Nightingale data analysis for White Swan

Zella King

*19/10/2018*

About this document

This report has been prepared for White Swan. It is an analysis of the data produced for Project Nightingale by the UMotif app, based on a download taken on 2018-10-19.

1. Number of user registrations

The total number of unique users who have registered for the app, excluding six project insiders, is 225.

The chart below shows the number of registrations per day from 2018-04-05, the date the app was launched, up to 2018-10-19.

Figure 1a: Number of registrations per day
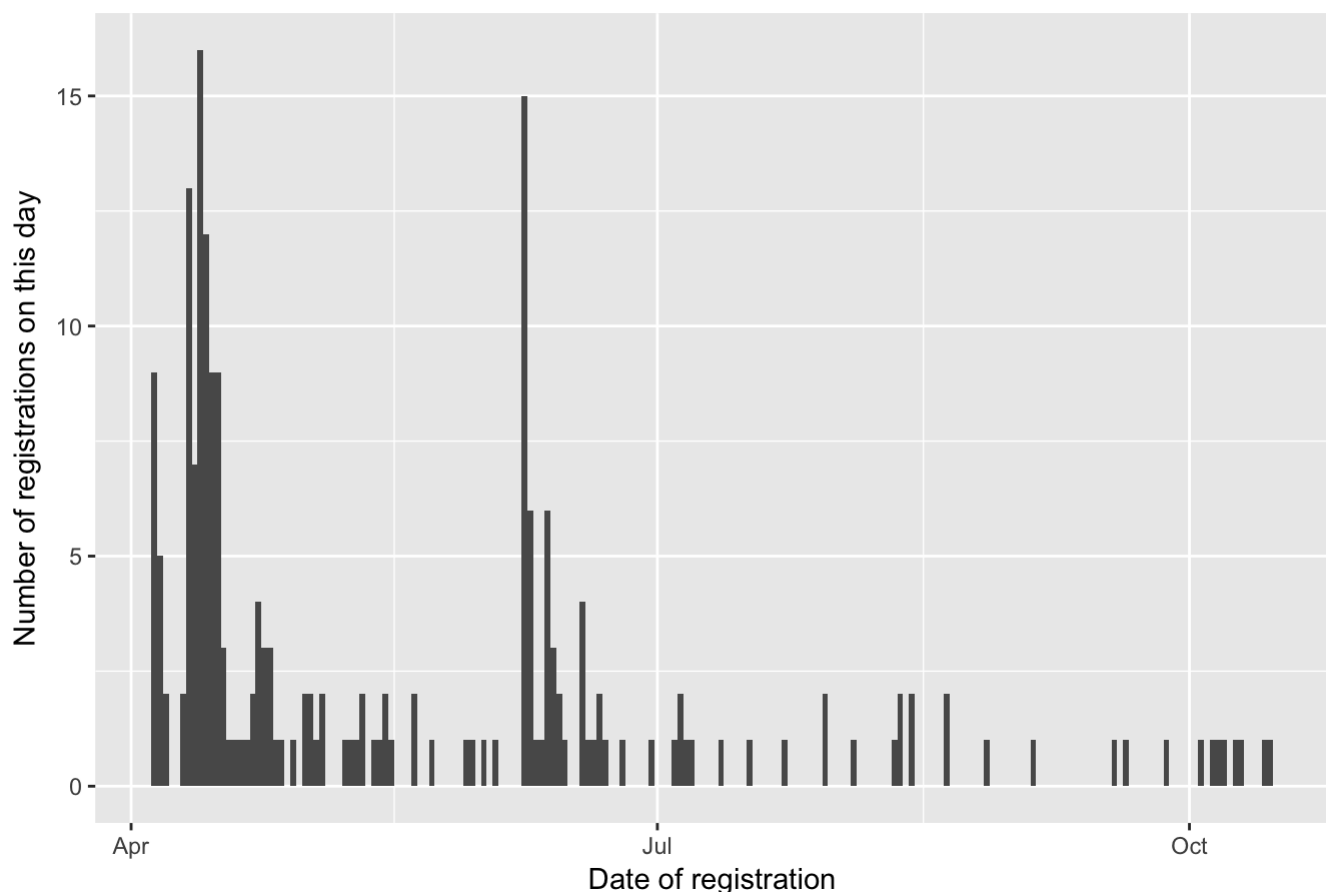
2. Attached devices

The number of users who have attached a device is 96 and between them they have 116 devices, an average of 1.2 devices per user. Some devices, shown as NA, are not identified.

