Strava Fitness Data Analytics: Python Insights Report

By: SHRUTI VINOD NAIR

This report provides a complete look at Strava user data. It covers activity (steps), calories burned, sleep duration, intensity levels, and user engagement patterns. The goal is to find useful insights that can help shape product, marketing, and engagement strategies for the Strava platform.

DATA OVERVIEW

1. Data Sources:
    - Daily, hourly, and minute-level activity logs
    - Daily calories burned
    - Daily sleep data
    - Activity intensity (Very Active, Fairly Active, Lightly Active, Sedentary)
2. Analysis Period:
    - One month (May 2016)
3. User Base:
    - 33 unique users, with high retention and activity logging

KEY FINDINGS

A. Activity Patterns
- Daily Steps:
    - Most user-days cluster between 2,000 and 12,000 steps.
    - There is a clear right-skew: a large group of days with low activity (<2,000 steps), and fewer “power user” days reaching 15,000+ steps.
    - Average steps remain stable over time (typically 7,000–8,500 per day).
    - By day of week: Users average the most steps on Tuesdays (~8,000) and the least on Sundays (~6,500).
- Hourly Patterns:
    - Activity ramps up after 6 AM, peaking between 5–7 PM, then tapering off into the night.
- User Segmentation:
    - 9 users are sedentary (<5,000 steps/day)
    - 17 are lightly active (5,000–10,000 steps/day)
    - 7 are active (10,000–15,000 steps/day)
    - 0 are highly active (>15,000 steps/day)

B. Calories Burned

- Most users burn between 1,500–3,500 calories per day, peaking at 2,000–2,200.
- Calories and steps are strongly correlated (corr ≈ 0.57): More steps generally mean more calories burned.

C. Sleep Patterns

- Most users sleep 350–500 minutes/night (6–8.3 hours), with the distribution peaking around 430–450 minutes.
- No strong relationship between steps and sleep (corr ≈ -0.19).

D. Sedentary Behavior and Activity Intensity

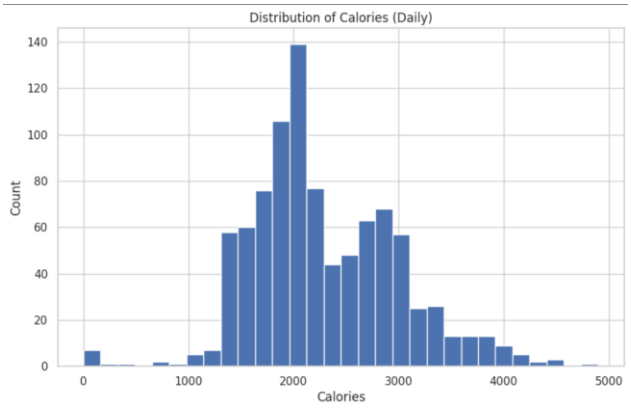
- Sedentary minutes per day are much higher than active minutes for most users (see "Average Daily Minutes by Activity Intensity" plot).
- Very Active Minutes are consistently low, with most users spending the majority of their day either lightly active or sedentary.

