Giulio Benedetti & Cristian Garcia

Data Science for Neuroscientists

Dec 5rth, 2023

Report of Apple Watch Data

Summary

With 52 variables, the Apple watch data offers a comprehensive overview onto a person's habits. The dataset we present in this report looks at several motion factors at different time stamps, which allows us to draw comparisons across the variables and understand how they come together as a piece. This report focuses on the short signal saved as "2023-11-06_09_24_37_Apple Watch.csv". Similar reports on the longer signals are available at the following address:
https://github.com/RiboRings/WatchAccelerometer/tree/main.

Methods

The analysis was conducted on Python v.3. The modules pandas and numpy were used to import and process data, pyplot and seaborn for visualisation, datetime for processing time data. Transformation from time domain to frequency domain was performed with the numpy.fft function, whereas spectrogram was generated with the spectrogram function from scipy.signal.

Results

Among the 58 variables, the monitored information includes:

- logging and timestamp [0, 1, 12, 16, 52]
- 3D location and corresponding accuracy [1:12]
- 3D acceleration [12:16]
- 3D motion (e.g. rotation and gravity) [16:39]
- pedometer and activity type [39:52]

Mean sampling frequency was 50.25 Hz. The sampling intervals seemed to vary more at the beginning and at the end of the measurement, when motion may be less constant. In the centre, sampling frequency became relatively stable at 50.25 Hz. We can also see an average of 20 milliseconds per sample. In total, the measurement was about 36-second long.

Table 1. Comparison across data sets of data recording features.

Data Set	Sampling Freq	Sample Interval	Total Interval
Short Signal	50.25 Hz	0.02 sec	36 sec
Long Signal 1	50.22 Hz	0.02 sec	757.79 sec
Long Signal 2	50.22 Hz	0.02 sec	859.38 sec

