

Data Analytics Bootcamp @ Neue Fische

# How to improve endurance sports performance

Capstone Project Presentation

by Aisling Horan, Katharina Kreß & Nono Weinzierl

3rd March 2025

# Amateur Triathlete

Our team member Nono is an amateur triathlete. Triathlon is a multi-sport where athletes swim, bike and run consecutively.



Swim

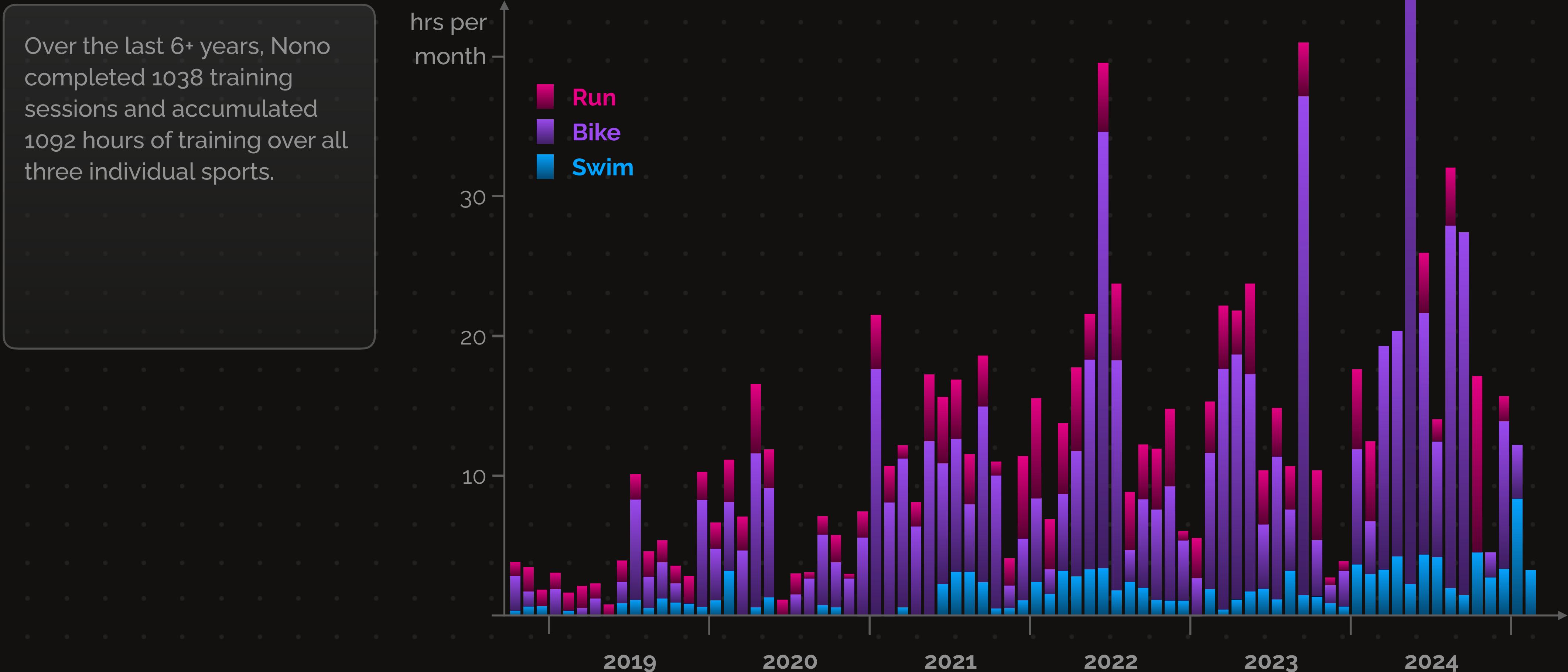


Bike



Run

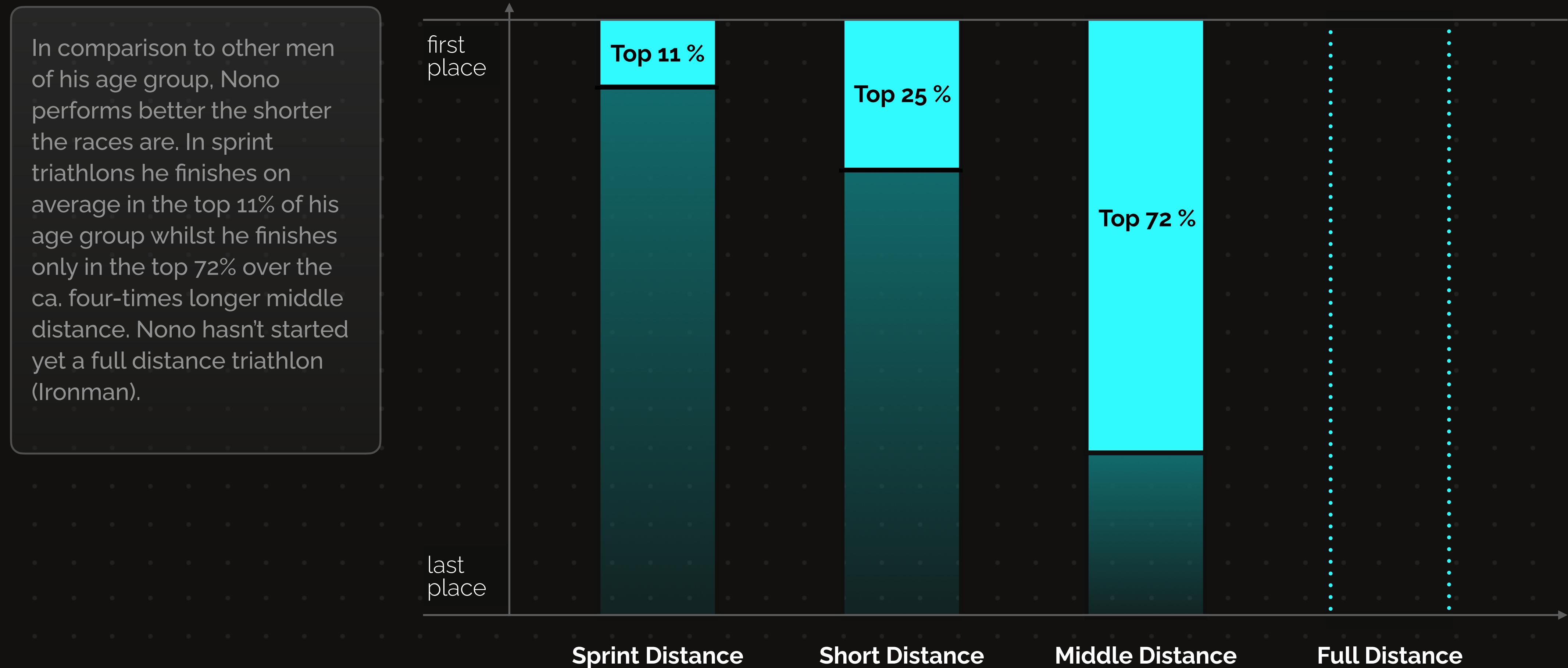
# More than 6 Years of Training



n = 1038 activities

# Placement per Event

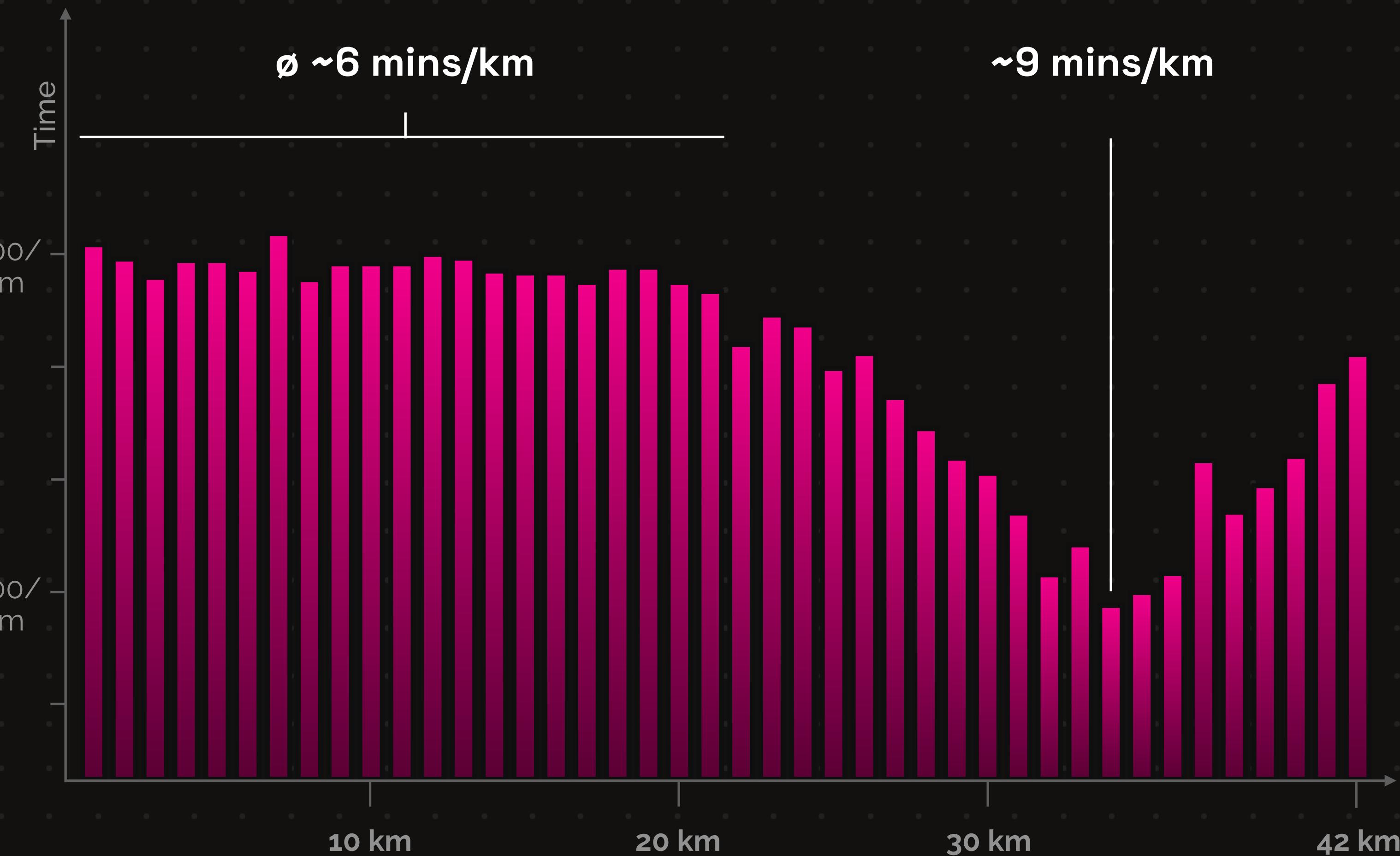
in comparison to the same age group



In comparison to other men of his age group, Nono performs better the shorter the races are. In sprint triathlons he finishes on average in the top 11% of his age group whilst he finishes only in the top 72% over the ca. four-times longer middle distance. Nono hasn't started yet a full distance triathlon (Ironman).