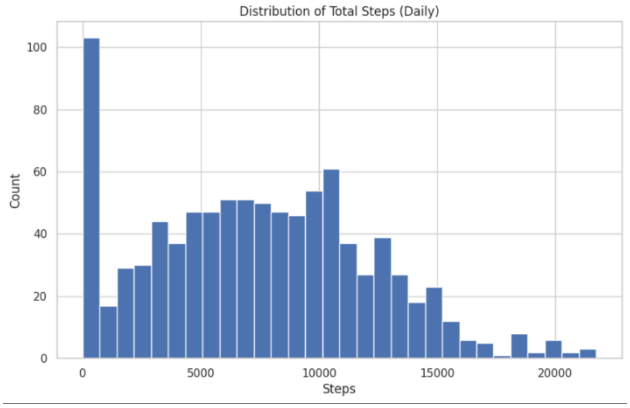
E. User Engagement & Retention

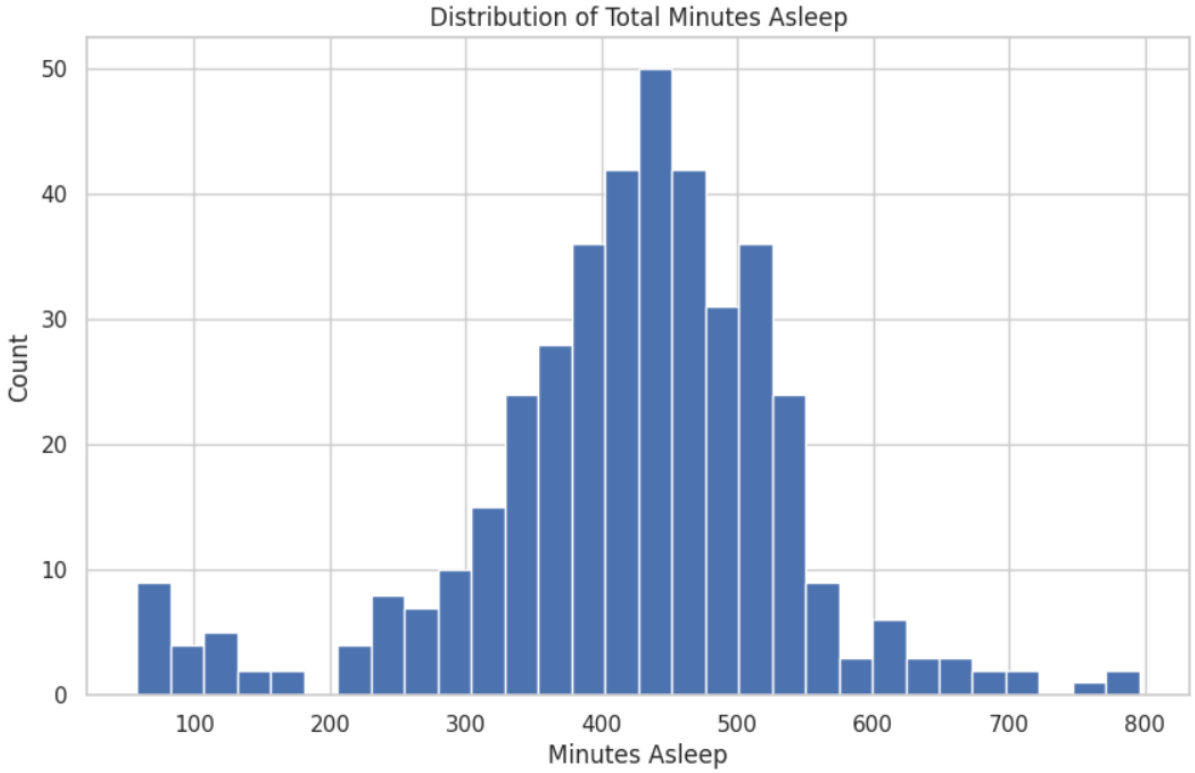
- Median active days per user: 31 out of 31—most users are highly engaged throughout the month.
- Distribution of active days: Almost all users log activity nearly every day, indicating strong retention.

VISUALIZATIONS & DATA HIGHLIGHTS

```
Activity (Daily) — Checking user ID column:
Nulls: 0
Unique IDs: 33
Sample IDs: ['1503960366' '1624580081' '1644430081' '1844505072' '1927972279']
ID dtype: int64

Steps (Daily) — Checking user ID column:
Nulls: 0
Unique IDs: 33
Sample IDs: ['1503960366' '1624580081' '1644430081' '1844505072' '1927972279']
ID dtype: int64

Intensities (Daily) — Checking user ID column:
Nulls: 0
Unique IDs: 33
Sample IDs: ['1503960366' '1624580081' '1644430081' '1844505072' '1927972279']
ID dtype: int64

Calories (Daily) — Checking user ID column:
Nulls: 0
Unique IDs: 33
Sample IDs: ['1503960366' '1624580081' '1644430081' '1844505072' '1927972279']
ID dtype: int64

Sleep (Daily) — Checking user ID column:
Nulls: 0
Unique IDs: 24
Sample IDs: ['1503960366' '1644430081' '1844505072' '1927972279' '2026352035']
ID dtype: int64
```







