## # A tibble: 1 x 4
##   avg_steps avg_calories avg_sleep avg_heart_rate
##       <dbl>        <dbl>     <dbl>          <dbl>
## 1     7638.        2304.      419.           78.6
```

Visualizations

1. Steps over Time

```
ggplot(data = combined_data) +
  geom_smooth(mapping = aes(x = activity_date,y = total_steps),color = "steelblue")+
  labs(title = "Total Steps Over Time",x = "Date",y = "Total Steps")
```
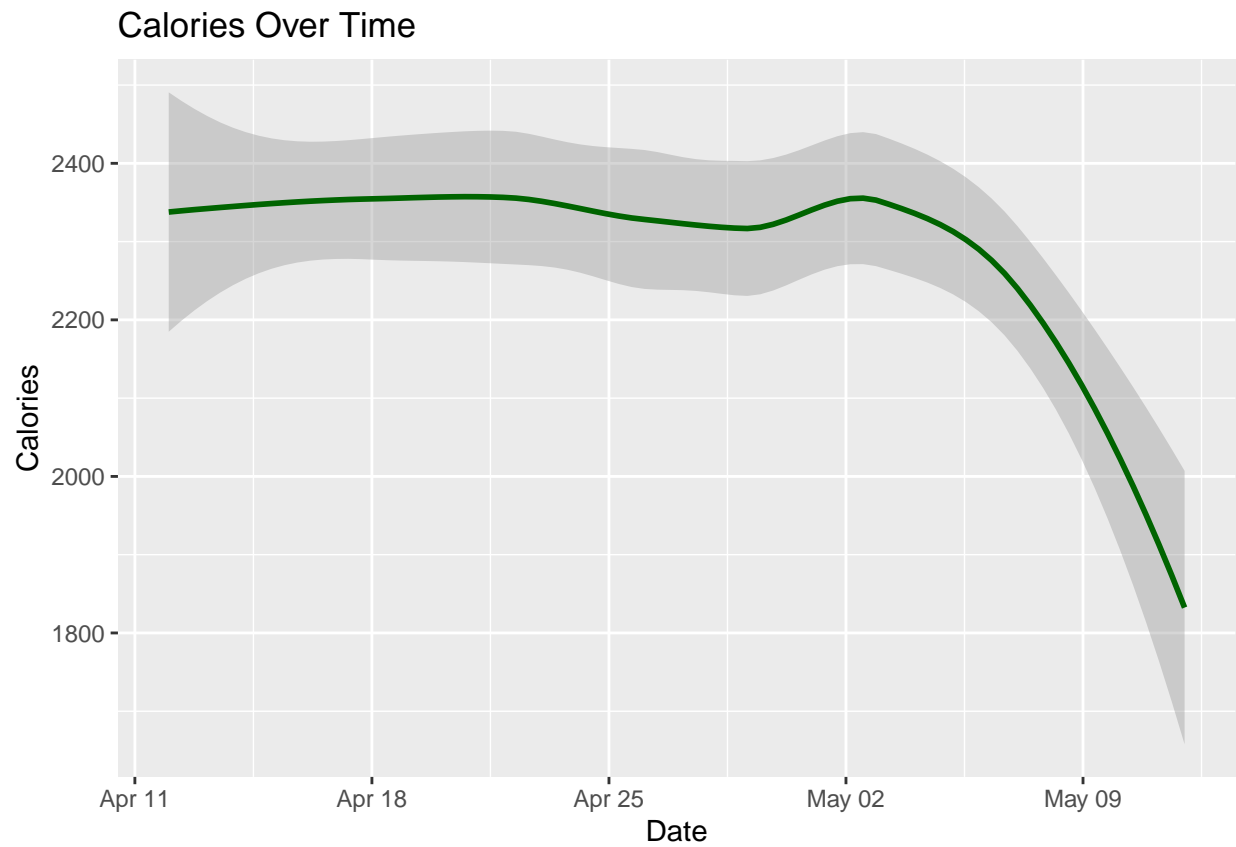
```
## `geom_smooth()` using method = 'loess' and formula = 'y ~ x'
```


2. Calories over Time

```
ggplot(data = combined_data) +
  geom_smooth(mapping = aes(x = activity_date,y = calories),color = "darkgreen") +
  labs(title = "Calories Over Time",x = "Date", y = "Calories")
```

```
## `geom_smooth()` using method = 'loess' and formula = 'y ~ x'
```

Calories Over Time

3. Calories vs Steps