# First Marathon

In order to improve his long distance proficiency, Nono trained for his first marathon which he finished in October 2024. Whilst the first half of the race went okay, in the second half Nono's pace dropped drastically for some unknown reason. In the end he finished around 45 minutes later than expected. A humbling experience. Even worse: he also picked up an injury in his thigh.







What  
went wrong ?

What  
has to change ?

# A team of sports and data enthusiasts to answer these questions



**Aisling Horan**

GAA Football,  
Ballet,  
**Run**



**Katharina Kreß**

**Yoga,**  
**Bike,**  
Ballroom Dance



**Nono Weinzierl**

**Swim,**  
**Bike,**  
**Run**

Data Analytics Bootcamp @ Neue Fische

# How to improve endurance sports performance

Capstone Project Presentation

by Aisling Horan, Katharina Kreß & Nono Weinzierl

3rd March 2025

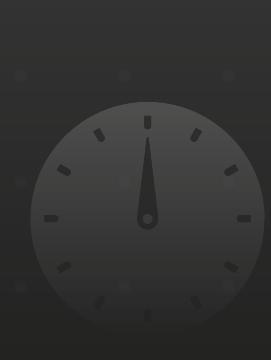
# Types of Endurance Performance

Different definitions of 'endurance performance' exist. In our project and this presentation, however, we will focus solely on cardio-vascular fitness as a means of assessing the performance of athletes.