Sleep Data Summary:

	Id	SleepDay	TotalSleepRecords	TotalMinutesAsleep	TotalTimeInBed
count	4.100000e+02	410	410.000000	410.000000	410.000000
mean	4.994963e+09	2016-04-26 11:38:55.609756160	1.119512	419.173171	458.482927
min	1.503960e+09	2016-04-12 00:00:00	1.000000	58.000000	61.000000
25%	3.977334e+09	2016-04-19 00:00:00	1.000000	361.000000	403.750000
50%	4.702922e+09	2016-04-27 00:00:00	1.000000	432.500000	463.000000
75%	6.962181e+09	2016-05-04 00:00:00	1.000000	490.000000	526.000000
max	8.792010e+09	2016-05-12 00:00:00	3.000000	796.000000	961.000000
std	2.060863e+09	NaN	0.346636	118.635918	127.455140

Weight Data Summary:

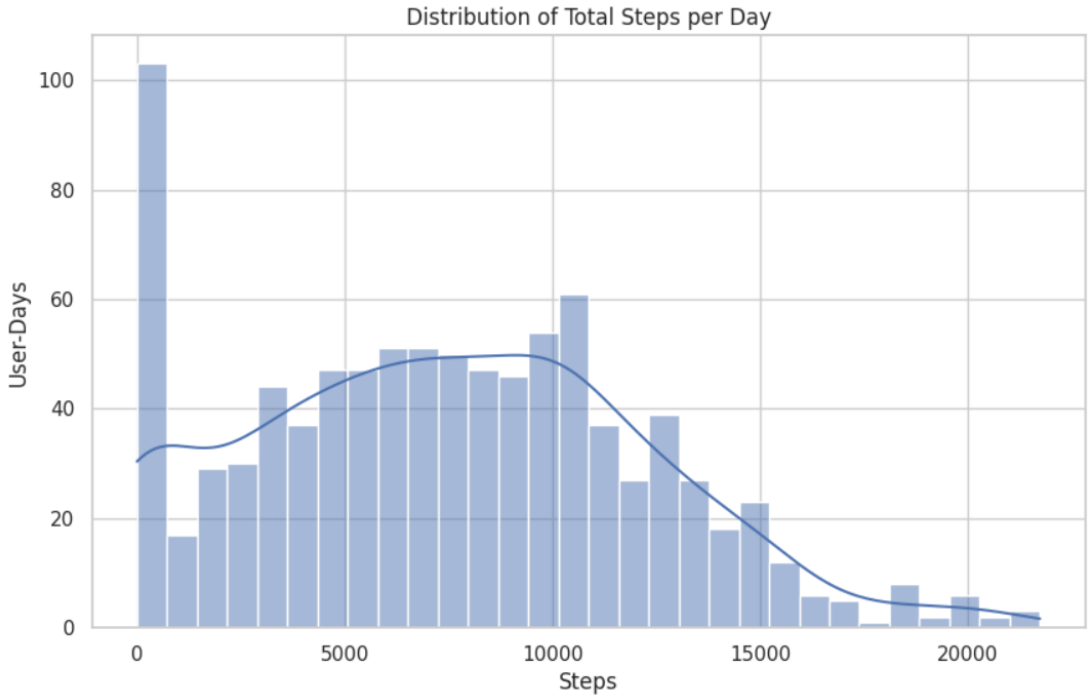
	WeightKg
count	67.000000
mean	72.035821
std	13.923206
min	52.599998
25%	61.400002
50%	62.500000
75%	85.049999
max	133.500000

