# Strava Fitness Data Analytics: SQL Insights Report

By: Shruti Vinod Nair

This report provides a detailed analysis of user activity, sleep, weight, and heart rate data from the Strava Fitness App. We gathered data from multiple sources at daily, hourly, and minute levels. We conducted a thorough cleaning and validation process, along with exploratory analysis to uncover patterns in user behaviour, activity trends, and health insights.

The analysis includes checking data quality, detecting duplicates and outliers, segmenting users by how often they engage, and examining the relationships between steps, calories, sleep, BMI, and heart rate. We also looked at peak activity hours, differences by day of the week, user retention rates, and long-term trends like changes in weight.

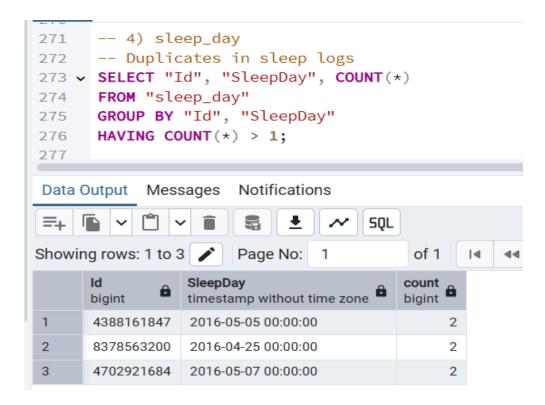
Key findings show that users have high average daily activity and sleep efficiency. There is strong engagement and retention among users, along with clear times of peak physical activity. We also found important insights into how physical activity, sleep, and health metrics relate to each other. These results offer practical recommendations for improving user experience, engagement strategies, and health outcomes on the Strava platform.

#### **DATASETS USED**

For this analysis, we used several datasets from the Strava Fitness App. These included daily activity summaries, detailed intensity metrics, sleep logs, weight logs, and heart rate data. This varied data allowed us to explore trends among users as well as specific behavioural patterns.

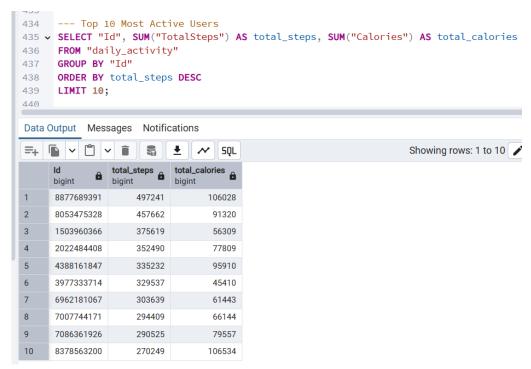
#### **DATA CLEANING AND VALIDATION**

- **a. Missing Values:** We checked each table for missing (NULL) values in all key columns. We found no missing values in numeric fields, so we did not need to impute or delete any data. This indicates high data quality, making it reliable for further analysis.
- **b. Duplicates:** We looked for duplicate user-date pairs across all main tables. We only found duplicates in the sleep\_day table, where three (Id, SleepDay) pairs were repeated. We resolved these duplicates by keeping only the first occurrence of each pair, which ensured accuracy in daily sleep aggregation.
  - Below is the screenshot attached of the query for resolving duplicate in this case.