```
ggplot(data = combined_data) +
  geom_point(mapping = aes(x = total_steps, y = calories),color = "tomato") +
  labs(title = "Calories Burned vs Steps Takes", x = "Total Steps",y = "Calories")
```

Calories Burned vs Steps Takes
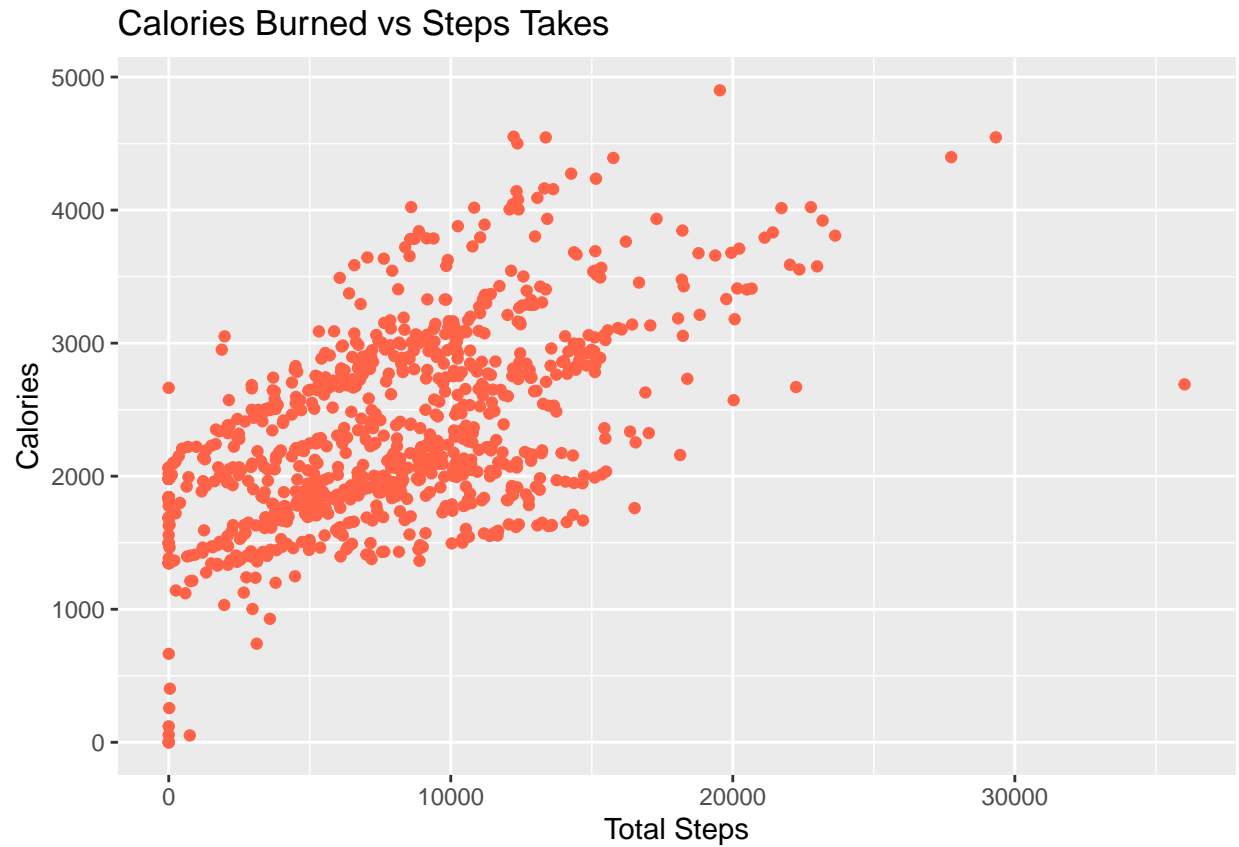
4. Average Steps by Weekday

```r
weekday_steps <- combined_data %>%
  mutate(weekday = factor(weekday,levels = c("Monday","Tuesday","Wednesday","Thursday","Friday","Saturda
  group_by(weekday) %>%
  summarise(avg_steps = mean(total_steps,na.rm = TRUE), .groups = "drop")

