DASU20 Final Report

Group 5 - Challenge 1, 2 and 3

Sander Donninger	Niels van Gils	Max de Jongh	Anika Kok	Anne van Staveren
TU/e	TU/e	TU/e	TU/e	TU/e
Eindhoven, Netherlands	Eindhoven, Netherlands	Eindhoven, Netherlands	Eindhoven, Netherlands	Eindhoven, Netherlands
h.a.donninger@student.tue.nl	n.h.t.v.gils@student.tue.nl	m.d.jongh@student.tue.nl	a.k.kok@student.tue.nl	a.s.v.staveren@student.tue.nl

INTRODUCTION

This report contains the three assignments of the course DASU20 (Data Acquisition and Visualisation through Embodied Sensors). These assignments were meant to learn how to acquire and analyse relevant data. The first assignment was about amateur marathon runners and their personal marathon data. With this data, observations and recommendations were made for the individual runners. Assignment 2 required detailed analysis of a GPX file that contained information about a single individual who participated in a marathon. Finally, the third assignment contained a large dataset of GPX files from which a conclusion needed to be drawn that would lead to a concept for a product that could aid the runner during future runs.

ASSIGNMENT 1

For the first assignment five recreational runners of the Eindhoven Marathon of 2017 were researched. These runners were selected based on the time they needed to complete the marathon. Research was conducted into their pace, their facial expression, posture, technique, their equipment and their history in other marathons. This was done using pictures provided by the Eindhoven Marathon database and by using other data found on other relevant websites. From the analysis can be concluded that all of the runners have a steady speed that drops slightly as the marathon progresses. The average speed of the analysed runners is approximately 3 km/h slower than the general average speed. The analysis of the photos shows that the runners wear appropriate clothing but that the technique of the runners can be improved. This most likely influences the final result of the runners, because a good technique potentially improves the consistency in pace and increases the speed. The main learning point from this assignment was how to combine different data points from multiple data sources. In this case the quantitative data of the marathon, combined with quantitative data from other online resources and more qualitative data in the form of photos and videos.

ASSIGNMENT 2

The second assignment included an in-depth analysis of one detailed dataset provided by the course. Because of the level of detail of this dataset, there was an opportunity to do specific and elaborate research. When analyzing the dataset, different indicators have been analyzed; The heart rate and speed of the participant, the outside temperature, and the elevation.

Heart rate

When looking at the heart rate of the participant, some interesting data changes can be seen. In the beginning of the marathon, the heart rate rises from 100 to 145 in two minutes. This shows the start of physical effort and can also have to do with possibile adrenaline.

After 12,5 km within the marathon, the heart rate decreases from around 160 to an average of 140 with highs and lows. This continues until around 36 km. After that the heart rate rises to an average of around 150, that can indicate a final sprint. As could be expected, the heart rate decreases after the finish of the marathon.

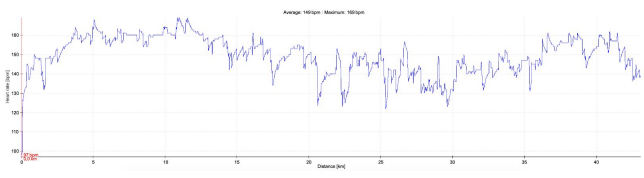

Figure 1, Heart rate measured during the run

Temperature

As can be seen from Figure 2, the temperature has decreased drastically over time. At the start of the marathon, the measured temperature was 27° C. Over the course of the marathon, this measured temperature dropped to 18° C, which is quite a significant difference.

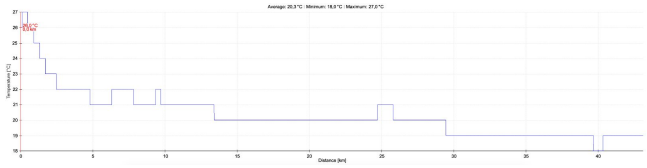

Figure 2, Temperature measured over the distance during the run

Despite the fact that it was rather difficult to find the exact start time, a video was found about the marathon in Eindhoven from 2017. In this video it can be seen that the marathon took place around sunset, which would explain a drop in temperature. However, the data gathered from the KNMI database contains completely different temperature values, namely a maximum of 15,1° C and a minimum of 9,9° C, as can be seen in Figure 3.