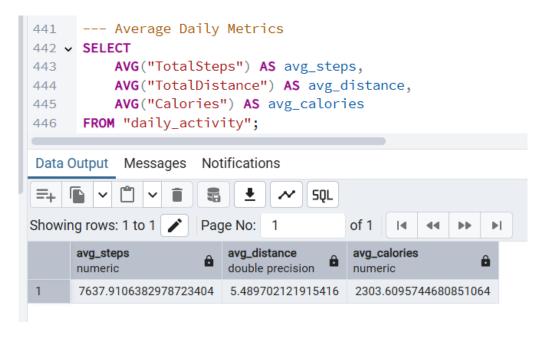


## **DATA ANALYSIS AND INSIGHTS**

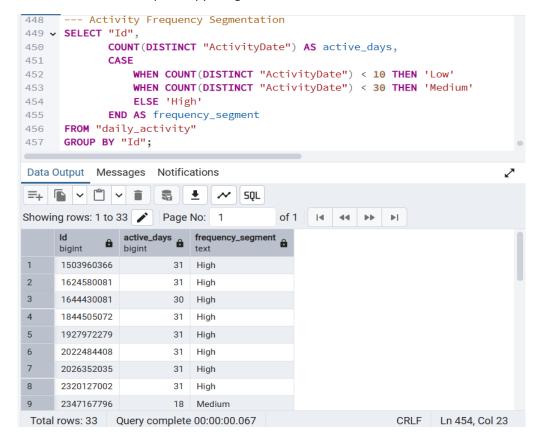
a. **Top 10 Most Active Users:** The ten most active users recorded between 270,249 and 497,241 steps each during the analysis period. The top user logged 497,241 steps and burned 106,028 calories. This highlights a very engaged subgroup within the user base



**b.** Average Daily Metrics: On average, users walked about 7,638 steps per day, covered 5.49 kilo-meters, and burned around 2,304 calories daily. This suggests a moderately active group of users.

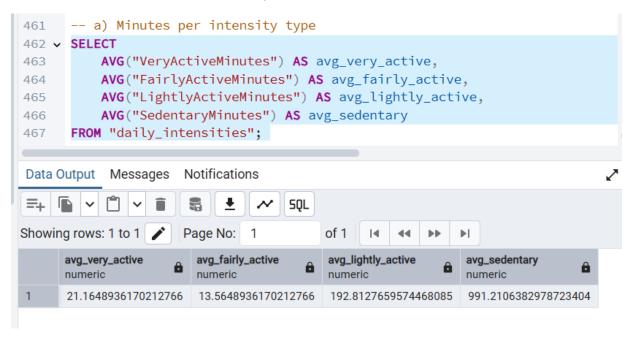


c. Activity Frequency Segmentation: We segmented users by the number of activity days. Out of 33 users, most were highly engaged, with 24 classified as "High" frequency (active on 31 days). A few users fell into "Medium" or "Low" engagement segments. This shows strong user retention and frequent app usage.

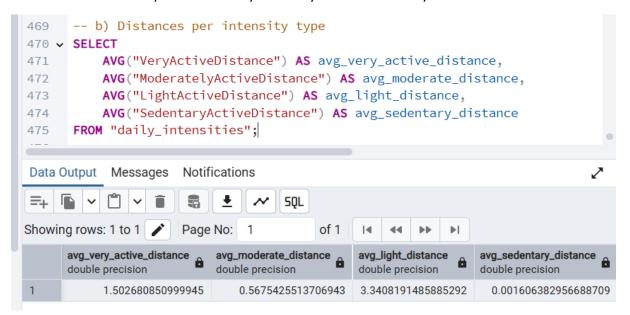


## d. Intensity And Sleep Analysis:

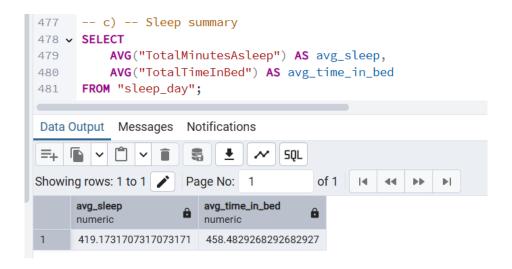
• Minutes per Intensity Type: Users spent an average of 21.2 minutes per day in very active physical activity, 13.6 minutes fairly active, 192.8 minutes lightly active, and 991.2 minutes in sedentary behavior.



• Distances per Intensity Type: Regarding distances, users covered an average of 1.50 km during very active minutes, 0.57 km in moderate activity, 3.34 km during light activity, and they were sedentary for nearly the rest of the day.