ggplot(data = weekday_steps) +
  geom_col(mapping = aes(x = weekday,y = avg_steps),fill = "mediumpurple") +
  labs(title = "Avearge Steps by Weekday", x = "Weekday", y = "Average Steps")
```

Avearge Steps by Weekday
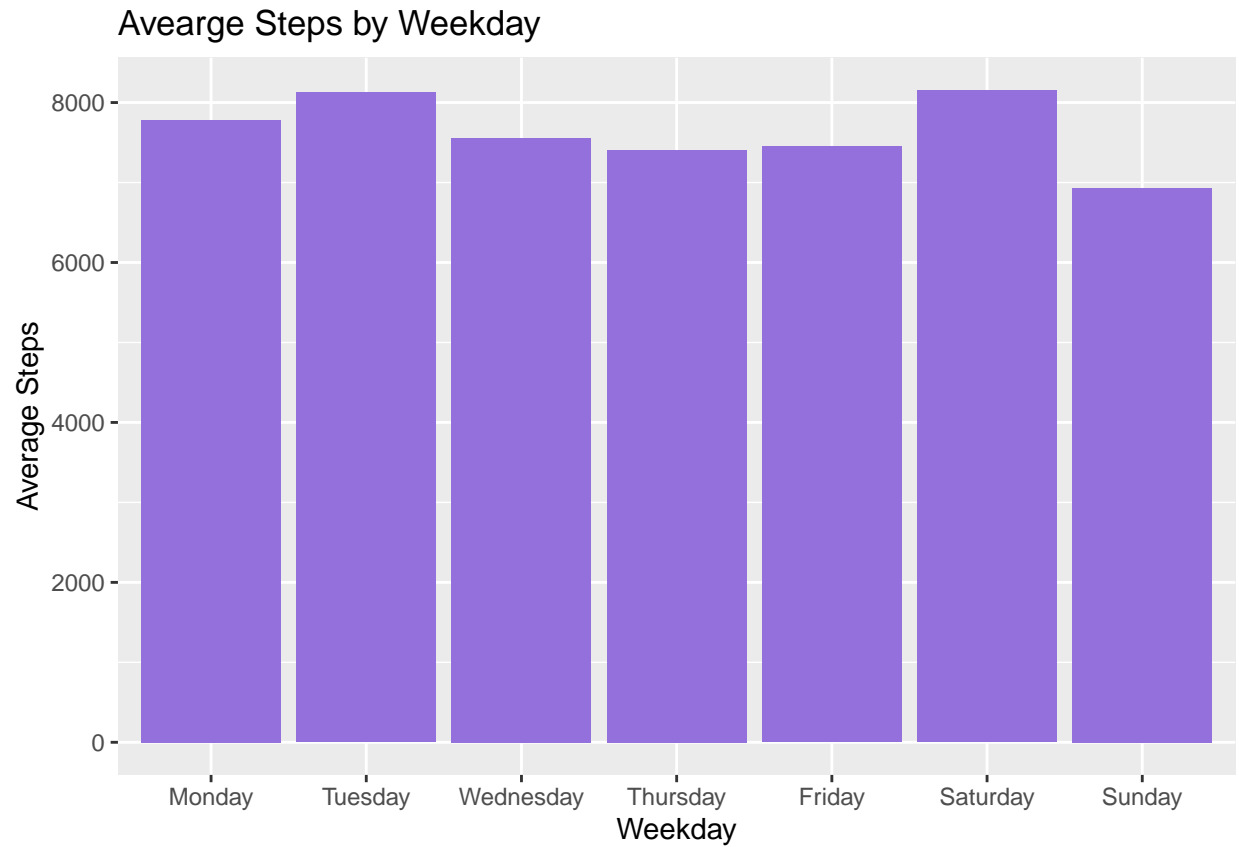
5. Distribution of Total Steps