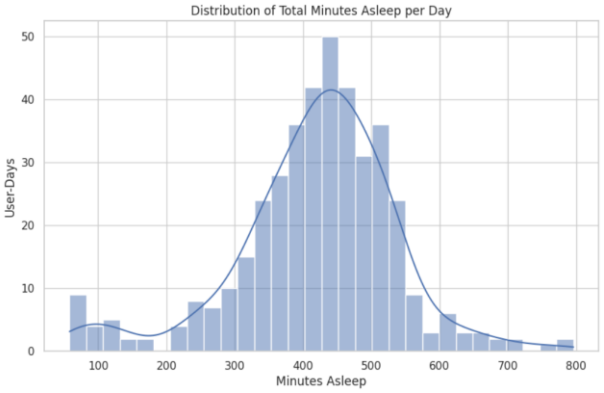
dtype: float64

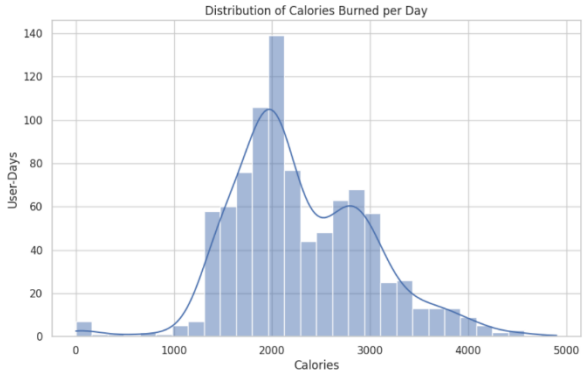
Calories Summary:

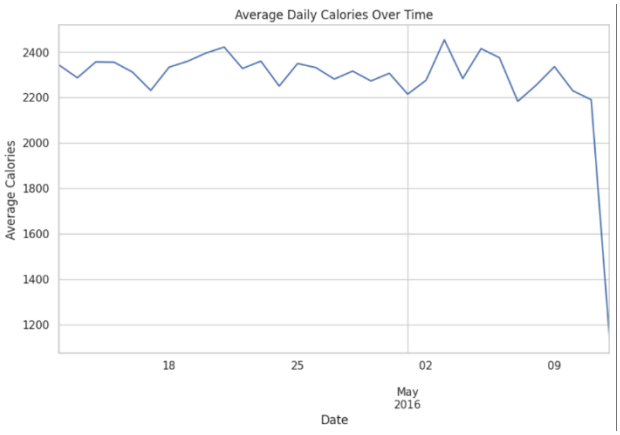
	Calories
count	930.000000
mean	2288.835484
std	704.992971
min	0.000000
25%	1826.250000
50%	2125.500000
75%	2781.750000
max	4900.000000