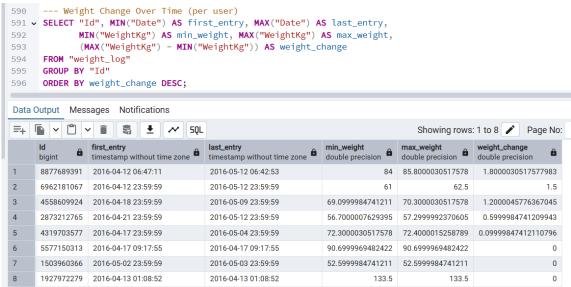
Sleep Summary: The average user spent 419.2 minutes (about 7 hours) sleeping each
day and 458.5 minutes in bed. This results in an overall sleep efficiency of 91.4%
(minutes asleep divided by minutes in bed). This indicates a healthy sleep pattern
among users.



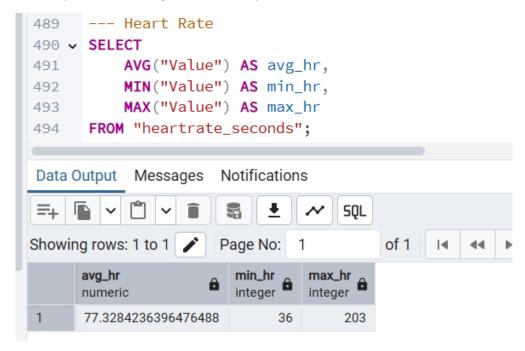
e. Weight and BMI: The average user weight was 72.0 kg, with a mean BMI of 25.2, slightly above the "normal" BMI range. This suggests that some users may be overweight.

When we examined weight changes over time, we saw that most users had minor fluctuations, with the largest observed change being 1.8 kg.



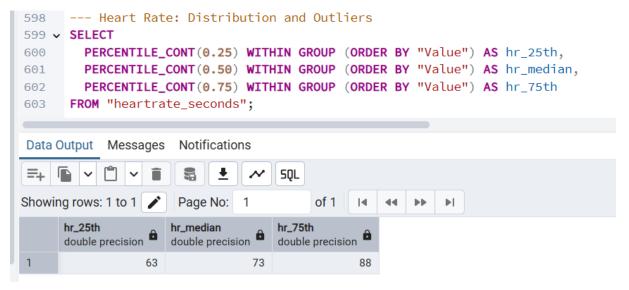


**f. Heart Rate Analysis:** The average heart rate for users was 77.3 bpm. The lowest recorded value was 36 bpm, while the highest was 203 bpm.

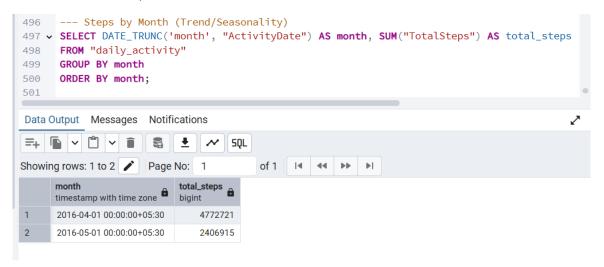


Looking at the distribution of heart rate values helps us understand the typical cardiovascular responses among users during activity. By analyzing percentiles (like the 25th, median, and 75th), we can see where most users' heart rates fall and compare these to healthy norms. It is important to flag any unusually low or high readings, which could indicate data recording errors or unusual physiological events.