```
ggplot(data = combined_data) +
  geom_histogram(mapping = aes(x = total_steps),binwidth = 1000, fill = "skyblue",color = "black") +
  labs(title = "Distribution of Total Steps", x = "Steps", y = "Count")
```
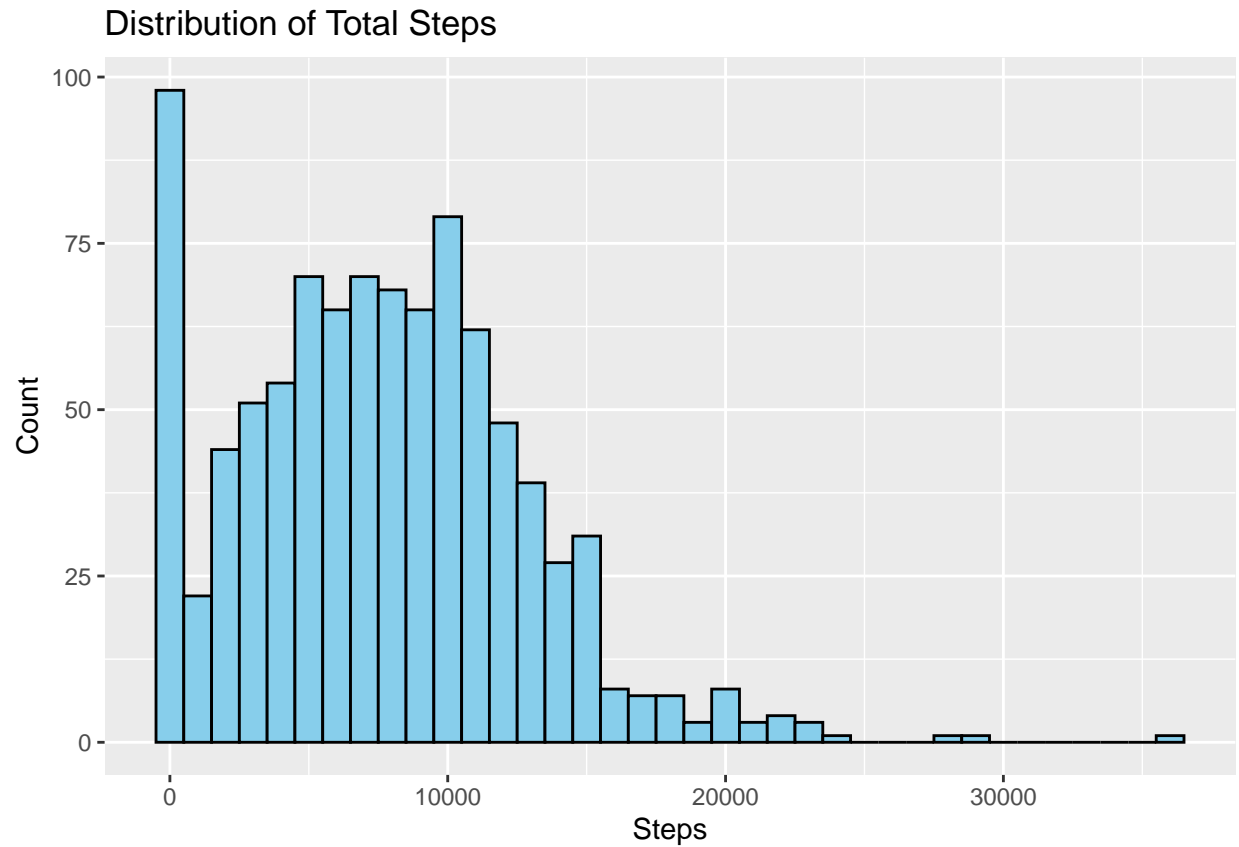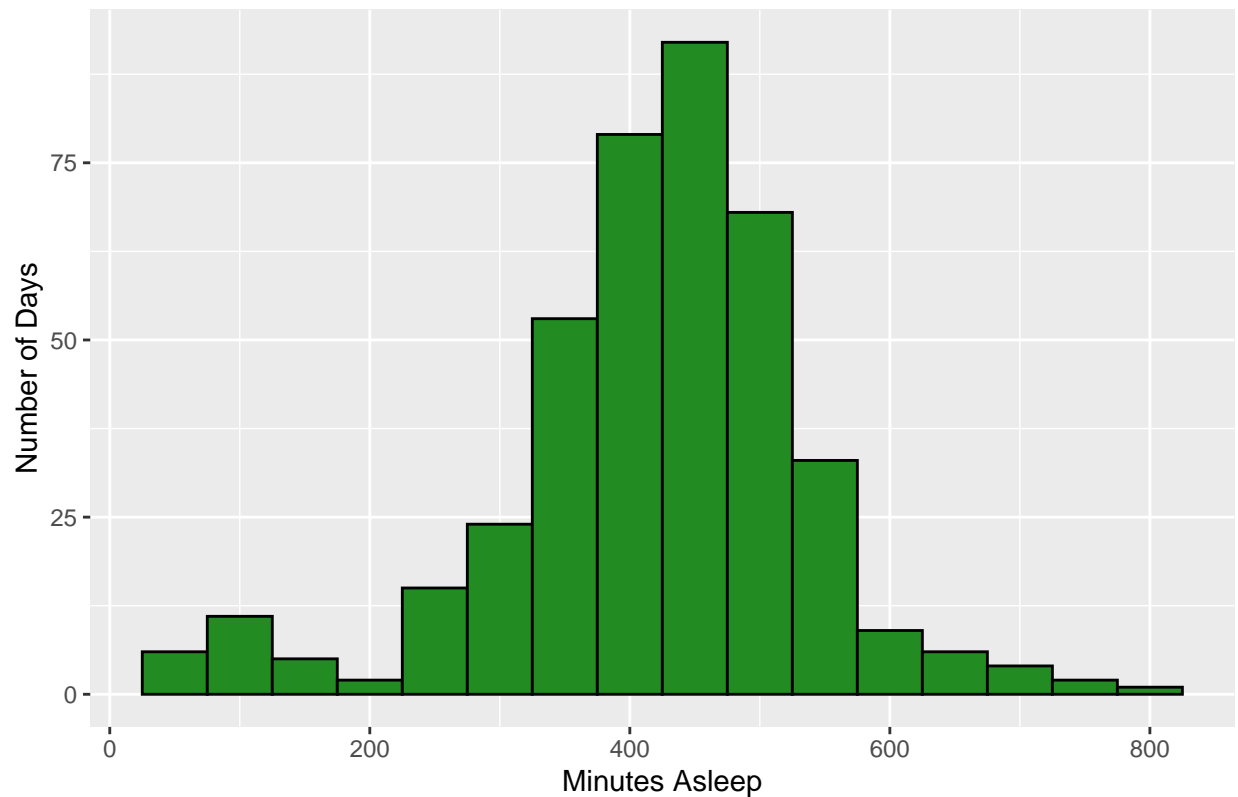
6. Sleep Duration Distribution

```
ggplot(data = combined_data) +
  geom_histogram(mapping = aes(x = total_minutes_asleep),binwidth = 50, fill = "forestgreen", color = "l
  labs(title = "Sleep Duration Distribution", x = "Minutes Asleep",y = "Number of Days")
```