Daggegevens van het weer in Nederland

Het weer op zondag 8 oktober 2017 te Eindhoven © KNMI

Temperatuur		**Normaal**	**Neerslag**		
Gemiddelde	12.0 °C	12.4 °C	Hoeveelheid	1.1 mm	
Maximum	15.1 °C	16.6 °C	Duur	0.9 uur	
Minimum	9.9 °C	8.2 °C			
Zon, bewolking & zicht			**Wind**		
Duur zonneschijn	2.4 uur		Gemiddelde snelheid	2.1 m/s	(2 Bft)
Rel. zonneschijnduur	21 %	33 %	Maximale uurgemiddelde snelheid	3.0 m/s	(2 Bft)
Gem. bedekkingsgraad	7 octa's		Maximale stoot	6.0 m/s	
	Vrijwel geheel bewolkt				
Minimaal zicht	1.2 km		Overheersende richting	270 °	(W)
Relatieve luchtvochtigheid			**Luchtdruk**		
Gemiddelde	94 %	86 %	Gemiddelde luchtdruk	1017.9 hPa	

Figure 3, Weather data of the day of the marathon [1]

As the KNMI has advanced measuring equipment and experienced employees, we expect the difference in data to be an error that occurred in the dataset of the runner. This difference in measured temperatures can be caused by multiple factors. One possibility would be inaccuracy of the runner's measuring device. Another possibility would be incorrect placement of the temperature sensor. As the values are a lot higher than the actual outside temperature (and assuming the measuring device is accurate), it is likely that the sensor was placed in a way that it has picked up heat generated by the runner's body (e.g. on an arm). This would also explain the steep drop at the start of the graph as there is no wind passing the sensor and all runners are warming up and standing closely together. As soon as the marathon starts, the runners spread out and wind starts to pass the temperature sensor, thus dropping the temperature rapidly.

In short, both the decreasing outside temperature and the possible inaccurate measuring equipment and/or technique can explain the strong decrease in temperature.

Speed

The average speed of the runner for the whole marathon is 11 km/h. This means he was slightly slower than the average of all participants of the marathon.

His maximum speed was 35,1 km/h and as this does not seem to be a plausible speed, further research on this speed will be done. His pace was on average 5:28 min/km.

There are some peaks in the data that need further analysis to determine what happened here, and why such extreme speeds were measured, this can be seen in Figure 4.

19,5 km/h at 0,5km; At the moment of this data point, the runner was underneath a highway, so an explanation for this speed increase could be a temporary loss of signal for the GPS. The same goes for the data points 20,5 km/h at 4,1km and 35,1 km/h at 20,2 km. The latter of which the runner was running underneath a railroad.

14,4 km/h at 36,9 km; Even though this is a clearly noticeable increase in speed compared to the entire run, there is no similar explanation like in the previous cases where the speed increased rapidly. Also, this speed is manageable for a runner, so this might be a small sprint to catch up, or overtake someone.

There are also some noticeable low speeds measured. For instance: 6,5 km/h at 1,4 km. This was at the very beginning of the marathon, even though the runner had found his pace earlier, the slower speed might be caused by running in a large group of people and being held up by them.

1,8 km/h at 10,5 km; At this point the runner was almost standing still. Possible explanations for this are either a water break, or some other reason for him to have a break and slow down by a lot for a very short amount of time.

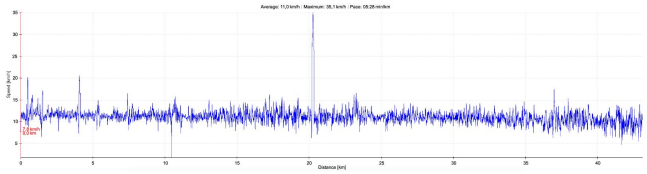

Figure 4, Speed in km/h over the distance during the run

Elevation

The elevation map is not very accurate, as it claims that the overall elevation difference is 19 meters (370m-351m), while in reality this is not the case. There is likely a small deviation in the measuring device, as the effect can only be seen long-term. The short-term elevation differences are caused by going either through tunnels or over bridges. The very small increases and decreases in the graph are caused by the body moving while running, so they can be ignored when looking at the effects of elevation on running.