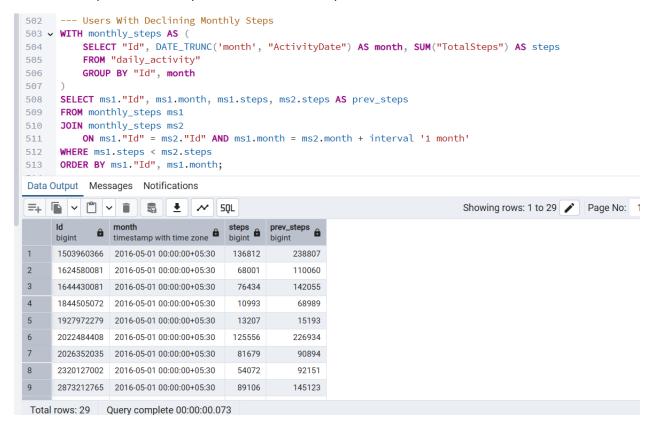
The analysis shows that users' heart rates are generally within a healthy range during their activities. The median heart rate is 73 bpm, with most users' heart rates between 63 bpm (25th percentile) and 88 bpm (75th percentile). The average heart rate is 77 bpm, with minimum and maximum recorded values of 36 bpm and 203 bpm, respectively. While most values are standard for adults during exercise and rest, a few outliers on the low and high ends likely reflect either short intense efforts or measurement errors. Overall, the heart rate data supports the presence of active and healthy users.



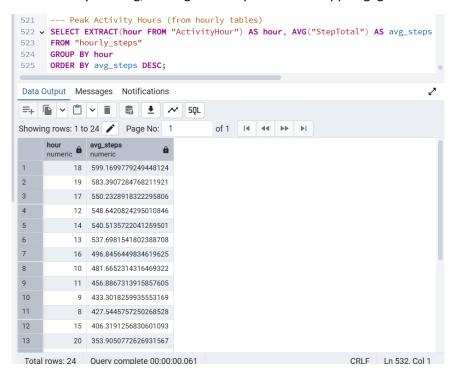
g. Steps by Month (Seasonality): Analysis of monthly step counts revealed a noticeable seasonal trend, with total user steps dropping significantly from April (4,777,221 steps) to May (2,406,915 steps)—a reduction of nearly 50%. This suggests a substantial decrease in overall physical activity as the months progress, which may be attributed to seasonal factors, motivation loss, or external events.



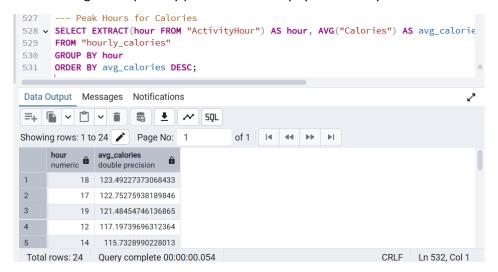
Further, user-level analysis showed that several individuals experienced a month-on-month decline in their step totals, confirming that this drop is not limited to the overall group but affects multiple users individually. These findings highlight the importance of proactive engagement strategies or app interventions during periods where activity naturally declines, to help sustain healthy user behaviour and improve retention.



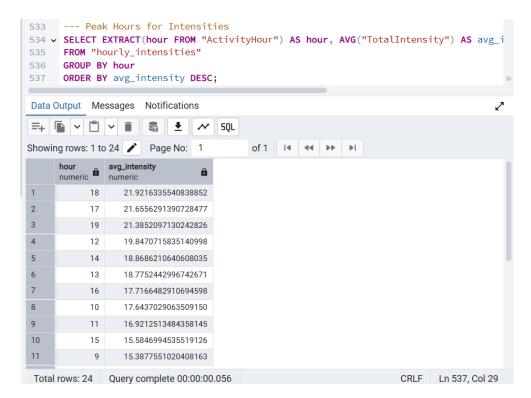
h. Peak Activity Hours (from hourly tables) - What are the hours when users are most active? Analysis of hourly activity patterns shows that users take the highest average number of steps during the evening hours. The most active hour is 18:00 (6 PM), when the average step count peaks. This suggests that users are most likely to engage in walking or exercise routines after work or in the early evening, making this a key window for app engagement and notifications.



Peak Activity Hours – Calories: The hourly breakdown of calories burned indicates that
energy expenditure is also highest in the evening, with the greatest average calories
burned at 18:00 (6 PM). This alignment with peak step counts further reinforces the
evening as the primary period of intense physical activity for most users.