```
## Warning: Removed 530 rows containing non-finite outside the scale range
## ('stat_bin()').
```
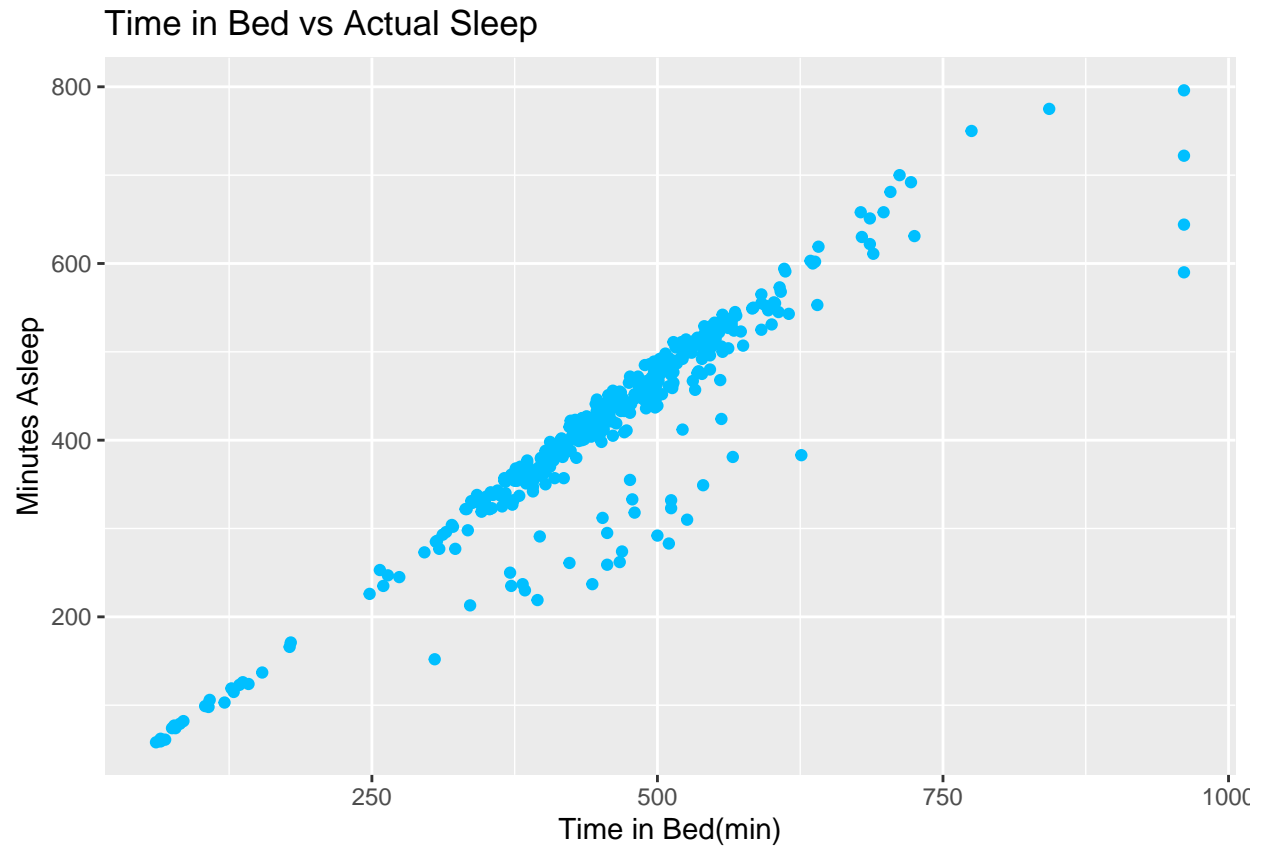
Sleep Duration Distribution

7. Sleep vs Time in bed

```r
ggplot(data = combined_data) +
  geom_point(mapping = aes(x = total_time_in_bed,y = total_minutes_asleep),color = "deepskyblue") +
  labs(title = "Time in Bed vs Actual Sleep", x = "Time in Bed(min)", y = "Minutes Asleep")
```

