**AVG daily total dailyActivity by day of the week**

This is a wide chart with daily averages of activity levels sorted and ordered by the day of the week on which they occur. Metrics include AVG total distance in km/day, AVG total distance in km at different energy levels/day, AVG time spent in minutes at different activity levels/day, AVG calorie burned/day.

**AVG daily total distance in km per user overall**

The following query shows that there is a sizable variance between the highest (13.21) average and the lowest (0.63). This query is useful to show the variance in average daily distance between the most active user to the least active user.

**AVG daily total distance in km overall for the test pool per day**

This query shows that the test pool daily average is 5.5km/day. However, that is total distance and does not group by activity intensity.

**AVG Active minutes per day of the week when corresponding activity minutes are > 0**

The two linked queries below show some interesting differences, but also some interesting similarities. The column for 'AVG Sed Active Mins/Day' is almost identical with only a slight difference on Thursday. 'AVG Lt Active Mins/Day' shows some moderate differences. 'AVG Mod Active Mins/Day' and 'AVG Very Active Mins/Day' both show significant differences. These differences occur when days with '0' minutes of activity in a certain level is logged.

The question is whether days with a '0' should be treated as an outlier or included in the calculated results.

My initial opinion is that the 0’s should be treated as outliers and eliminated from the AVG results. I figure that there are two ways of logging a 0. One is that the user was not wearing their device and the other is because the user was sedentary all day and doing nothing. I would think that the vast majority of the results would be explained by the former rather than the latter.

There are a few things that are important to note looking at the chart for these results. One is how the sedentary activity is a statistical 0 for distance, and this makes sense. Another is how the remainder of the results are distributed. The average moderately active distance hovers around 0 all week. The average very active distance is roughly between 2.4 and 3 with a spike on Sunday and lows on Monday and Friday. The average light active distance is the highest between 3.2 and 4 with its spikes on Tuesday and Saturday, and with lows on Sunday and Wednesday.

It is also important to note that this chart represents average distances achieved at those levels. Virtually all users will be represented in the sedentary and light active distance, with lesser numbers in the moderately active distance, and even fewer still representing the elite users in the very active distance. The point to keep in mind is that each user will represent a range that incorporates multiple activity time and distance levels.

**AVG active distance in km by (corresponding) active minutes/distance per day of the week**

I wrote two different types of queries that took averages of distances based on activity levels. One type of query counted and averaged distances based on whether the corresponding activity level time was greater than 0. The other type counted and averaged distances based on whether that activity level distance was greater than 0.

Both types eliminated 0’s and treated them as outliers.

\*NOTE\*

In the end, whether the comparison is based on active levels of time or distance, the differences are almost negligible. That being stated, I feel that it is more appropriate to use active levels of distance as the comparative metric in this case.

**What are the AVG steps per hour** – shows a 24-hour distribution of average steps taken per hour on any given day. The data shows that the highest AVG steps per hour was at 18:00 with 599 steps, and the lowest was at 03:00 with 6 steps.

\*NOTE\*

Some observations and inferences taken from the hourly chart are that the average hourly step counts ramp up quickly between 05:00 and 08:00 and stay between 400 and 600 steps per hour until roughly 20:00. The assumption is that people are starting to wind down at the end of the day.

There is also an interesting “dip” at 15:00 which I assume has something to do with afternoon commute times.

**Distribution of AVG calories burned per hour of the day**

This is a 24-hour distribution that shows the average amount of calories burned per hour.

The shape is very similar to the 24-hour steps per hour chart with similar times for high, low, and a 15:00 dip.

**Steps to calorie comparison**

As could be expected, there appears to be a correlation between average daily steps and average calories burned.

The higher the average steps taken corresponds with higher average calories burned.

**Average Steps per day of the week**

Weekly distribution that shows an average number of steps for that day of the week.

No surprise that Saturday is the winner, but that Sunday is the loser, and Tuesday is a CLOSE second place.

**Set up bins for different levels of step counts based on CDC recommendations**

https://www.cdc.gov/media/releases/2020/p0324-daily-step-count.html

https://www.healthline.com/health/how-many-steps-a-day#How-many-steps-should-you-take-a-day?

Inactive: less than 5,000 steps per day

Average (somewhat active): ranges from 7,500 to 9,999 steps per day

Very active: more than 12,500 steps per day

**\*NOTE regarding sleep data\***

Sleep data was not collected daily and/or universally by all users. This is most likely explained by users not wearing their devices when they are sleeping.

**Comparison of total minutes asleep and total time in bed using the sleepDay table**

This gave overall averages of the test pool. The average minutes asleep was 419 (6.98 hours) and the average minutes in bed was 458 (7.63 hours). This equates to a difference of 39 minutes. The data also shows that people spend 91% of their time in bed sleeping.

**\*Don’t make the job harder than it has to be\***