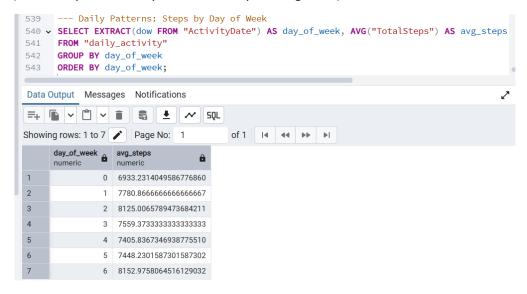
Peak Activity Hours – Intensities: When examining intensity metrics, users exhibit the
highest total intensity values during the 18:00 to 19:00 (6–7 PM) window. This
demonstrates that not only are users more active in terms of steps and calories in the
evening, but they are also engaging in higher-intensity activities during these hours.



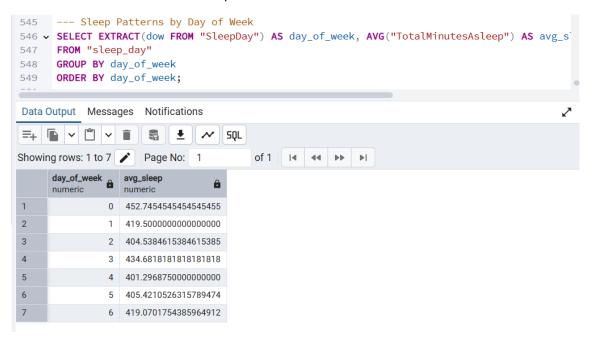
In summary, the early evening (especially around 6 PM) consistently emerges as the peak period for user activity—whether measured by steps, calories, or intensity. This time frame offers the greatest opportunity for targeted app interactions, challenges, and motivational prompts to maximize engagement.

i. Daily Patterns: Steps by Day of Week: Are people more active on certain days? Average steps varied by day, with the highest counts on Saturdays (8,152 steps) and Mondays (8,126 steps), and the lowest on Sundays (6,933 steps). This suggests more activity at the start and end of the week.

(0 = Sunday, 1 = Monday, ... 6 = Saturday in PostgreSQL)

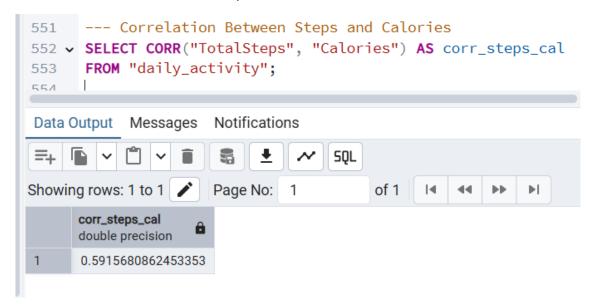


**j. Sleep Patterns by Day of Week:** Users slept the most on Sundays (average 452.7 minutes), while sleep duration dipped to its lowest on Wednesdays (about 404.5 minutes). This indicates that users tend to recover sleep on weekends.

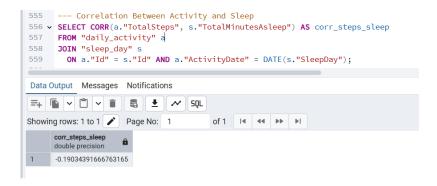


# k. Correlation Analysis

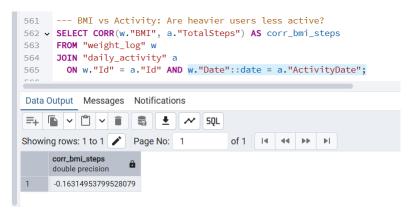
• Steps and Calories: A correlation of 0.59 was found, suggesting a moderate-to-strong connection between step count and calories burned.