## Fitness

cardio-vascular readiness to perform



## Speed

time over distance



## Placement

in races

# What is triathlon?

# Swim



# Bike



# Run



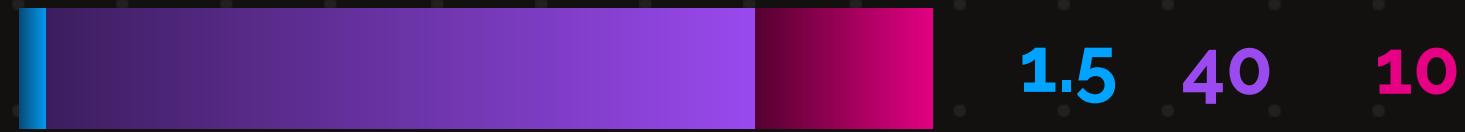
# Four classic triathlon distances

in kilometres

## Sprint Distance



## Short Distance



## Middle Distance

Half-Ironman



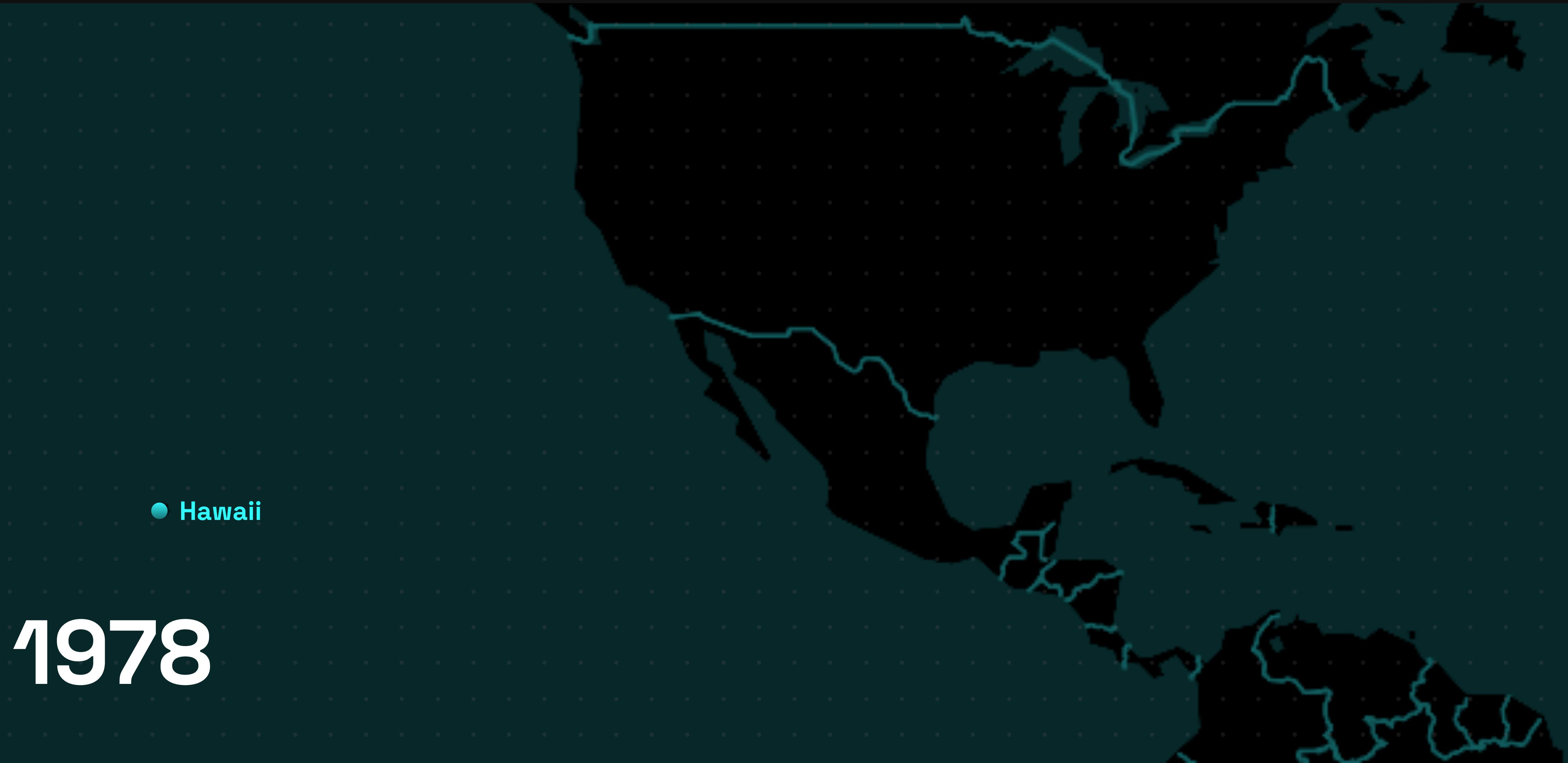
## Full Distance

Ironman



# Growth of Events

Ironman races per decade



# Growth of Events

Ironman races per decade

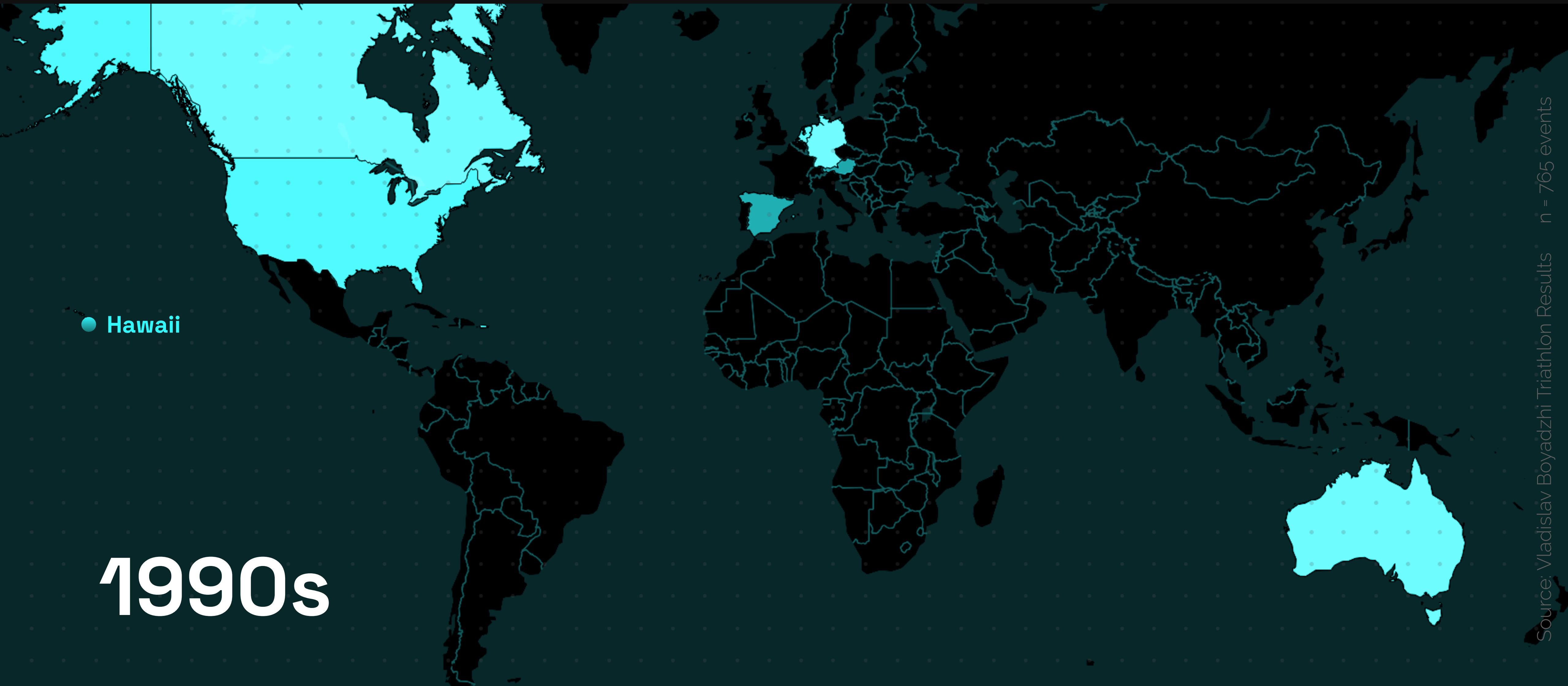
Hawaii

1980s

Source: Vladislav Boyadzhi Triathlon Results n = 765 events