The relationship between elevation and heart rate and elevation and speed do not show any outstanding or surprising results. Figure 5 gives a visual representation of the data measured on elevation. It seems that the runner speeds up while descending and slows down during ascending, but theses effects average out to become irrelevant for the overall speed. During the ascend at 17.5km there is a clear link between the increase in elevation and the heart rate, as the heart rate starts rising when the ascend starts, and stops rising when the descend starts. For the other elevation peaks it is nearly impossible to see the effect of elevation on heart rate in the graphs, which could be caused by there being no effect, the effect being minimal or the measurements not being accurate.

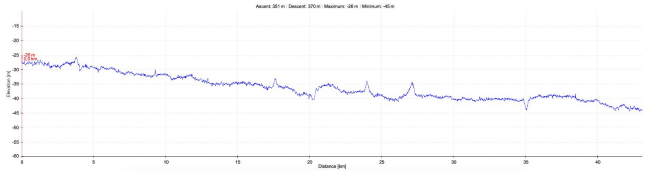

Figure 5, Elevation over the distance during the run

How can individual runners be supported given their real-time performance?

Individual runners can be supported by giving them insight into data that is gathered live, during the activity but also after the activity. The most important aspect is the meaning given to this data and the corresponding conclusions that are drawn. This can be hard for inexperienced runners that do not know what standard values are for their body type, age, and fitness level. To help this target group, it is important to gather their data and then translate these with the correct parameters, like the age and fitness level, so that useful conclusions can be drawn.

Age	Target HR Zone 50-85%	Average Maximum Heart Rate, 100%
20 years	100-170 beats per minute (bpm)	200 bpm
30 years	95-162 bpm	190 bpm
35 years	93-157 bpm	185 bpm
40 years	90-153 bpm	180 bpm
45 years	88-149 bpm	175 bpm
50 years	85-145 bpm	170 bpm
55 years	83-140 bpm	165 bpm
60 years	80-136 bpm	160 bpm
65 years	78-132 bpm	155 bpm
70 years	75-128 bpm	150 bpm

Table 1: Maximum and Target Heart Rate per age (average)[2]

For example, Table 1 shows the maximum and target heart rate which as can be seen differs per age. Besides, there is also a difference between men and women. Because so many factors play a role in determining what normal data is, interpreting personal data can therefore be difficult.

A solution would be a tool that is equipped with personal data so that it can subsequently provide personal information and advice. The more personal data, such as gender, weight, fitness level, etc. are provided, the better the translation of the tool and the more accurate the feedback will be for that specific person.

ASSIGNMENT 3

In the third and final assignment of this course, the challenge was to analyze information from a large and detailed dataset provided to us. With this analysis the goal was to extract valuable information from the data, in order to create a concept that helps socially engaged runners to improve themselves.

With this data we want to create a visualization based upon this socially engaged profile, which will benefit the runners who fit this particular profile.

User group

Runners can be divided into four categories: individual fitness, individual competitive, social competitive and social. For this assignment, the social/socially engaged target profile was chosen because this seems to be a profile

that is generally less focused upon. With the current technologies and solutions for runners, it seems like the main focus is on socially competitive or individually competitive runners. There are however, far less products that help social runners that are not competitive.

Ethical considerations

Because working with personal data was required in this course ethical considerations have been made to protect the privacy of the runner that provided the data. Within these assignments, there was no use of online data converters that have a risk of saving the data for own purposes. In addition, the data was only shared via secure platforms.

Because the data was provided in GPX format and was not converted to different formats like CSV, a GPX viewer named GPXSee was used to analyse the data.

Data subsets

The complete dataset that was provided consisted of 196 detailed files with information on location, elevation, speed, heart rate, cadence and temperature. In order to analyse the data, the files within the dataset were first categorised into individual and social runs. We did this by using external data from the website of a running group named Blixemsnel[3] because we found out that the researched runner is a member of this running group. This additional data resulted in specific insights on the runners of the group, their competitions, training schedules and routes.

To determine whether the participant was running alone, or took part in a group training, we compared the training schedule on the website of Blixemsnel with the information of the runs in the provided dataset. We have assumed that the runner participated in the Blixemsnel training if he had the same type of run, on the same day, at the same location as the schedule stated.

Figure 6 shows how we categorized the dataset so that we could determine which runs were important for our analysis.