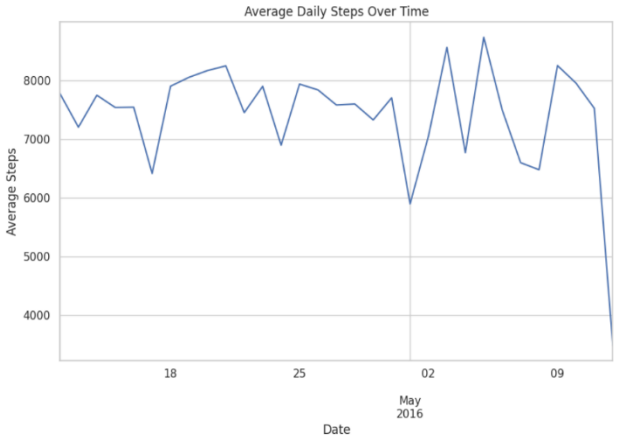
dtype: float64

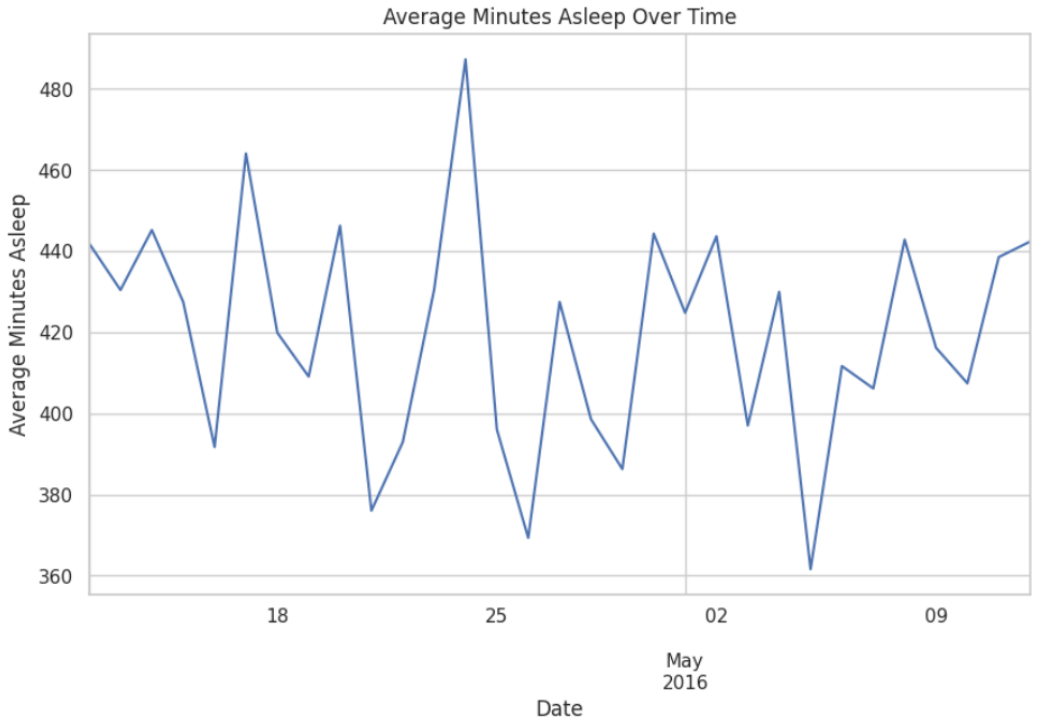


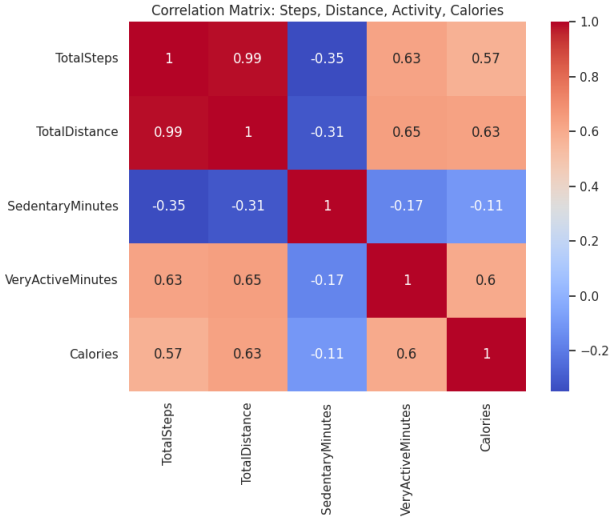

Distribution of Total Minutes Asleep per Day
User-Days
Minutes Asleep
0
10
20
30
40
50
100
200
300
400
500
600
700
800