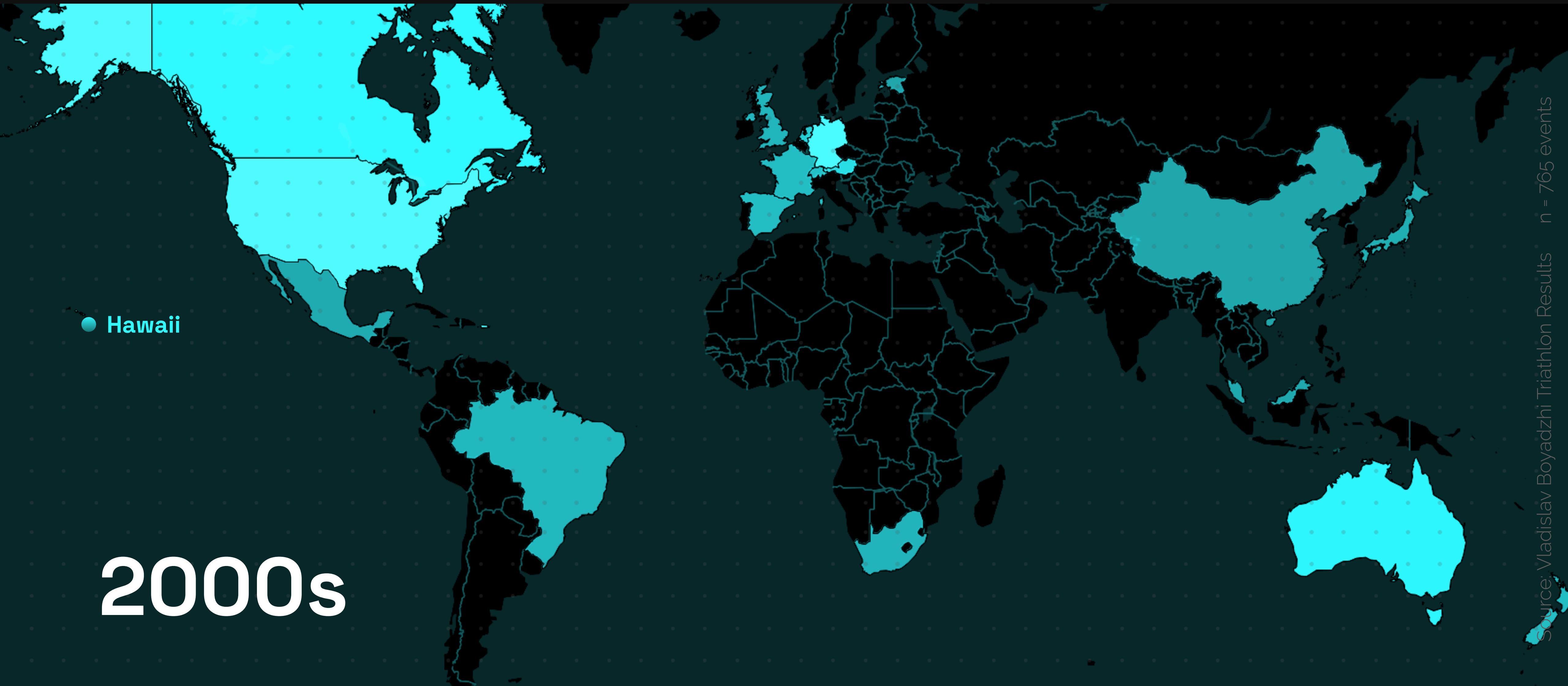
# Growth of Events

Ironman races per decade



# Growth of Events

Ironman races per decade



# Growth of Events

Ironman races per decade

Hawaii

2010s

Source: Vladislav Boyadzhi Triathlon Results n = 765 events

# Growth of Events

Ironman races per decade

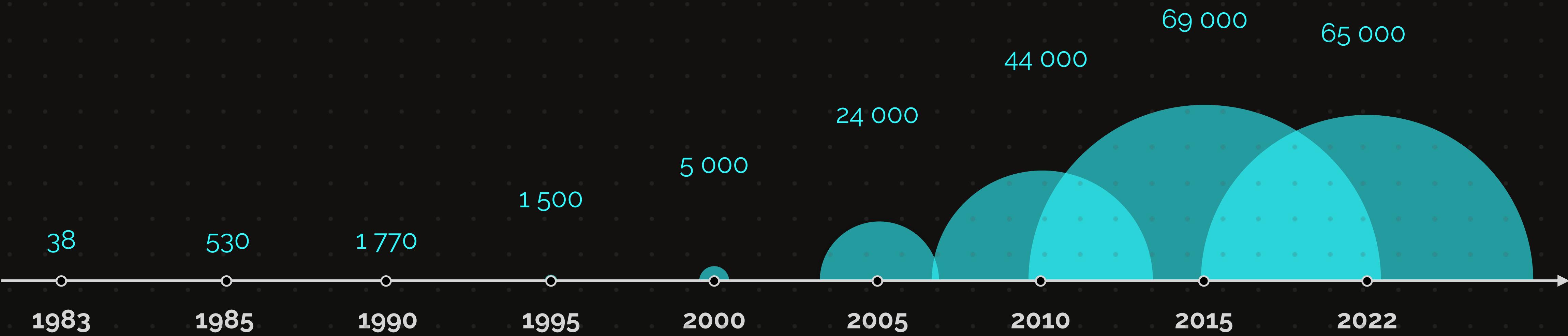
Hawaii

2020s

Source: Vladislav Boyadzhi Triathlon Results n = 765 events

# Participants

Ironman races

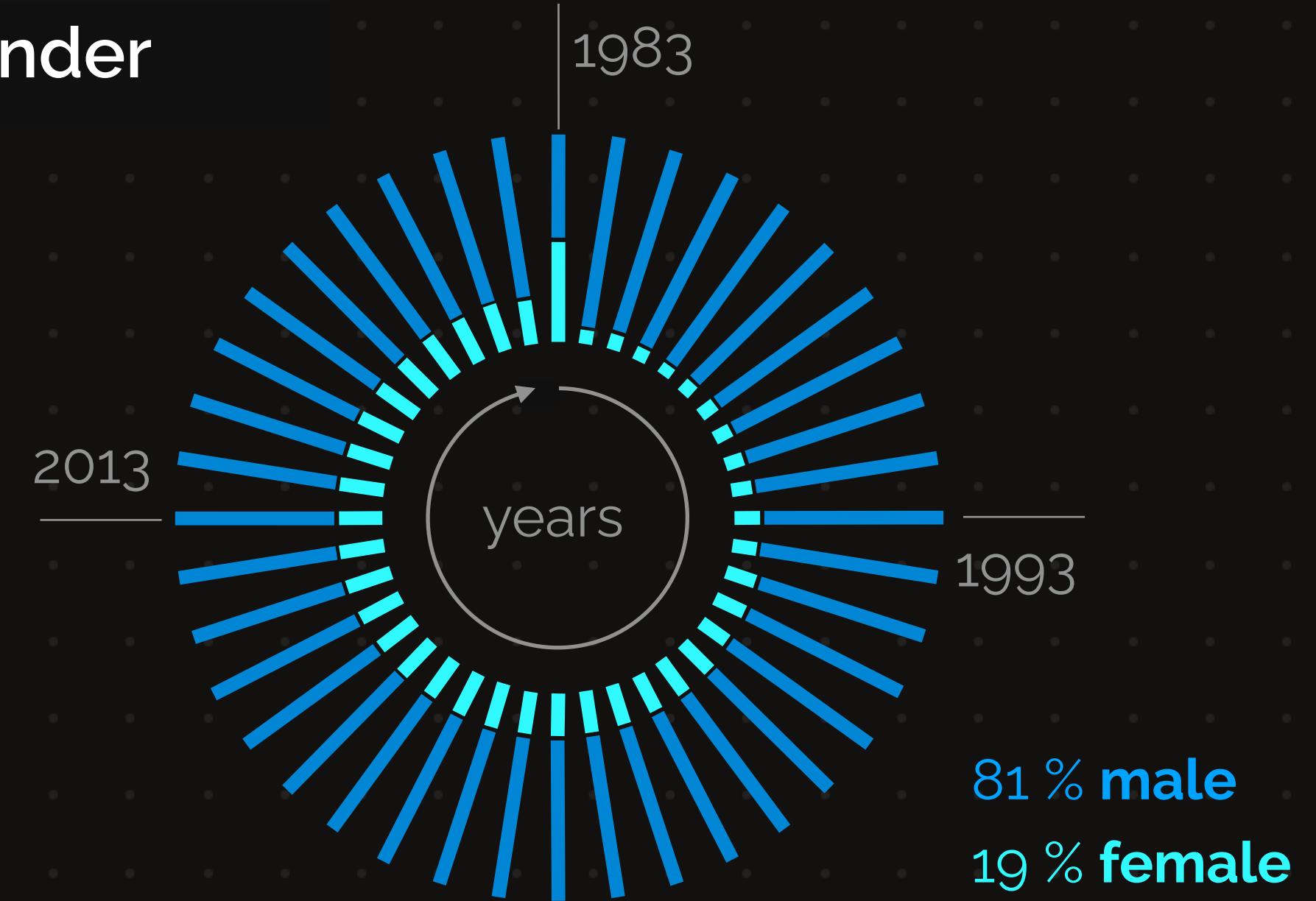


numbers rounded

# Gender & Age

Ironman races

Gender

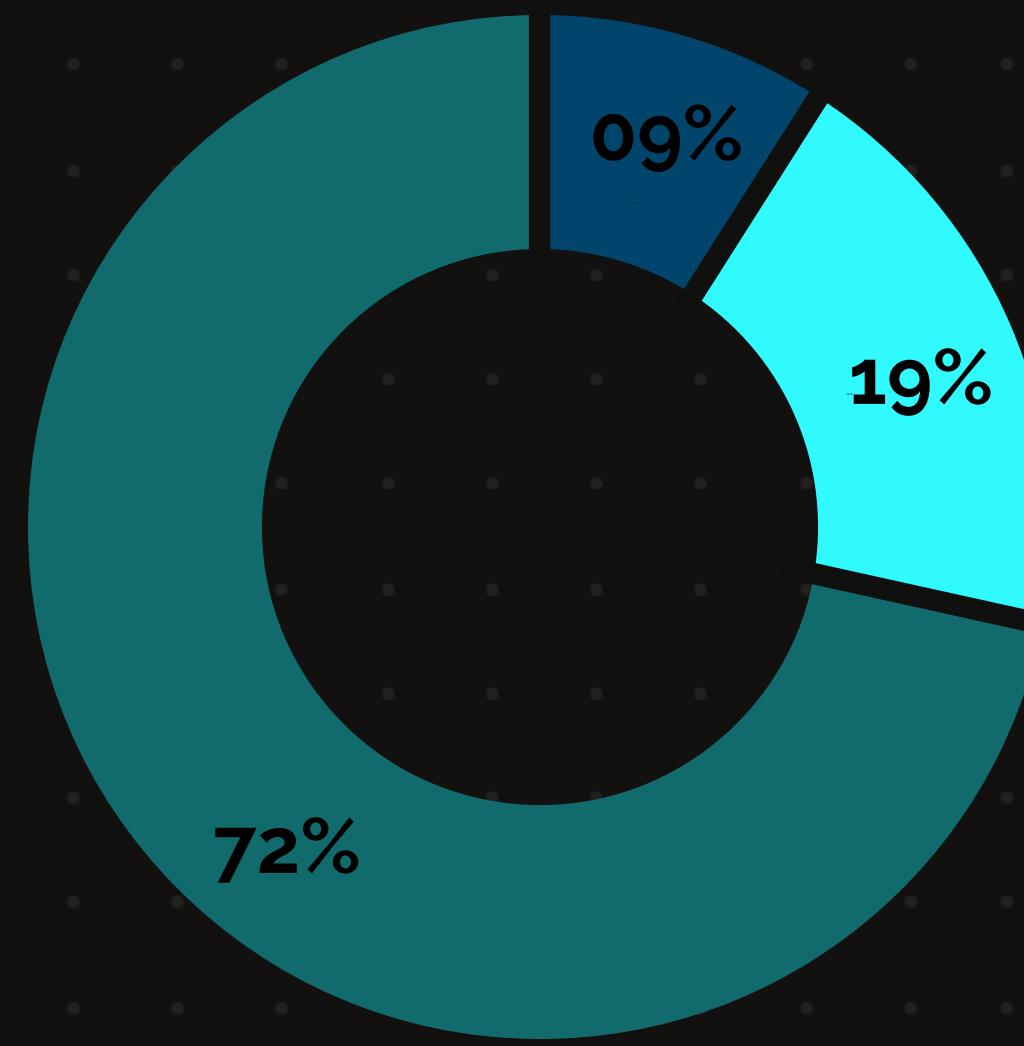


only 19 % Ironman starters are female