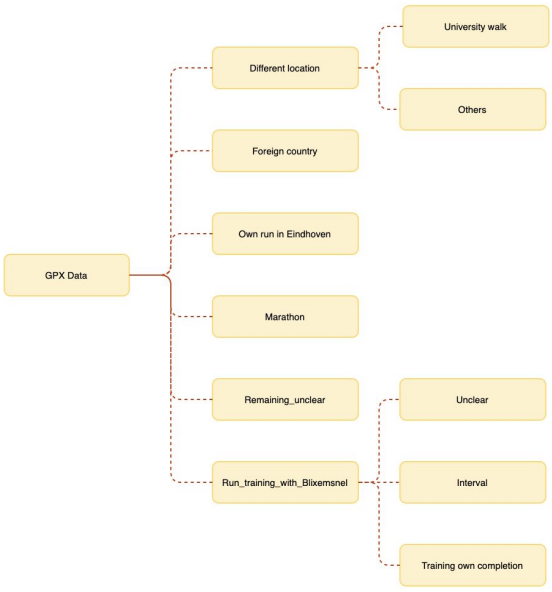

Figure 6: Subsets of the data

Data visualization

After categorising the data, we had two final subsets of the data: one contained the structured trainings with Blixemsnel and the other contained what we categorized as individual runs. These two subsets were loaded into the graphs that can be seen in Figure 7 and 8. As the subset that contained individual runs consisted of too many files, only 12 files were selected for the visualisation in order to make sure the graph would not be too cluttered. Both of these visualisations have been made using GPXSee.

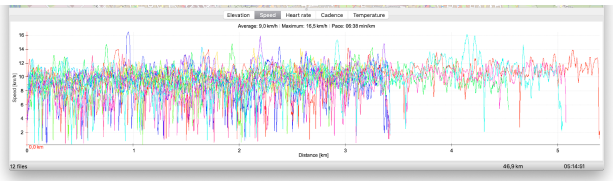

Figure 7: Visualisation of the individual runs

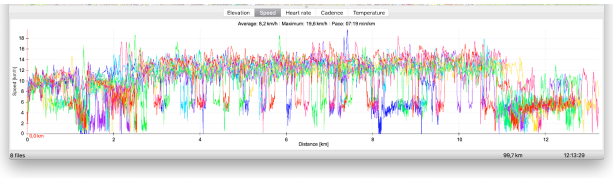

Figure 8: Visualisation of the runs with Blixemsnel

Findings

From the data can be concluded that the participant has a consistent training schedule that includes multiple runs per week. Even when the runner is abroad, the training continues. In addition, it can be concluded that there is a lot of variation in the training sessions of the participant. He runs different distances, at varying locations. He also varies between interval and duration runs within his schedule. The starting point of the runs is often the same, but when analysing the running routes, it was found that the routes are almost never identical.

By analyzing the website and training schedules of the running group Blixemsnel and combining this data with the provided dataset, it can be concluded that the runner is a light-interval trainer at Blixemsnel. The website of Blixemsnel states the following:

Almost every training is an interval training, because they yield a lot of progression and they are the most difficult to do individually. Duration runs are left for people to do individually.

("Loopgroep Blixemsnel," n.d.)

Based on this information from the website it can be assumed that long duration runs will mostly be individual, while the interval trainings are presumably executed in training sessions with two or more runners.

When comparing the assumed social runs and individual runs, differences between these two types of runs can be found. It can be concluded that the individual runs are generally short and less structured than the social runs. The scheduled training sessions from the running group are usually around 13 km long and structured, including a build-up, interval and cooling down.

In conclusion, we found that running in a social context seems to improve the structure of a training, which in return is beneficial for the runners improvement.[4] Increasing the opportunities for runners to engage in social runs therefore potentially improves their performance.

Design

Based on our research we found that runners have more structured runs when they run together with someone else, and more structure in a training usually results in a more efficient training.[4] We assume that when they run with someone with a lower level, they can not perform at their maximum level and thus can not improve their results in the most efficient way. For that reason, it is important that a runner is matched with someone that performs at the same level.

Based on our research, it can be concluded that social runners benefit from structured training sessions with at least one other runner. To facilitate this type of training and help runners find someone else to run with, we have created the concept MatchRun.