```
## Warning: Removed 530 rows containing missing values or values outside the scale range
## ('geom_point()').
```
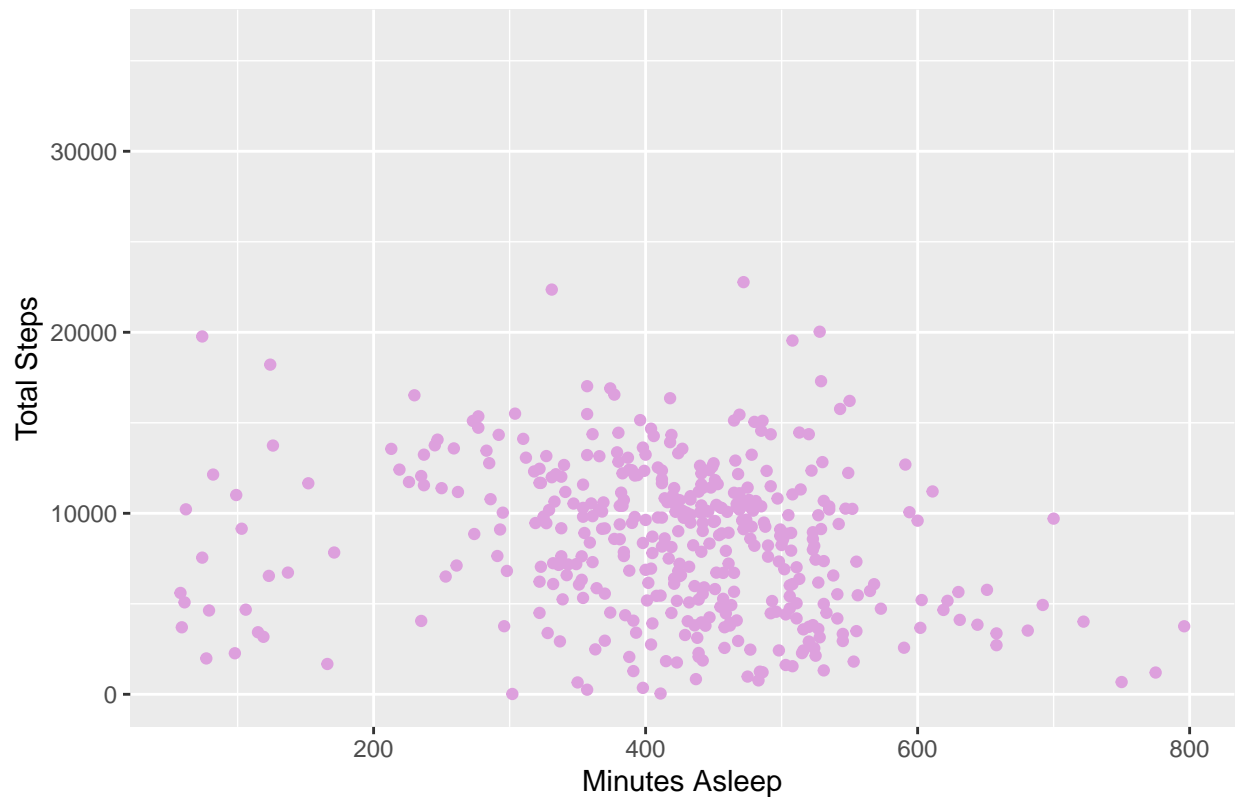
Time in Bed vs Actual Sleep

8. Sleep vs Steps