1983-2022

Age

18-30 years  
30-49 years  
50+ years



19% are older than 50 years

1983-2022

# Growth in Professionalism

1980s



1990s



2020s



Growing Need for Data

Let's dive  
into the **data**

We'll be looking at three different aspects:

intensity

consistency

injury

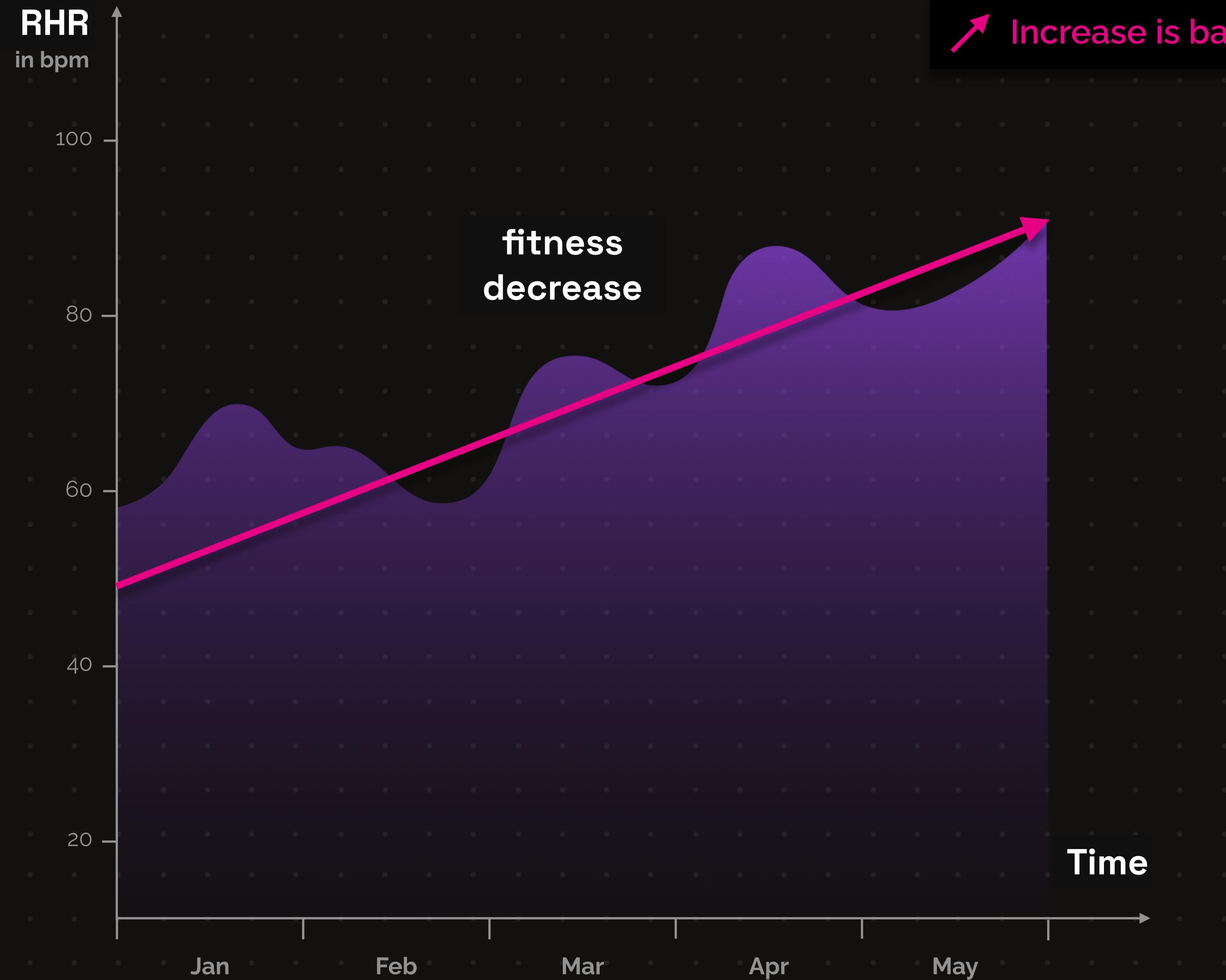
Does **high** or **low** intensity  
training improve fitness?

Intensity

# Resting Heart Rate RHR

A common biomarker used to measure fitness is resting heart rate (RHR) which is measured in beats per minute (bpm). The lower the heart rate the fitter a person is.