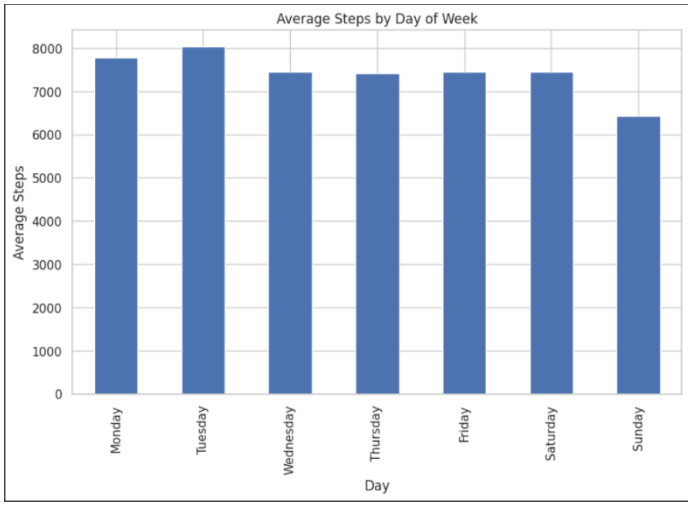

Distribution of Calories Burned per Day
User-Days
Calories
0
20
40
60
80
100
120
140
0
1000
2000
3000
4000
5000

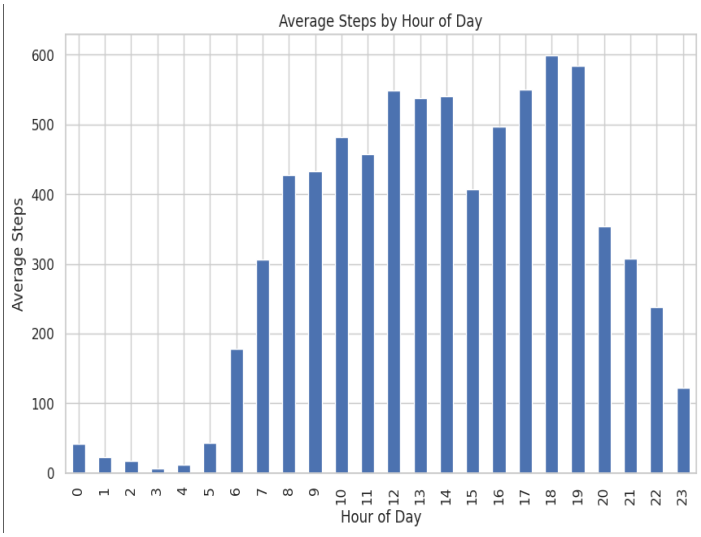

Average Daily Calories Over Time
Average Calories
Date
May
2016
1200
1400
1600
1800
2000
2200
2400
18
25
02
09

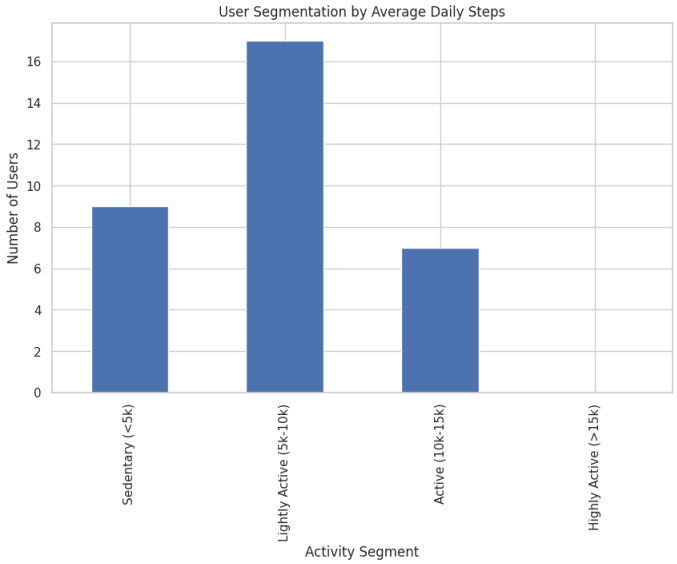

Average Daily Steps Over Time
Average Steps
Date
May
2016
4000
5000
6000
7000
8000
18
25
02
09