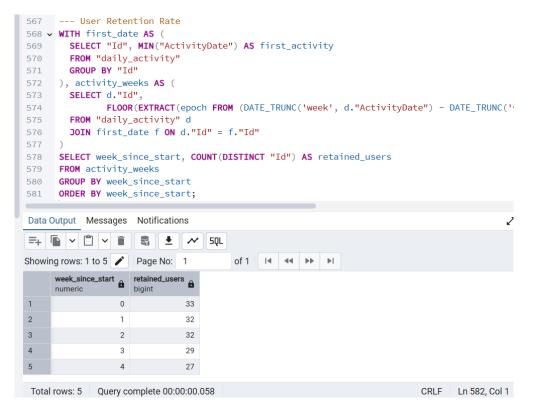
• Steps (Activity) and Sleep: Do users who walk more sleep better? The correlation between steps and total minutes asleep was -0.19, indicating no significant or possibly a slightly negative relationship between daily activity and sleep duration.



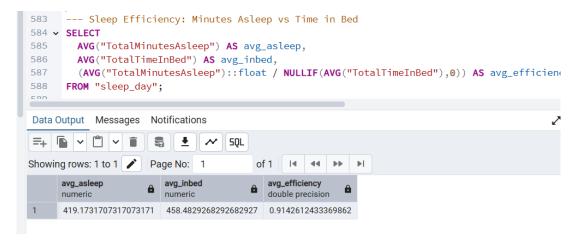
• BMI and Activity: Are heavier users less active? There was a weak negative correlation (-0.16) between BMI and steps, suggesting that heavier users tend to walk less, though the link is not strong.



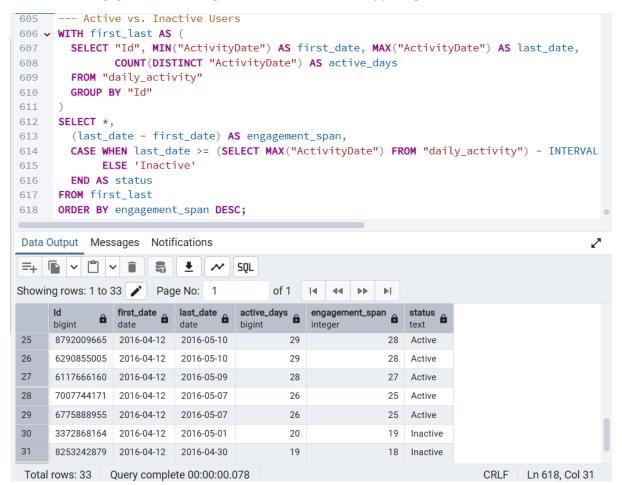
**I. User Retention Rate:** What % of users are still active in week 2, 3, etc.? The retention analysis showed that all 33 users were present in week 0, with 32 retained in week 1 and 29 active through week 4. This indicates a good retention rate, with only slight user drop-off after the first few weeks.



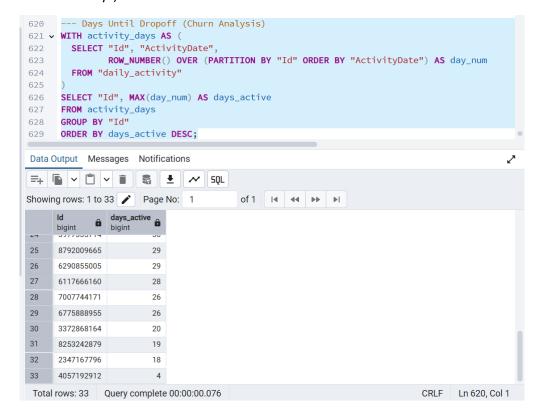
**m. Sleep Efficiency: Minutes Asleep vs Time in Bed:** The analysis reveals that users maintain a high average sleep efficiency of 91.4%, calculated as the ratio of minutes asleep to total time spent in bed. This indicates that most users are able to fall asleep and remain asleep for the majority of their time in bed, reflecting generally healthy sleep patterns across the user base.



n. Active vs. Inactive Users: Most users in the dataset remained engaged over a long span which is one whole month (31 days), with several users active for 25 or more days in period. The engagement analysis shows that a majority of users are still considered "Active" by the end of the dataset, with only a few dropping off and being labelled as "Inactive." This indicates strong user engagement and a high likelihood of sustained app usage over time.