indicator for cardio-vascular fitness

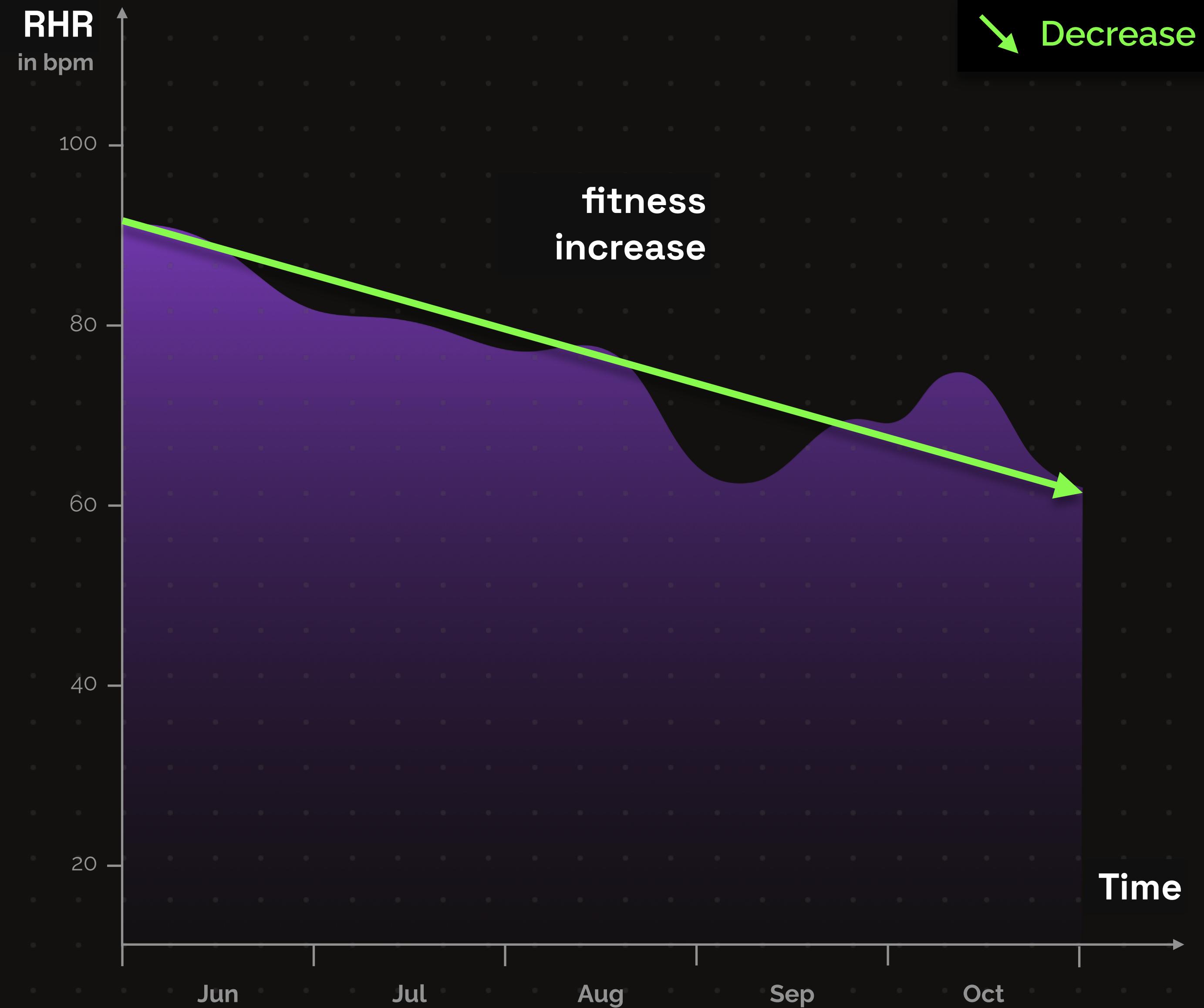


Intensity

## Resting Heart Rate RHR

A common biomarker used to measure fitness is resting heart rate (RHR) which is measured in beats per minute (bpm). The lower the heart rate the fitter a person is.

indicator for cardio-vascular fitness



## Trial Data

# Intensity Trial Data

In order to better understand the effect of high and low intensity training we are having a closer look at a public dataset gathered for a Harvard trial.

### SOURCE

Cardiovascular and metabolic profiles of moderately trained individuals following a 6-week HIIT vs. MICT intervention

participants



65% female  
35% male

time period



2 groups



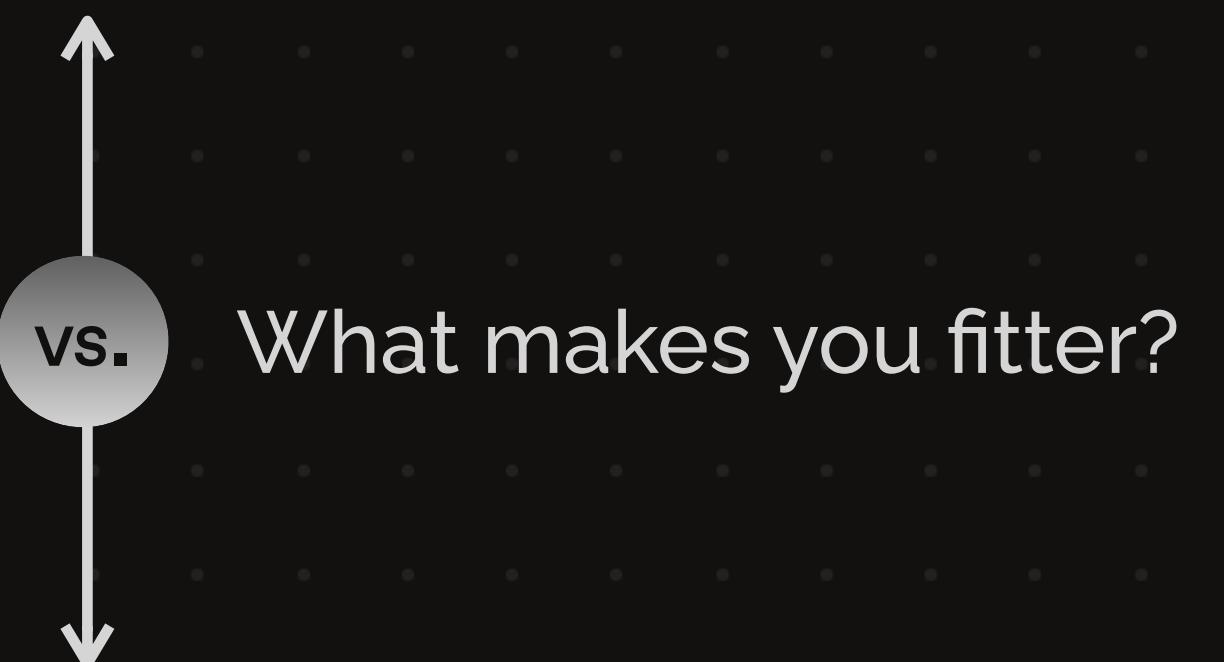
Trial Data

# Intensity Trial Data

In order to better understand the effect of high and low intensity training we are having a closer look at a public dataset gathered for a Harvard trial.



**high intensity** 6 weeks



What makes you fitter?



**low intensity** 6 weeks

## SOURCE

Cardiovascular and metabolic profiles of moderately trained individuals following a 6-week HIIT vs. MICT intervention