Device type	Number of devices attached

Device type	Number of devices attached
NA	44
fatsecret	1
fitbit	54
garmin	6
moves	2
myfitnesspal	4
strava	5

3. Number of users regularly using uMotif

Although 225 unique user IDs exist on UMotif, data is not complete for all of these, as some have no symptom data. Excluding those with missing data leaves 213 users.

Not all users enter symptoms every day. We can define an active user in different ways:

Have entered symptom data in the last 4 weeks: 104
Have entered symptom data in the last week: 90
Have entered symptom data at least 5 days each week during the time they have been using the app: 99

4. Activity analysis

The following charts show descriptive statistics for activity data (sleep, steps and distance). Where a user has attached more than one device, activity data (sleep, steps and distance) has been averaged across all the devices they attached which provided data on that day.

The charts show the distribution by months as a box-and-whisker chart. The median (middle quartile) marks the mid-point of the data and is shown by the line that divides the box into two parts. Half the scores are greater than or equal to this value and half are less. The pink "box" represents the middle 50% of scores for the group. The upper and lower whiskers represent scores outside the middle 50%. The width of the box represents the number of observations that month. There is little data for October so far, hence the narrow box for month 10. The median of zero may be due to few devices having synced their data so far this month.

Steps

The mean number of steps is 6384 per day across all days of data and users. The median is 5794
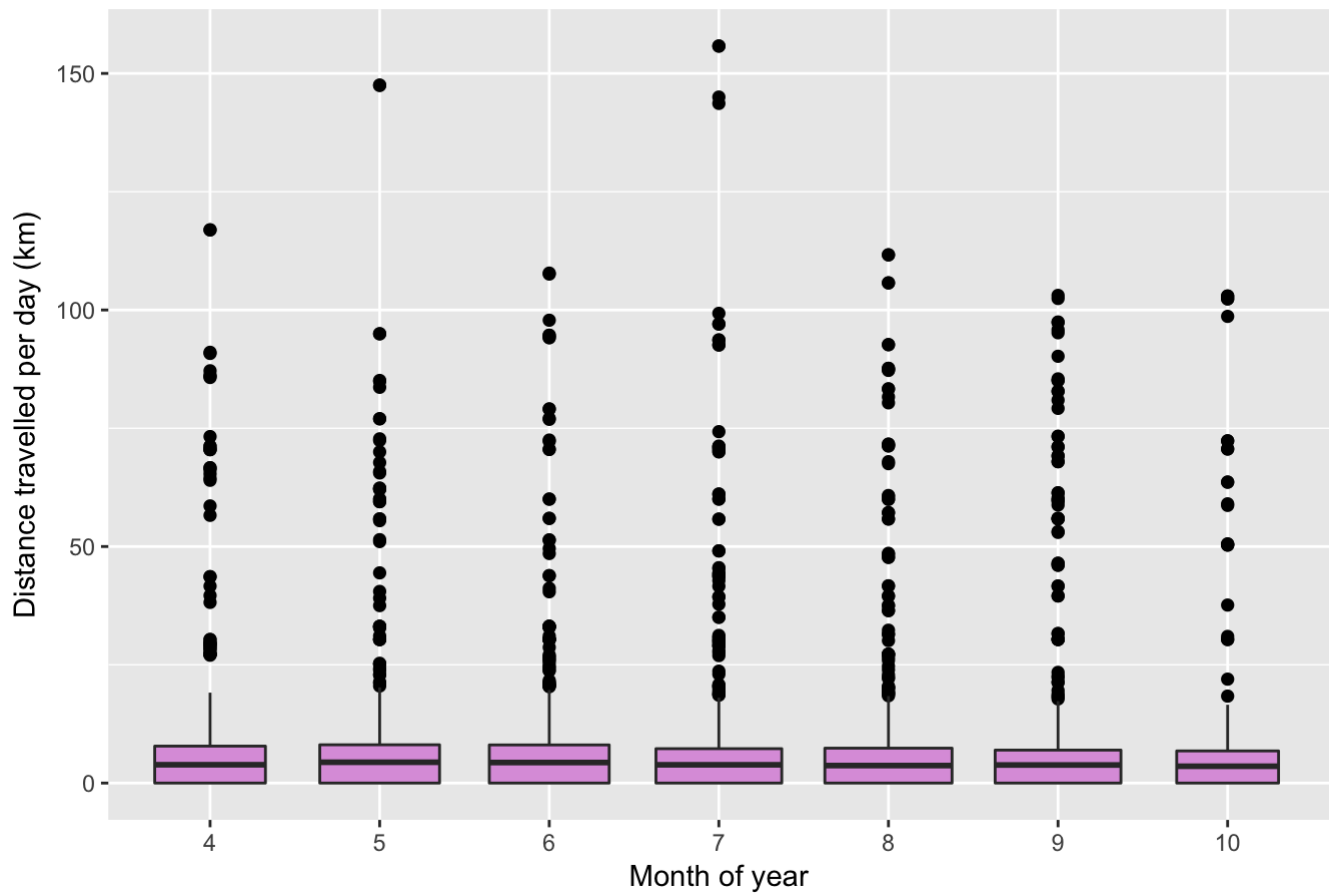
Figure 4a: Steps by month
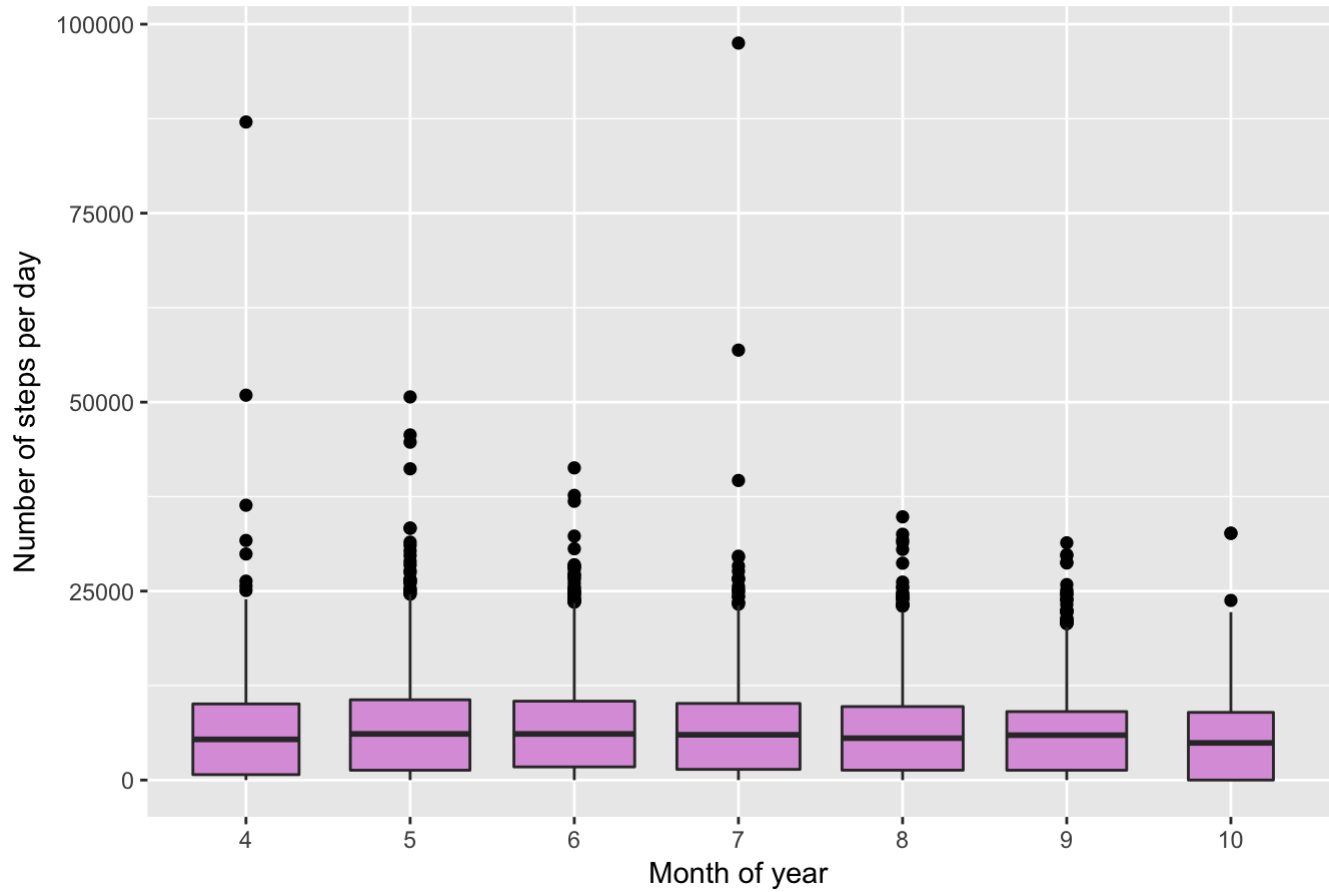
Distance