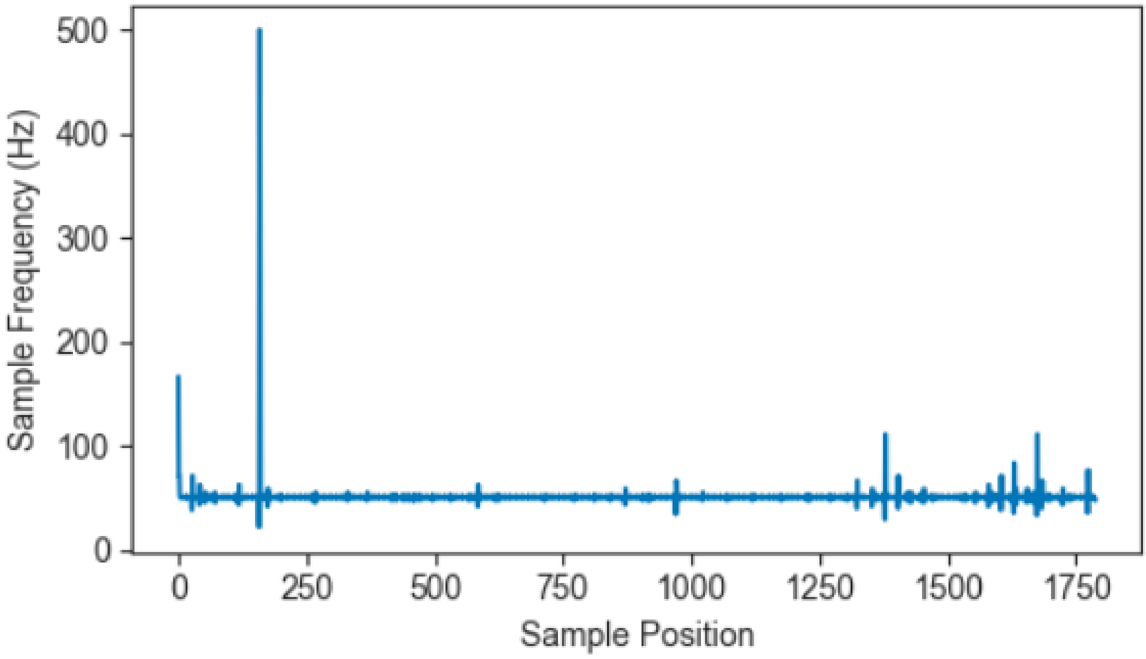
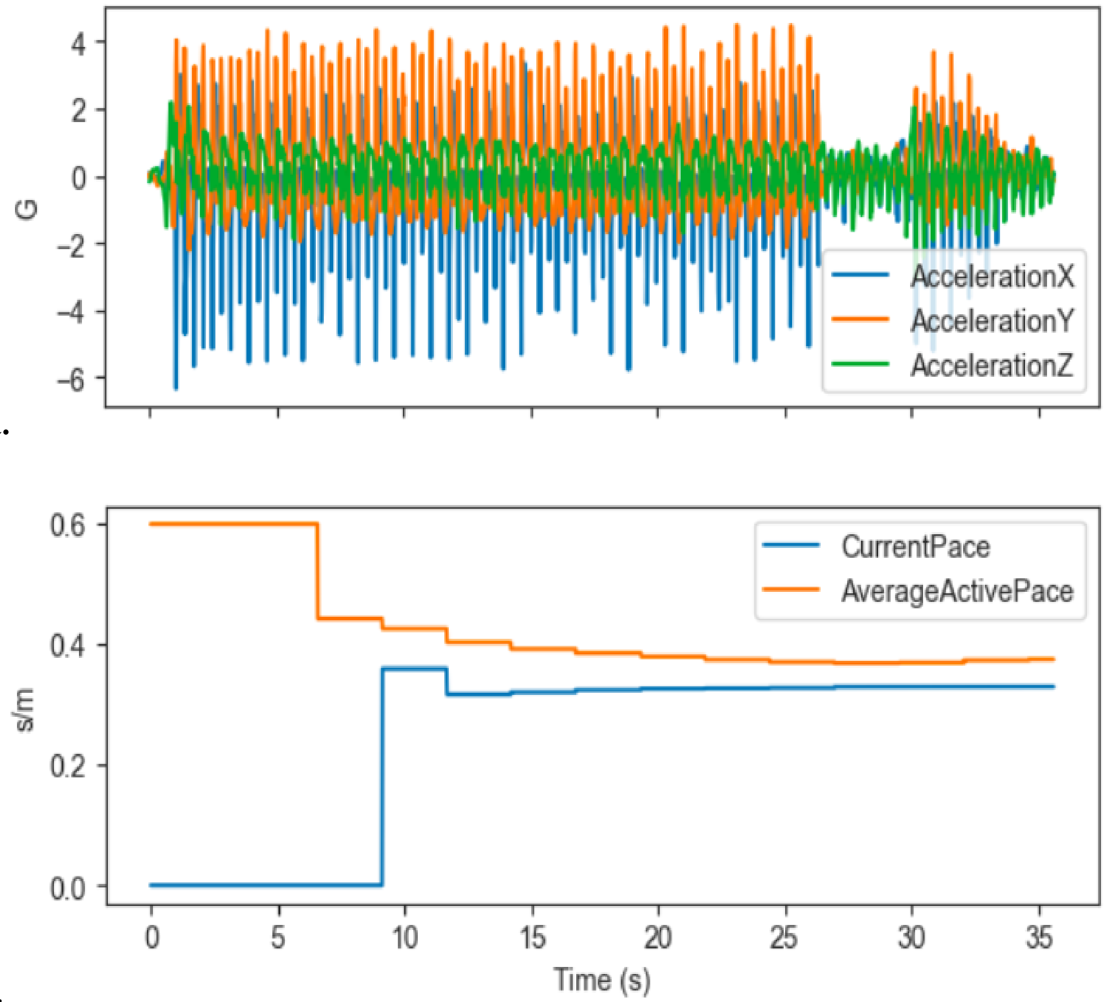


Figure 1. Recording frequency of sample across its position in the short signal dataset. Although consistent around 50.25 Hz, the sample frequency occasionally oscillated with a minimum value of 21.74 Hz and a maximum value of 500 Hz.



a.

b.

Figure 2. Apple watch activity recordings across time. **a.** Across the three coordinates, acceleration oscillates from -6 G to 4 G, with an important rest between 25 and 30 seconds. Acceleration at the X coordinate is notably more negative and at the Y coordinate notably more positive. On the Z axis, acceleration is reliably around 0 G. **b.** Pedometer recordings show a sudden decrease in average active pace followed by an increase in current pace.