# high intensity

Trial Data

Trial Data

# Impact of high intensity training on RHR

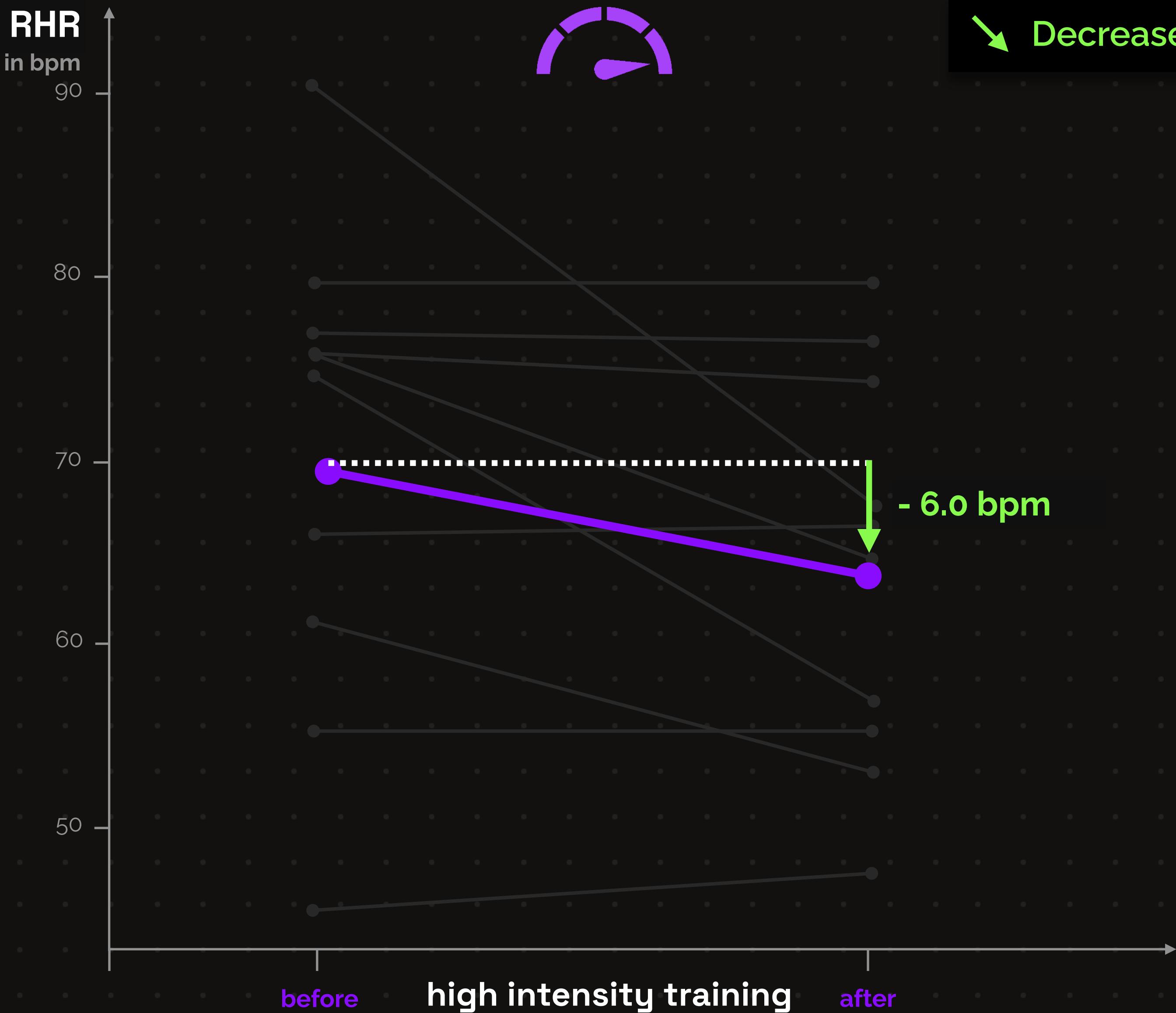
In moderately active individuals, high intensity training has a **positive effect** on RHR, which decreased on average by 6 bpm.

 medium effect size

Cohen's D: 0.6



Decrease is good



n = 10



# low intensity

Trial Data

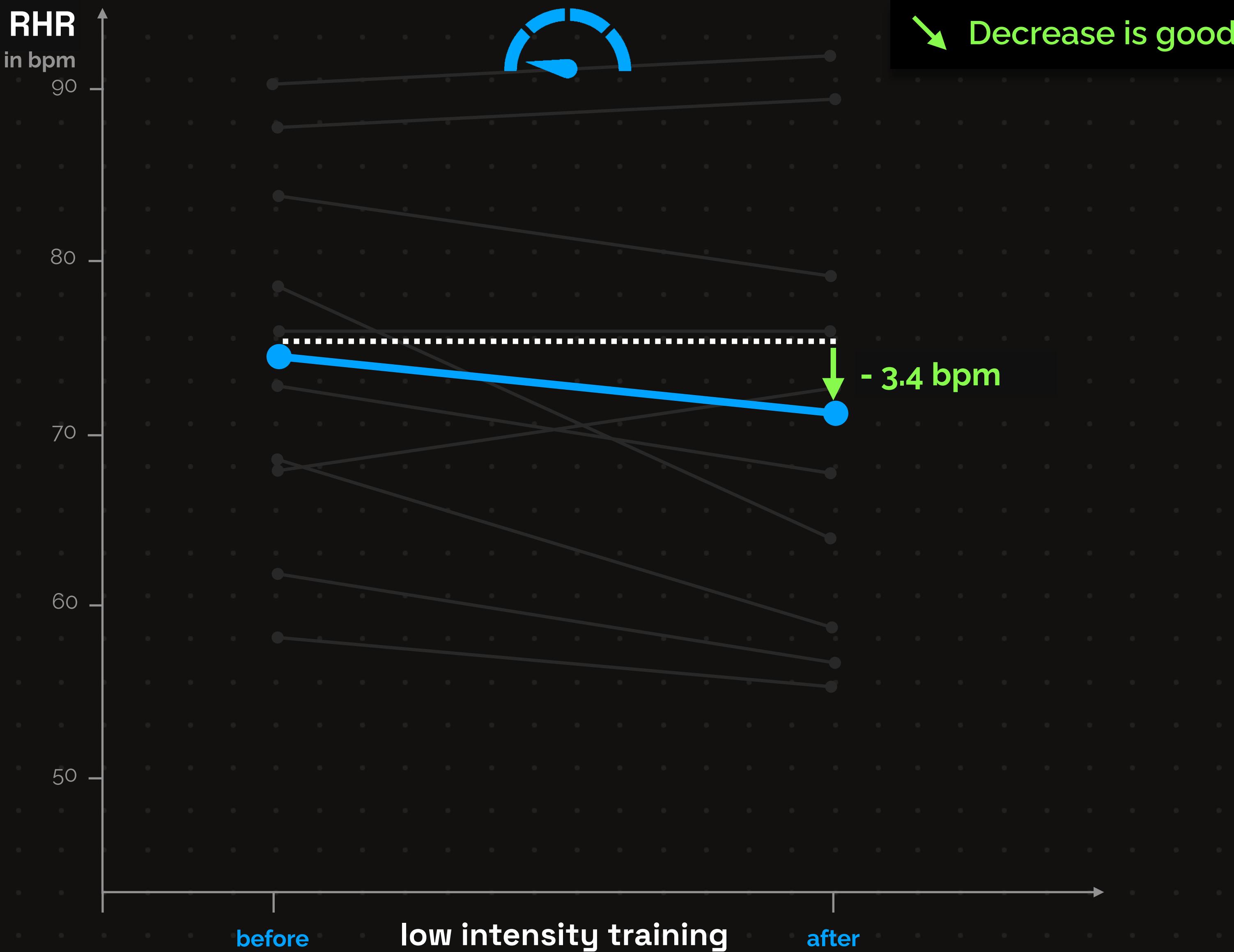
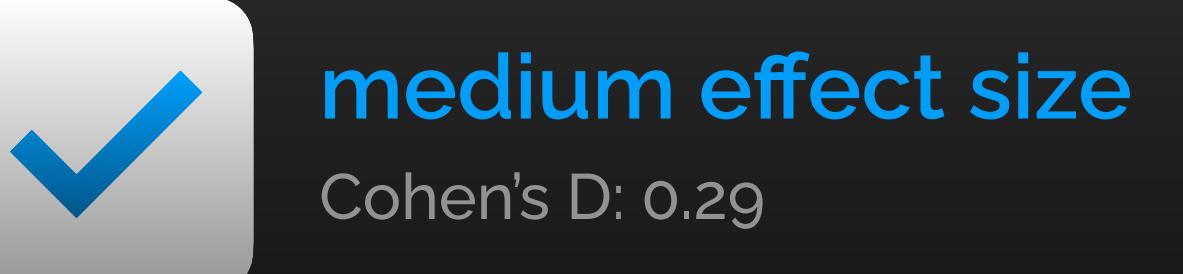
Trial Data

# Impact of low intensity training on RHR

In moderately active individuals, low intensity training has a **small positive effect** on RHR, which decreased on average by **3.4 bpm**.