The mean distance covered per day is 5.95 km and the median is 3.94. The chart below shows the distribution by month Clearly, there are some extreme distances covered.
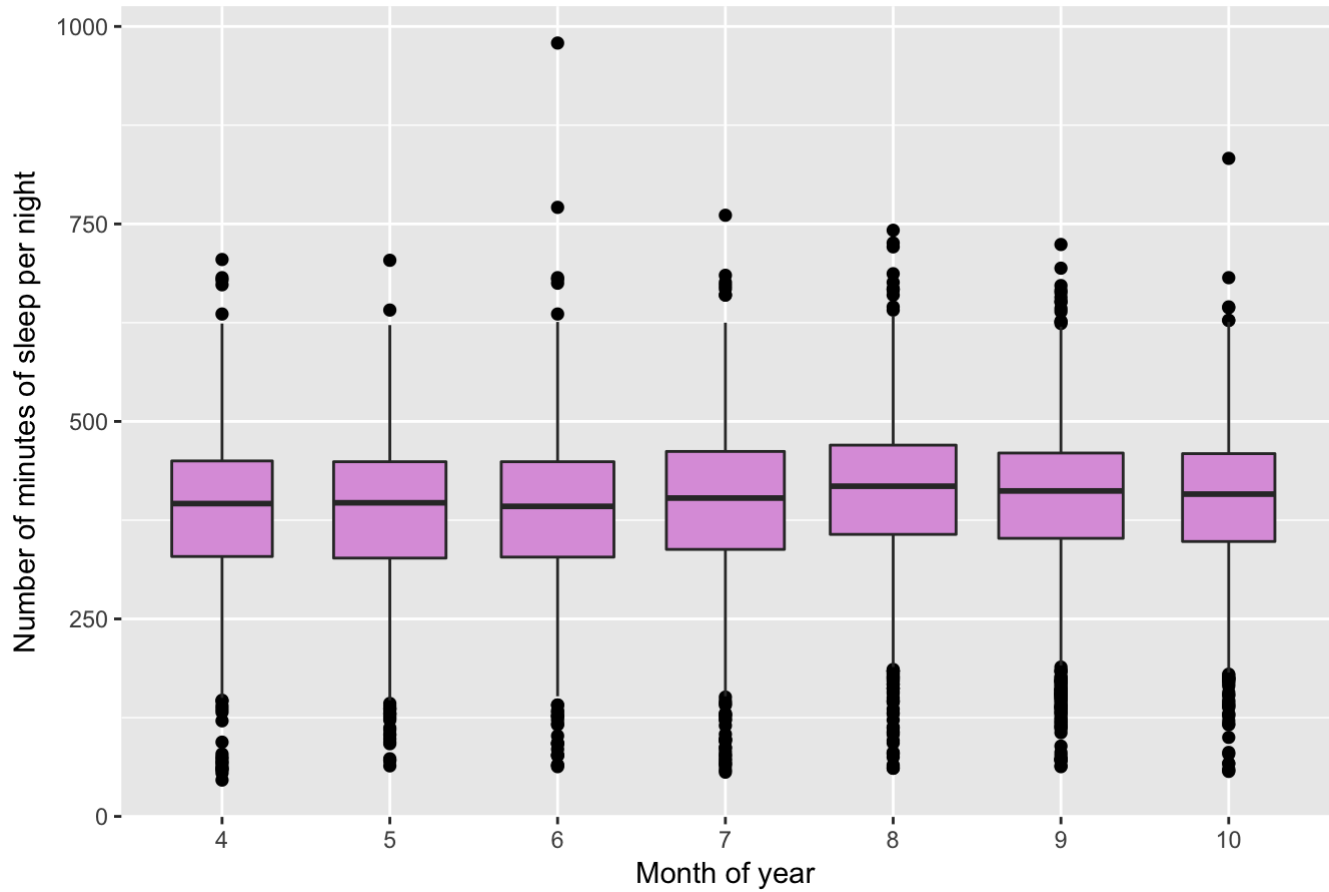
Figure 4b: Distance by month