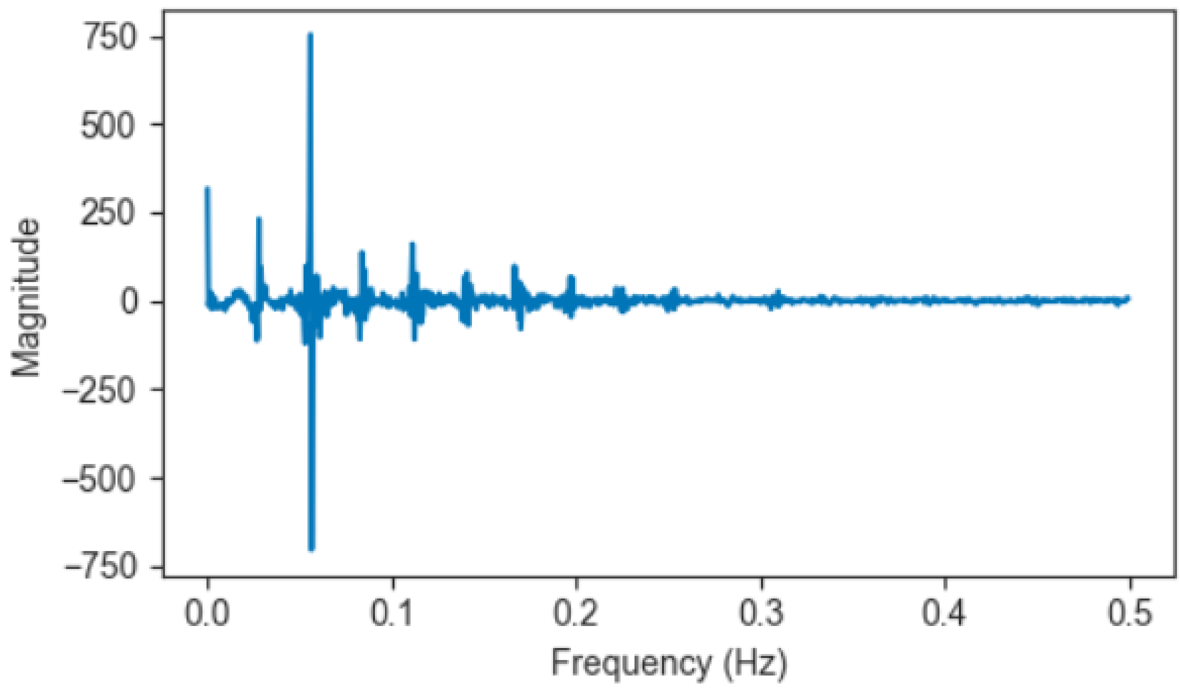


Figure 3. Frequency at the Y coordinate computed through FFT on motion acceleration data.