```
ggplot(data = combined_data) +
  geom_point(mapping = aes(x = total_minutes_asleep, y = total_steps),color = "plum") +
  labs(title = "Sleep Distribution vs Total Steps", x = "Minutes Asleep", y = "Total Steps")
```
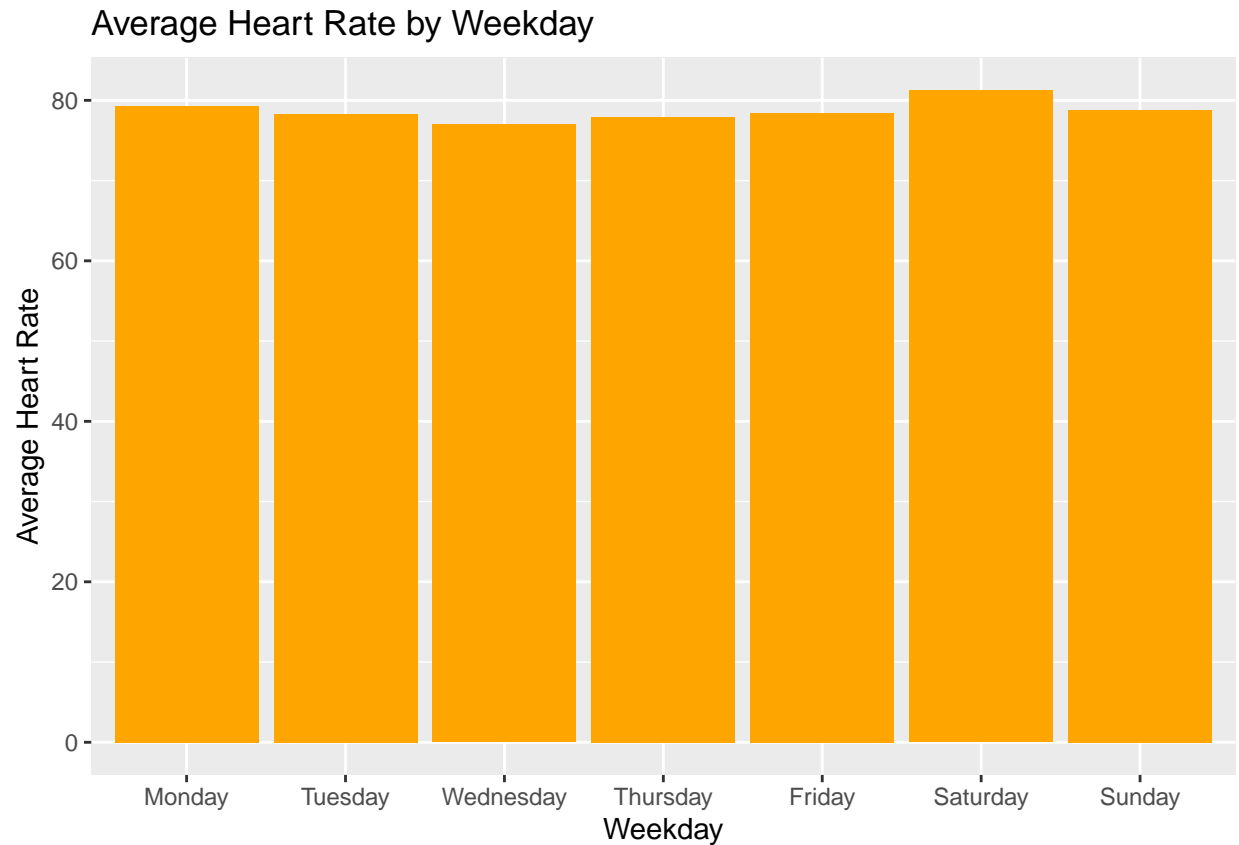
```
## Warning: Removed 530 rows containing missing values or values outside the scale range
## ('geom_point()').
```

Sleep Distribution vs Total Steps
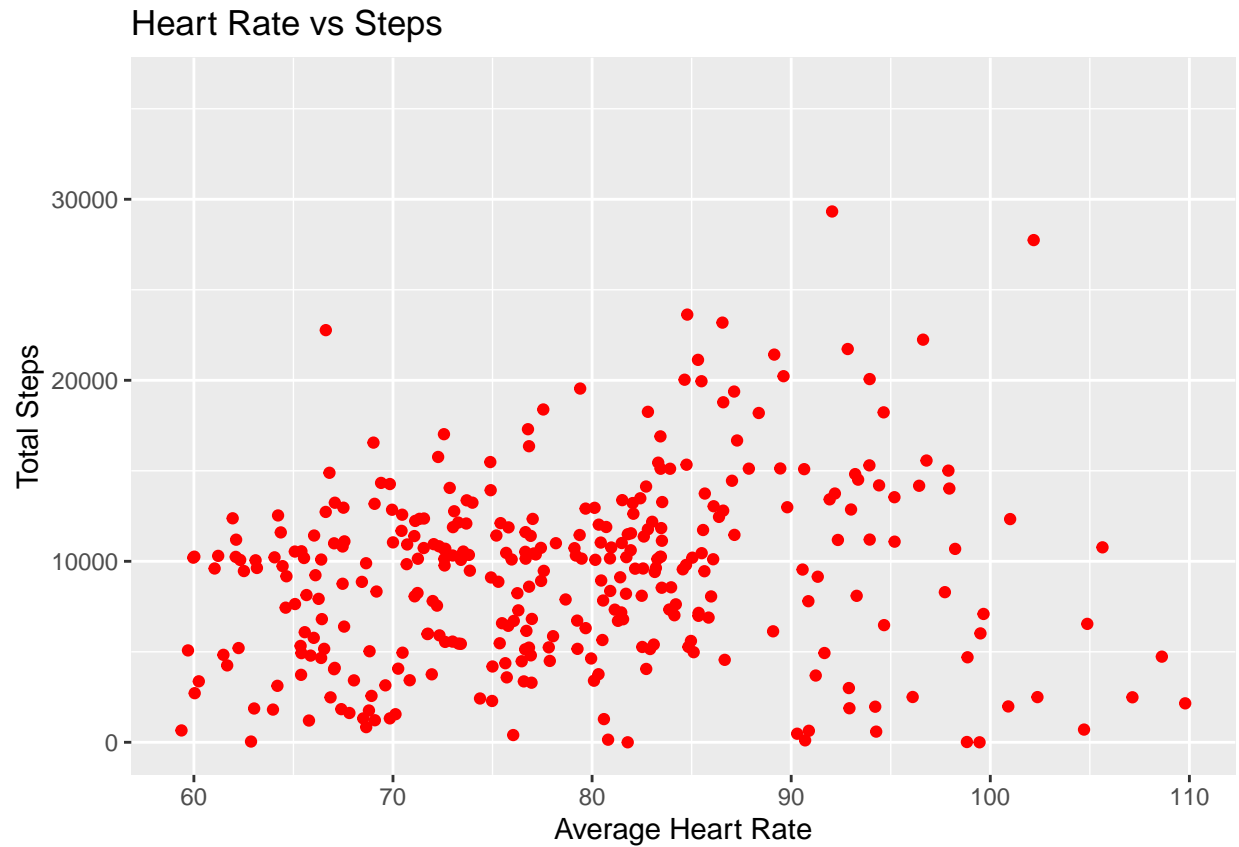

9. Average Heart Rate by Weekday

```
weekday_hr <- combined_data %>%
  mutate(weekday = factor(weekday,levels = c("Monday","Tuesday","Wednesday","Thursday","Friday","Saturda
  group_by(weekday) %>%
  summarise(avg_hr = mean(avg_heart_rate,na.rm = TRUE), .groups = "drop")

ggplot(data = weekday_hr) +
  geom_col(mapping = aes(x = weekday,y = avg_hr), fill = "orange") +
  labs(title = "Average Heart Rate by Weekday", x = "Weekday",y = "Average Heart Rate")
```

Average Heart Rate by Weekday

10. Heart Rate vs Steps

```
ggplot(data = combined_data) +
  geom_point(mapping = aes(x = avg_heart_rate,y = total_steps),color = "red") +
  labs(title = "Heart Rate vs Steps", x = "Average Heart Rate",y = "Total Steps")
```