Sleep

The mean number of minutes of sleep per night is 392 minutes or 6.5 hours and the median is 406 minutes or 6.8 hours. The chart below shows the distribution by month
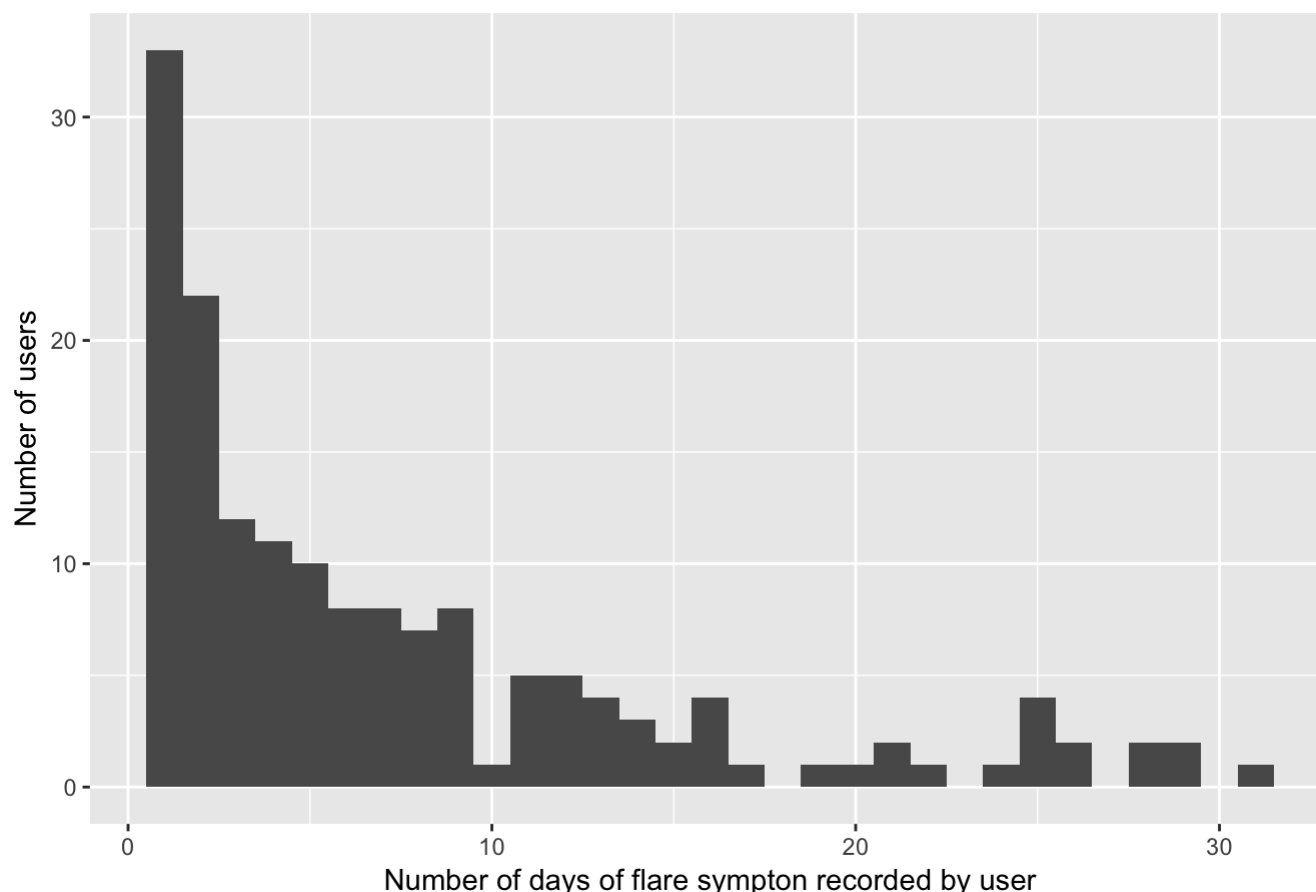
Figure 4c: Minutes of sleep per night by month