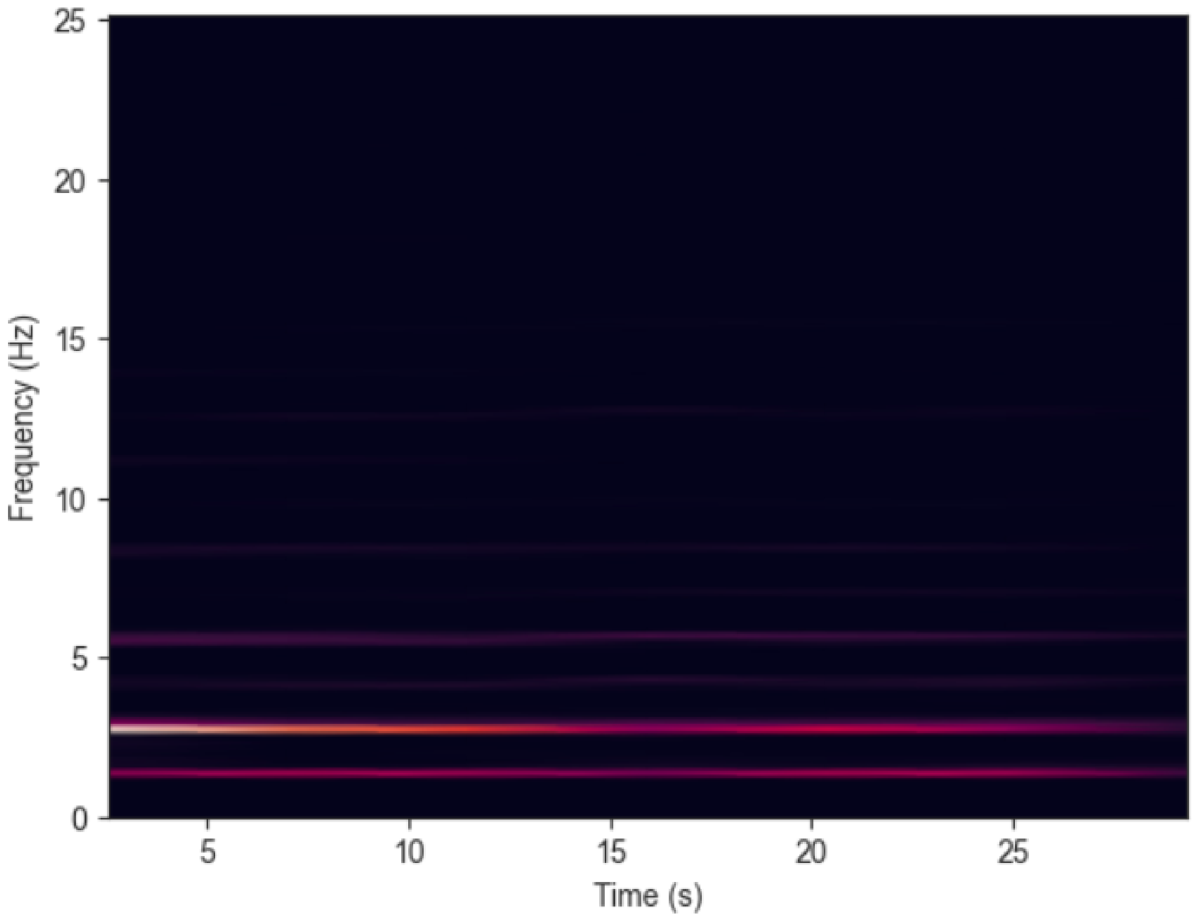


Figure 4. Spectrogram of motion acceleration data on X coordinate.

Table 2. Comparison across data sets of each calculated step frequency derived from motion acceleration data.

Data Set	Step Frequency	Interval Per Step
Short Signal	0.056 Hz	Approx 20 sec
Long Signal 1	0.020 Hz	Approx 50 sec
Long Signal 2	0.024 Hz	Approx 40 sec

Discussion

We define a stationary random process as a set of signals produced over time, each of whose points are random variables drawn from the same distribution regardless of their position in time. Therefore, it follows that the decision of considering a signal stationary or not chiefly depends on the time range under inspection: shorter signals may better reflect stationary processes, whereas longer signals may appear more subject to periodicity.

With that said, it is safe to consider walking and running sessions as stationary random processes, when looked at individually. Here, based on motionUserAcceleration we can infer that the portions of signal from 0 to 25 sec, and between 25 and 30 sec well represent stationary random processes (Figure 2a). However, the signal over its full length contains multiple sessions, and thus it is not stationary.

We define an ergodic process as a set of signals produced over time, whose points from the beginning to the end of each signal are drawn from a distribution that is identical to that of any random variable from the signal. Whereas all ergodic processes are also random stationary, not all of the latter meet the conditions to be ergodic. For example, if the random variables from two signals are drawn from a single bimodal distribution, but each signal contains random variables from only one mode of that distribution, the process is stationary but not ergodic.

With that said, we could consider the signals ergodic only if each signal contained either walking or running sessions, but not both. Because the two are actually combined, the process is not ergodic. However, if we narrow down the signals to one activity type, the process can be considered ergodic. Here, based on motionUserAcceleration the signal up to 25 sec and that from 30 sec on are ergodic with respect to one another. Taken all together, these observations indicate that Different periods of the signal represent different activity states, such as walking and running.

Although motionUserAcceleration(G) was mainly used, motionUserRotationRate(rad/s) also showed similar patterns that may prove relevant for further analysis. In addition, the information on gravity, pitch, yaw, roll and pedometer seemed to vary with activity mode, but most of them showed a non-stationary behaviour.