MatchRun is an app that offers runners the possibility to run with others in the same neighborhood. When creating an account in the app, the user can set goals that he or she wants to achieve. The user can then work together with other users of the app to achieve these goals. This can be done by finding runners that live closeby, or by selecting his or her running club. When using the app, a user can manually or randomly select another runner and invite them to go for a run. This will result in two or more users that run together and motivate each other to achieve their goals. After the goals of the runner are met, new goals can be set to keep improving the running. Since the app focuses on running *with* another user of the app, instead of competing *against* them, this app focuses on the target group, the socially engaged runners instead of the competitive runners. Figure 10 shows the different visuals of the app called MatchRun.

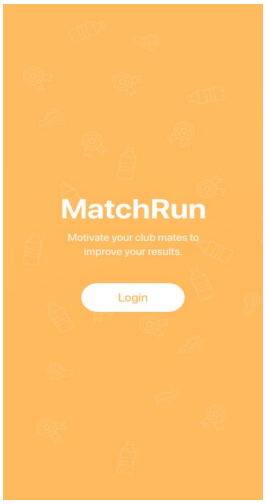

Figure 10.1: Login screen of the MatchRun app

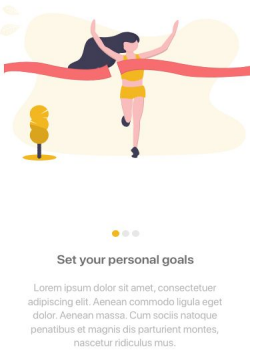

Figure 10.2: Set personal goals with MatchRun (onboarding)

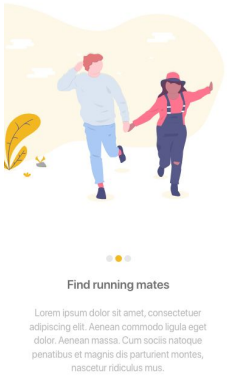

Figure 10.3: Find running mates with MatchRun (onboarding)

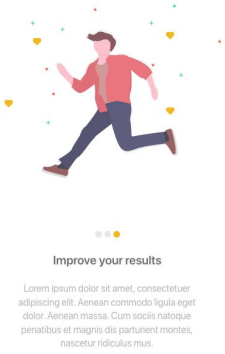

Figure 10.4: Improve your results with MatchRun (onboarding)

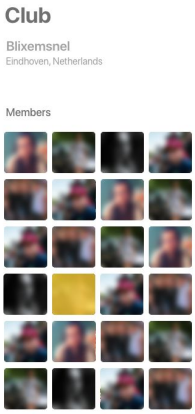

Figure 10.7: Random matching after choosing this option at screen shown in Figure 10.6

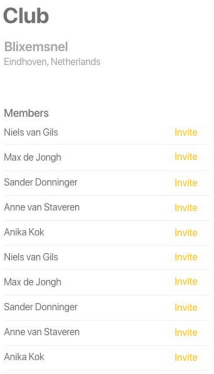

Figure 10.8: List of fellow club members after choosing 'invite' option at screen shown in Figure 10.6

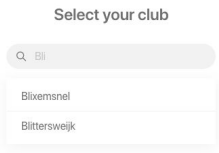

Figure 10.5: Select your club

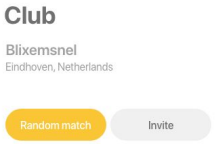

Figure 10.6: Choose whether to match randomly with a fellow club member, or to invite a fellow club member on a run

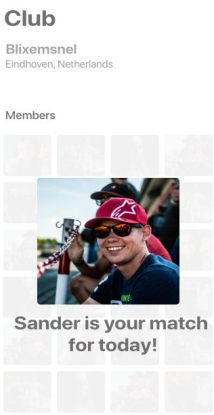

Figure 10.9: Match made in MatchRun

REFERENCES

[1] Knmi. (n.d.). Daggegevens van het weer in Nederland. Retrieved from https://projects.knmi.nl/klimatologie/daggegevens/index.cgi

[2] Know Your Target Heart Rates for Exercise, Losing Weight and Health. (n.d.). Retrieved from https://www.heart.org/en/healthy-living/fitness/fitness-basics/target-heart-rates

[3] Loopgroep Blixemsnel. (n.d.). Retrieved from https://www.blixemsnel.nl/

[4] Zinner C. (2016) Training Aspects of Marathon Running. In: Zinner C., Sperlich B. (eds) Marathon Running: Physiology, Psychology, Nutrition and Training Aspects. Springer, Cham