5. Flare analysis

In total 161 users reported flare symptoms on at least one day. This chart shows how many days of flare data each user provided so far, ranging from 1 to 31 days.

Figure 5a: No. days flare data provided by user

Flare duration

Flare duration is difficult to calculate as we don't have complete data for each patient every day. It was calculated by taking as start date any day where a patient reported a flare, and counting the number of days of complete data where the flare continued. Note that if the patient had incomplete data with days of flare symptoms missing, the last day in a continuous sequence was deemed to be the last day of the flare. It's possible of course that the flare continued afterward, but the patient did not report this. The median flare length was 1 and the mean was 2.6

The table below shows how the mean and standard deviation in flare duration varied each week for which we have data. (The week is the week in which the flare started and the duration is measured in days.)

Flare duration by week

Figure 5b: Flare duration by week

Week number	No of flares in week	Mean duration	Median duration
14	7	3.7	2
15	22	3.0	1
16	54	2.8	1
17	55	3.5	2
18	44	2.3	2
19	36	2.4	1
20	40	1.9	1

Week number	No of flares in week	Mean duration	Median duration
21	47	2.7	2
22	53	2.5	1
23	45	2.7	1
24	49	2.8	2
25	58	2.8	1
26	40	2.2	2
27	49	1.9	1
28	43	2.5	2
29	36	2.2	1
30	51	3.3	2
31	45	2.2	1
32	46	3.8	1
33	44	2.5	1
34	35	2.7	2
35	44	2.5	2
36	39	3.1	1
37	42	2.2	1
38	40	2.5	1
39	30	2.0	1
40	29	2.9	2
41	40	2.3	1
42	30	1.6	1

The chart below shows the same information as a box plot.

