Bellabeat Case Study – Step-by-Step Procedure Summary

This document outlines the chronological steps taken during the Bellabeat case study project, including data preparation, analysis in R, visualization, and reporting.

# 1. Data Analysis in R

Step 1: Load Required Libraries

• library(tidyverse) – Loaded the Tidyverse collection of packages for data manipulation and visualization.

Step 2: Read CSV Files

• activity = read\_csv('dailyActivity\_merged.csv')  
• sleep = read\_csv('sleepDay\_merged.csv')  
• weight = read\_csv('weightLogInfo\_merged.csv')  
These steps imported the Fitbit data into R for processing.

Step 3: Convert date columns to proper Date format

• activity = activity %>% mutate(ActivityDate = as.Date(ActivityDate, "%m/%d/%Y"))

• sleep = sleep %>% mutate(SleepDay = as.Date(SleepDay, "%m/%d/%Y"))

Step 4: Inspect and Clean Data

• Used glimpse(), summary(), and sum(duplicated()) to explore and detect duplicates.  
• Removed duplicates using distinct().

Step 5: Merge Datasets

• combined = inner\_join(activity, sleep, by = c('ActivityDate' = 'SleepDay'))  
This merged the activity and sleep datasets on the date field.

Step 6: View a snapshot of the merged data

• View (combined)

Step 7: Summary Statistics

• Calculated min, max, mean, median for TotalSteps, Calories, SleepMinutes, etc., using summary().

Step 8: Visualizations

• Used ggplot2 to create scatter plots (Steps vs Calories) and line charts (Sleep over time).

Step 9: Add sleephours column for easier analysis

• combined = combined %>% mutate(SleepHours = TotalMinutesAsleep / 60)

# 2. Excel Analysis and Visualization

Step 10: Export Combined Data to Excel

• Used write\_xlsx(combined, 'combined\_data.xlsx') in R to export the merged dataset.

Step 11: Load Data in Excel

• Opened 'combined\_data.xlsx' to analyze trends using Excel tools.

Step 12: Pivot Table Analysis

• Inserted PivotTables to summarize TotalSteps, Calories, and SleepHours.  
• Used Value Field Settings to compute average, max, and min values.

Step 13: Created Visual Charts

• Created Line Chart (Calories over time).

# 3. Reporting and Sharing

Step 14: Inserted Charts and Summaries into Word Report

• Copied visualizations from Excel and R into a Word template.  
• Wrote up insights, recommendations, and conclusions based on the data.

Step 15: Expanded Report with Advanced Analysis

• Added correlation analysis, regression, user segmentation, and limitations sections to the report.