```
## Warning: Removed 606 rows containing missing values or values outside the scale range
## ('geom_point()').
```

Heart Rate vs Steps

11. Time in Bed vs Calories Burned

```
ggplot(data = combined_data) +
  geom_point(mapping = aes(x = total_time_in_bed, y = calories),color = "darkblue") +
  labs(title = "Total Time in Bed vs Calories Burned", x = "Time in Bed(min)",y = "Calories")
```

```
## Warning: Removed 530 rows containing missing values or values outside the scale range
## ('geom_point()').
```

Total Time in Bed vs Calories Burned

12. Boxplot: Steps per Weekday

```
combined_data <- combined_data %>%
  mutate(weekday = factor(weekday,levels = c("Monday","Tuesday","Wednesday","Thursday","Friday","Saturd

ggplot(data = combined_data) +
  geom_boxplot(mapping = aes(x = weekday, y = total_steps),fill = "gold",outlier.color = "red",outlier.s
  labs(title = "Distribution of Steps per Weekday", x = "Weekday",y = "Total Steps")
```

Distribution of Steps per Weekday
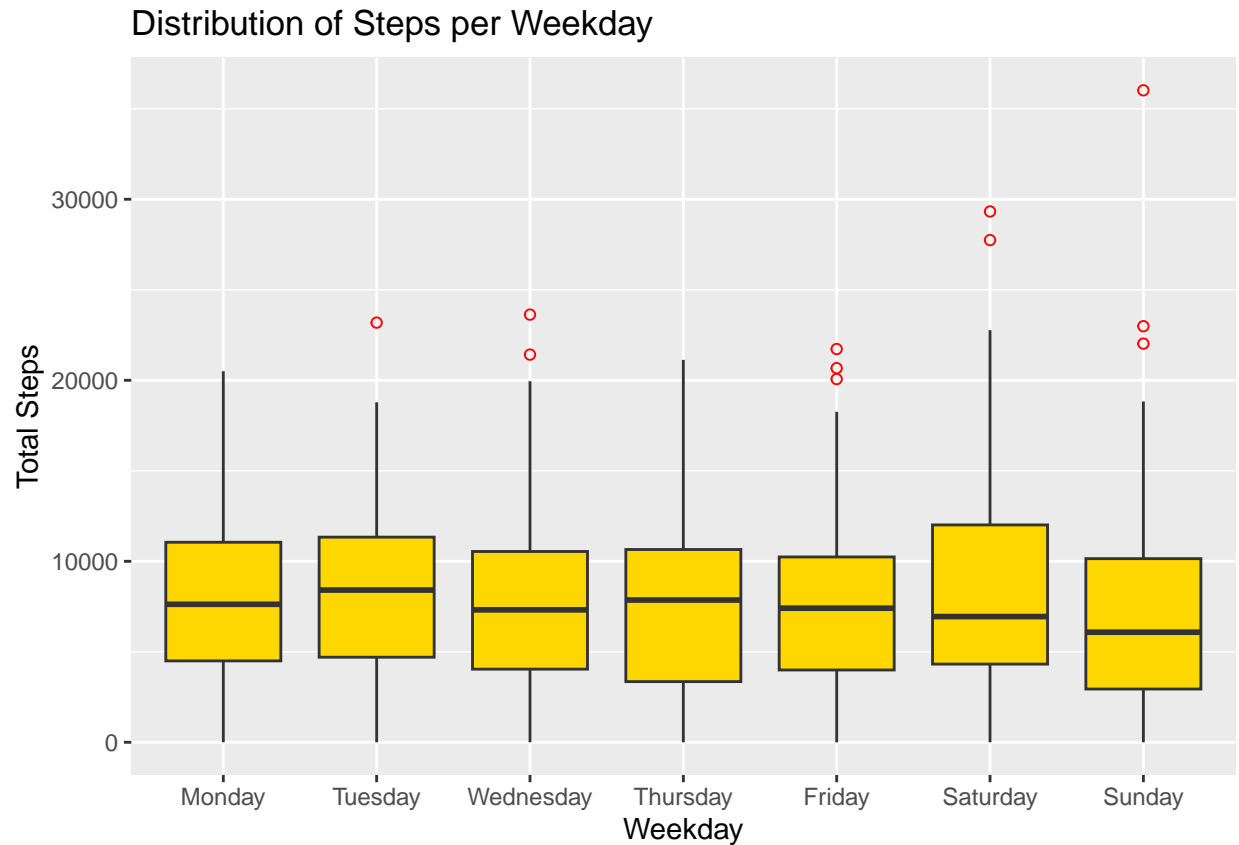
Recommendations

- Encourage users to walk more during low-activity days.
- Introduce reminders and sleep tips in the app based on sleep trends.
- Use heart rate + activity insights to tailor fitness plans.
- Focus marketing on days when users are most active.

Conclusion

The Fitbit data reveals user patterns in activity, sleep, and heart rate that Bellabeat can use to improve customer engagement and personalize features in their wellness app.