Figure 5c: Flare duration by week
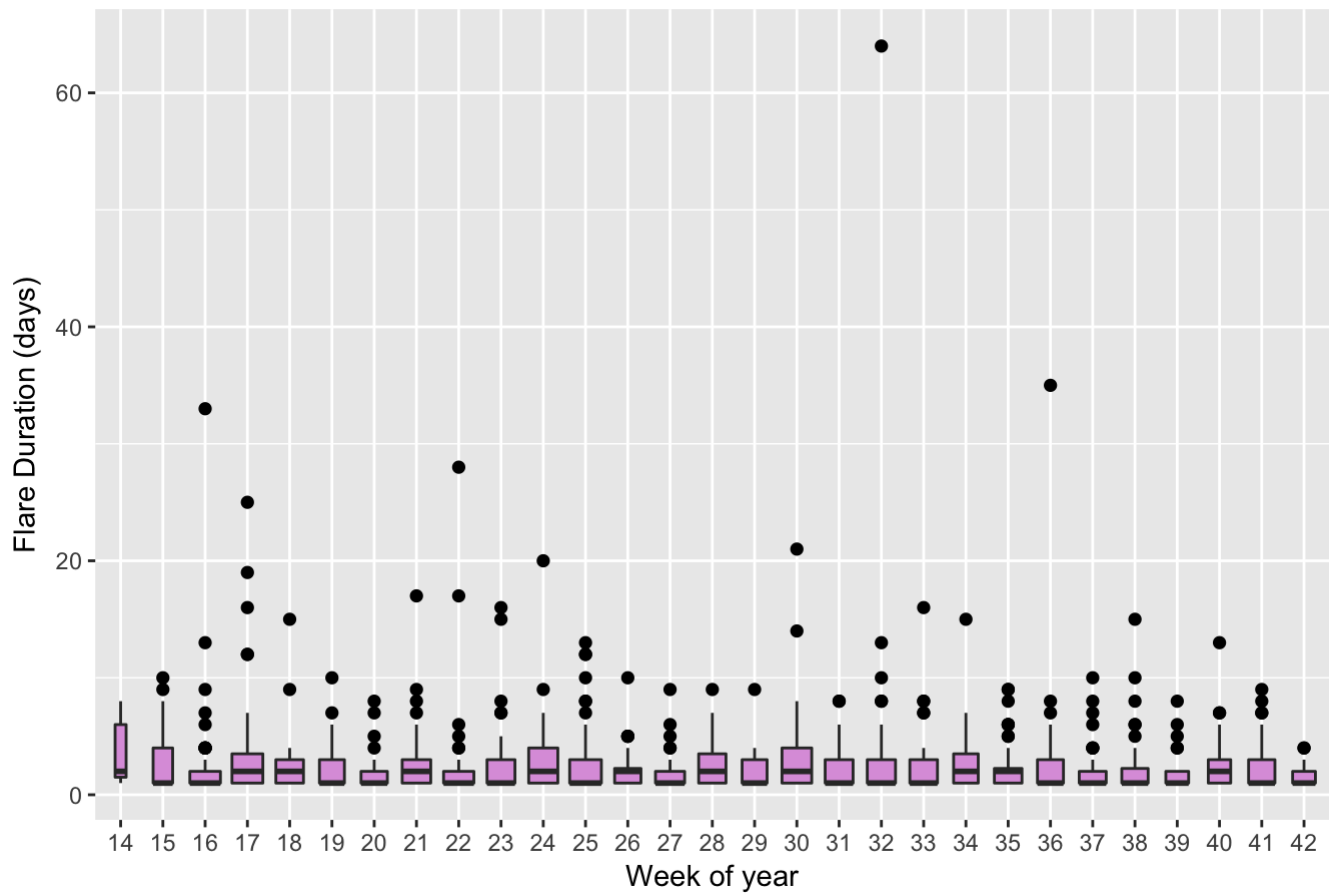
6. Complete user records

A crucial question for further analysis of the relationship between symptoms, activity (captured by devices) and flares is the number of patients who have provided enough flare symptom data to be confident that we know when their flares occured and that we have enough consistent data to do longitudinal analysis. Ideally they would have activity from devices too. The table indicates the number of users with more than 50 days of flare symptoms, and those with more than 50 days of steps data. It shows that 44 users have more than 50 days of both.

Figure 6: Patients with more than 50 days of data

	Patients with <50 days of flare data	Patients with >50 days of flare data	Total no. of patients
Patients with <50 days of step data	84	57	141
Patients with >50 days of step data	28	44	72
Total	112	101	213