Strava Fitness App Data Analysis Report

Manisha Kaushal

1. Overview

Strava is a popular fitness tracking app that records user activities like steps, sleep, calories burned, heart rate, and more. This project involves analyzing user-level fitness data to derive actionable insights. The main goal is to evaluate patterns in physical activity, sleep behavior, and physiological signals to understand user engagement, health patterns, and improve service personalization.

2. Business Task

The objective is to understand how users interact with the fitness app across various metrics like steps, calories, heart rate, and sleep. The insights aim to inform improvements in user engagement strategies, personalized recommendations, and health-based nudges.

Summary

We will explore daily user behavior, quality of data recorded, relationships between variables (e.g., sleep and activity), and identify top-performing users. This analysis provides both aggregate metrics and user-specific behaviors.

3. Dataset Descriptions

The project utilizes the following 11 datasets:

- **dailyActivity.csv**: Records daily step counts, calories burned, distance, and activity duration.
- **sleepDay.csv**: Contains total sleep minutes, time in bed, and sleep efficiency.
- **weightLogInfo.csv**: Logs weight in kg/pounds, BMI, and body fat percentage.

- **heartrate_seconds.csv**: Time-stamped heart rate measurements.
- **hourlyCalories.csv**: Calories burned each hour.
- **hourlySteps.csv**: Steps per hour.
- **hourlyIntensity.csv**: Hourly total intensity and average intensity.
- **minuteCaloriesNarrow.csv**: Minute-level calories burned.
- **minuteStepsNarrow.csv**: Steps taken per minute.
- **minuteIntensityNarrow.csv**: Minute-wise total and average intensity.
- **minuteSleep.csv**: Per-minute sleep status.

4. Data Cleaning and Manipulation Summary

- **Deduplication**: Removed duplicate entries based on IDs and timestamps.
- **Outlier Removal**: Heart rates outside 40–220 bpm range were flagged. Step counts above 100,000 and calorie values of 0 were nullified.
- **Type Casting**: Dates and timestamps were uniformly cast to `DATE` or `DATETIME` formats.
- **Merging**: Cleaned views were merged on `Id` and `ActivityDate` with proper alignment using date differences or time rounding.
- **Data Quality Flags**: Labeled entries as 'Complete', 'Partial', or 'Invalid' based on presence of core metrics.

5. Analysis and Insights

1: Daily Averages by User

Avg Steps	Avg Calories	Avg Sleep Minutes
8608	2393	421

Insight: Most users average about 8600 steps and 2393 calories/day, suggesting moderate activity levels.

Business Impact: Useful baseline for setting app nudges and personalized activity goals.

2: Users with Best Sleep Consistency

User ID	Avg Sleep Duration	Count of Sleep Days
4558609924	15	23
2026352035	79	30
2347167796	44	42
7086361926	67	49
1844505072	9	54
6962181067	91	61
3977333714	83	63
8792009665	44	64
4702921684	79	66
5577150313	73	67

Insight: Users with consistent sleep patterns can be targeted for sleep tracking feature enhancements.

Business Impact: Promotes premium features (like smart alarm, sleep analysis) to highly consistent users.

3: Calories Burned vs Activity Minutes

Active Minutes	Calories Burned
366	1985
257	1797
222	1776
272	1745
267	1863
291	1921
345	2035
245	1786
223	1827

Insight: Positive relationship between active minutes and calories burned.

Business Impact: Can be used to provide custom calorie burn predictions or fitness plans.

4: High Heart Rate Days vs. Activity Levels

HR Category	Day count
Moderate	13
High	7
Low	3

Insight: Segmenting HR days gives a view into how intense the user's fitness habits are. High frequency of "High HR" days suggests active or intense routines.

Business Impact: Help users balance training with recovery, or flag signs of overtraining if too many consecutive high-HR days appear.

5: Average Resting and Maximum Heart Rate

Resting HR	Max HR
51	149

Insights

- Ideal values are 60–80 bpm for adults. Lower RHR usually indicates better cardiovascular fitness.
- Max Heart Rate: High values reflect workout intensity. If average Max HR is consistently above 85

Business Impact: Recommend recovery suggestions for high HR, flag unusual patterns (e.g., too high RHR), and personalize fitness zones or workout recommendations.

6: MaxHeartRate vs Calories Correlation

Value: 0.708

Insight: Strong correlation between higher max HR and calories burned.

Business Impact: Suggests high-effort workouts are effective calorie-burn drivers, useful in coaching.

6. Conclusion

This analysis of Strava Fitness App user data reveals significant patterns in user behavior, physiological responses, and app engagement:

- Most users show moderate physical activity (avg. 8600 steps/day), with consistent calorie burn and sleep time.
- Users with high sleep consistency are ideal candidates for premium sleep tracking features.
- There is a clear positive correlation between active minutes and calories burned, as well as max heart rate and calorie burn (corr = 0.708), validating the app's tracking capabilities.
- Resting and maximum heart rate averages indicate a generally healthy user base.

These insights can be leveraged to:

- Design personalized fitness recommendations
- Target promotions for sleep analysis tools
- Improve engagement with smart goals based on historical activity and heart rate patterns