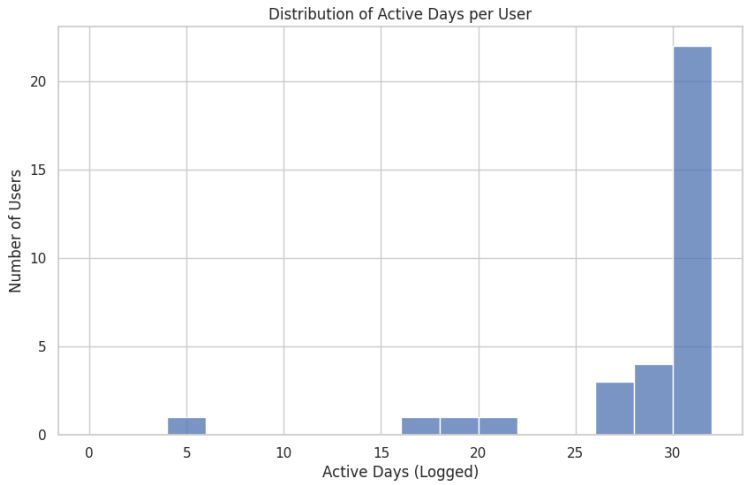

Average Minutes Asleep Over Time
Average Minutes Asleep
Date
May
2016
360
380
400
420
440
460
480
18
25
02
09

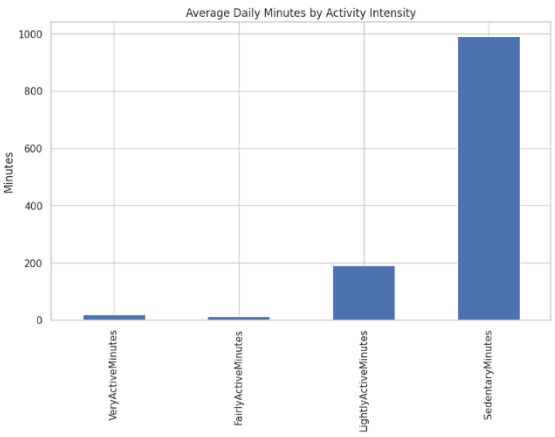

Correlation Matrix: Steps, Distance, Activity, Calories
TotalSteps
TotalDistance
SedentaryMinutes
VeryActiveMinutes
Calories
1
0.99
-0.35
0.63
0.57
0.99
1
-0.31
0.65
0.63
-0.35
-0.31
1
-0.17
-0.11
0.63
0.65
-0.17
1
0.6
0.57
0.63
-0.11
0.6
1
1.0
0.8
0.6
0.4
0.2
0.0
−0.2

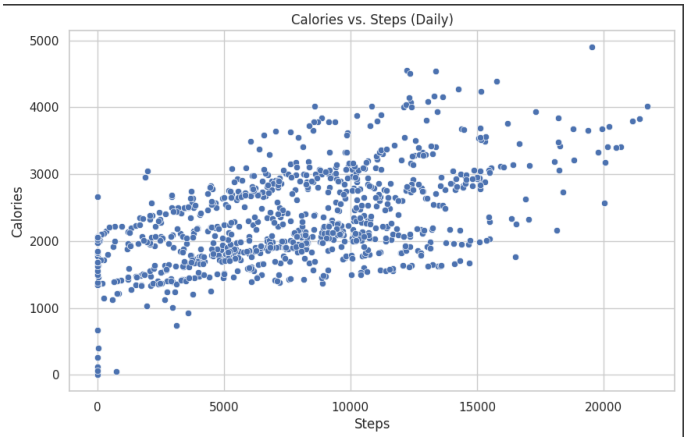

Average Steps by Day of Week
Average Steps
Day
Monday
Tuesday
Wednesday
Thursday
Friday
Saturday
Sunday
0
1000
2000
3000
4000
5000
6000
7000
8000