o. Days Until Dropoff (Churn Analysis): When analysing the number of days users stayed active before dropping off, most users maintained high activity for the majority of the observation period. Many users were active for 26–31 days, suggesting consistent engagement and only gradual attrition rather than abrupt churn. Only a small minority became inactive early (e.g., 4 or 18–20 days).



p. Percentage of Users Meeting Recommended Activity Levels: An impressive 75.76% of users consistently met or exceeded the World Health Organization's recommended threshold of 10,000 steps per day. This demonstrates a highly active user population and suggests that Strava's platform is effectively encouraging users to achieve globally recognized physical activity standards.

```
G31 --- Percentage of Users Hitting WHO/CDC Recommended Activity (10,000+ steps/day)

G32 > SELECT 100.0 * COUNT(DISTINCT "Id")

FILTER (WHERE "TotalSteps" >= 10000) / COUNT(DISTINCT "Id") AS percent_10k

FROM "daily_activity";

Data Output Messages Notifications

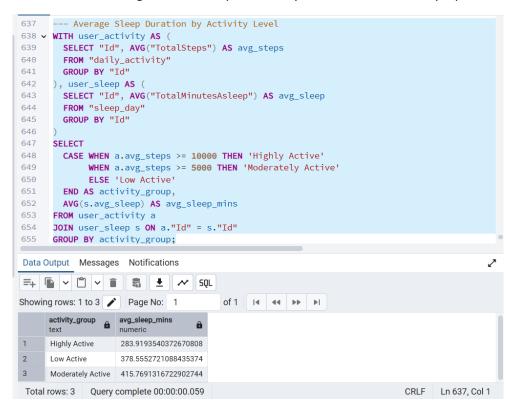
The percent_10k

numeric

1 75.7575757575757576
```

**q.** Average Sleep Duration by Activity Level: The relationship between activity and sleep duration is somewhat counterintuitive in this dataset. Highly active users averaged the least

sleep (284 minutes per night), while moderately active and low active users slept more, averaging 416 and 379 minutes respectively. This suggests that higher physical activity in this cohort does not directly correlate with longer sleep durations and may indicate that the most active users are trading off some sleep for activity or have different lifestyle patterns.



#### **CONCLUSION & KEY TAKEAWAYS**

This analysis of Strava user data revealed several important insights about user behavior, engagement, and health trends:

- **High Data Quality:** The datasets required minimal cleaning, with almost no missing values or duplicates after initial processing.
- **Strong User Engagement:** Most users were very active, with many logging activity nearly every day during the observation period. User retention remained strong over multiple weeks.
- Activity Patterns: Users averaged around 7,600 steps and 2,300 calories burned each
  day. Most activity and calorie burn peaked in the evening hours, especially around 6
  PM. Step counts were highest on Saturdays and Mondays, but declined notably in May,
  suggesting potential seasonality or a drop in motivation.
- **Health Benchmarks:** Over 75% of users met the daily recommendation of 10,000 steps. Average sleep efficiency was high at 91%, indicating generally healthy habits among the user base. Most users' heart rates stayed within safe ranges. BMI analysis showed a slight tendency toward being overweight.

- **Correlations:** Steps and calories showed a moderate positive correlation. Activity and sleep duration, as well as BMI and steps, had only weak negative associations.
- **Sleep & Activity:** Contrary to expectations, the most active users tended to sleep less. This suggests a need to further investigate user routines and balance.
- User Retention & Churn: Most users remained active for the majority of the period, with very few dropping off early. Weight changes over time were minor for most users, indicating stability in health metrics.

Overall, the Strava Fitness App user base in this sample showed high engagement, healthy behaviors, and strong retention. There are clear opportunities for further engagement during periods of declining activity or seasonal drops. These insights can guide future feature development, personalized interventions, and strategies to improve user experience and outcomes.