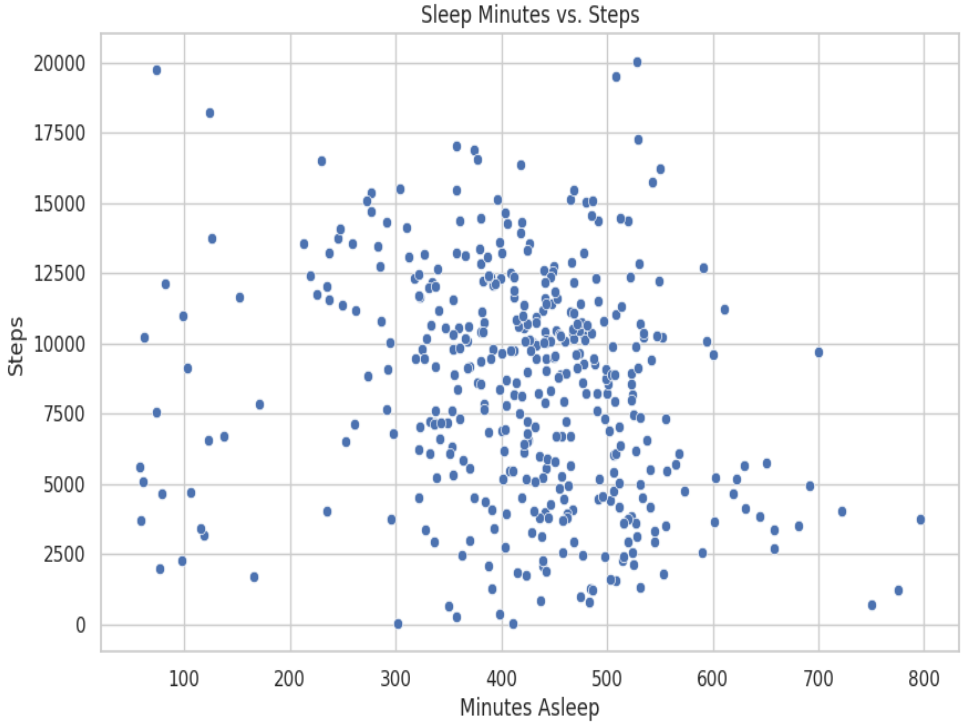

Average Steps by Hour of Day
Average Steps
Hour of Day
0
100
200
300
400
500
600
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23


User Segmentation by Average Daily Steps
Number of Users
Activity Segment
Sedentary (<5k)
Lightly Active (5k-10k)
Active (10k-15k)
Highly Active (>15k)
0
2
4
6
8
10
12
14
16


Distribution of Active Days per User
Number of Users
Active Days (Logged)
0
5
10
15
20
25
30







RECOMMENDATION

1. Target Sedentary and Low-Activity Users:
   - Encourage <5k step users with gentle reminders, easy daily challenges, or motivational content.
   - Use app nudges to prompt users to break up sedentary time, especially mid-day.

2. Capitalize on Peak Activity Times:
   - Schedule app challenges and notifications around the evening (5–7 PM), when most users are active.

3. Support and Progress Lightly Active Users:

- Introduce “level-up” programs or badges for users in the 5k–10k range to encourage them toward higher daily steps.

4. Promote Better Sleep Habits:
   - Offer in-app sleep education, streak tracking, or tips, as many users are below the recommended 7–8 hours/night.
5. Retain Consistent Users:
   - Celebrate daily streaks and reward consistency, since most users are logging daily activity.
6. Continue Outlier and Data Quality Monitoring:
   - Flag and review any extremely high activity/calorie days to ensure device sync and trust.

Thank you for reading this analysis!