 medium effect size

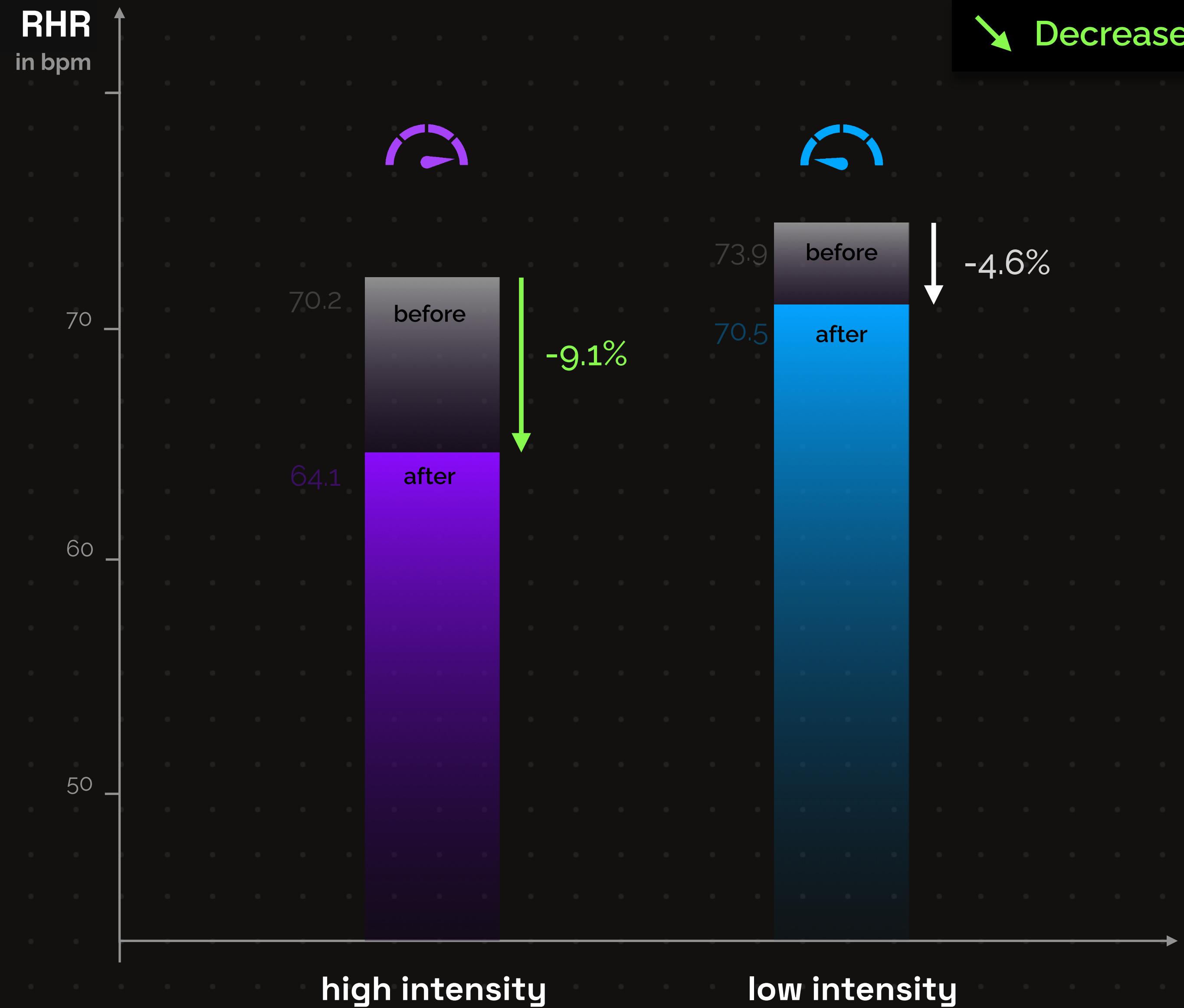
Cohen's D: 0.29



Trial Data

Larger improvements after high intensity training

High intensity training has a larger impact on fitness on moderately active individuals



n = 20



# high intensity

Nono's Data

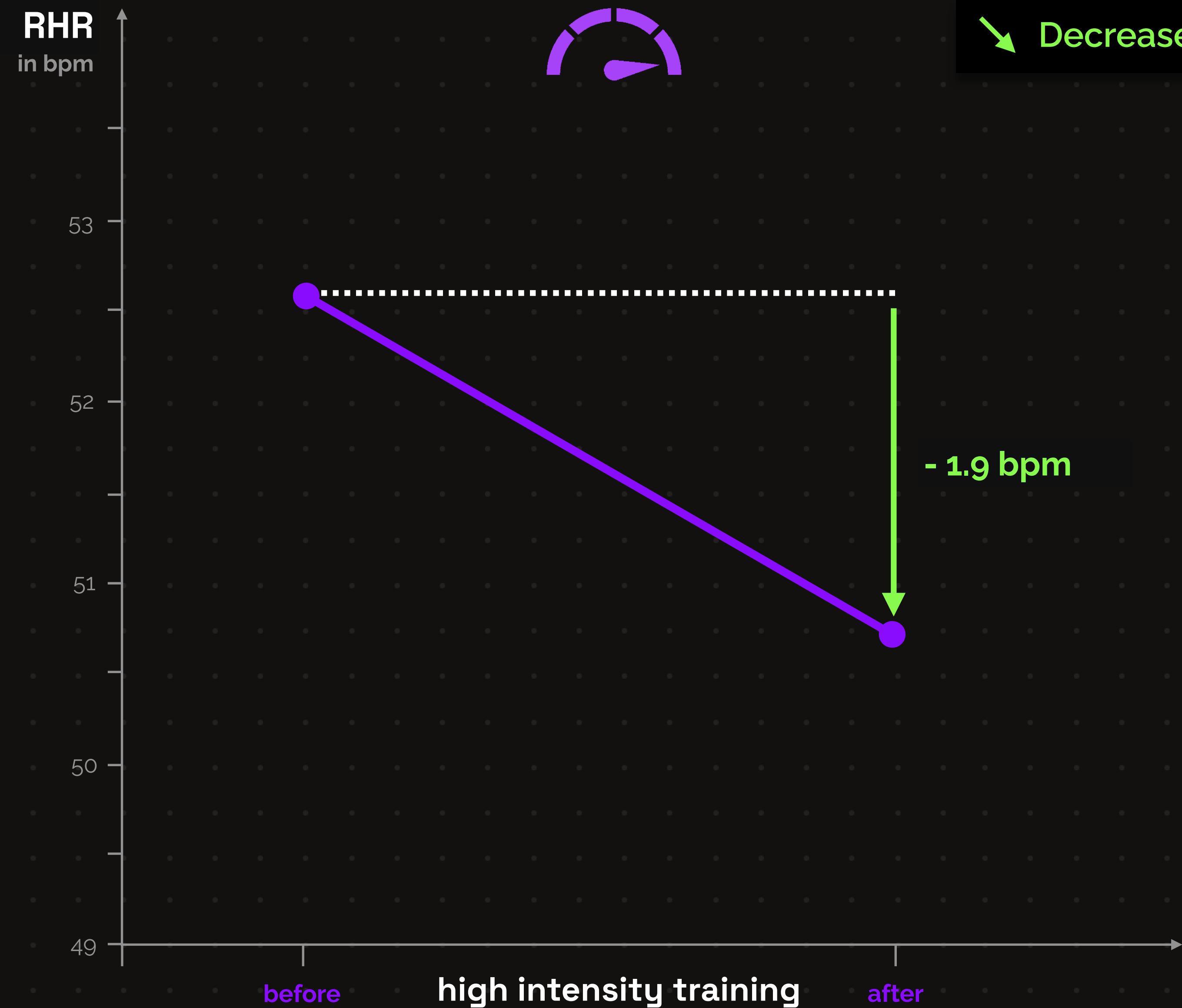
Nono's Data

## Impact of high intensity training on RHR

For Nono, after the four weeks of high intensity training, we can measure a **positive effect** on RHR, which decreased by 1.9 bpm.



**significant**  
p-value: 0.004



n = 58 activities



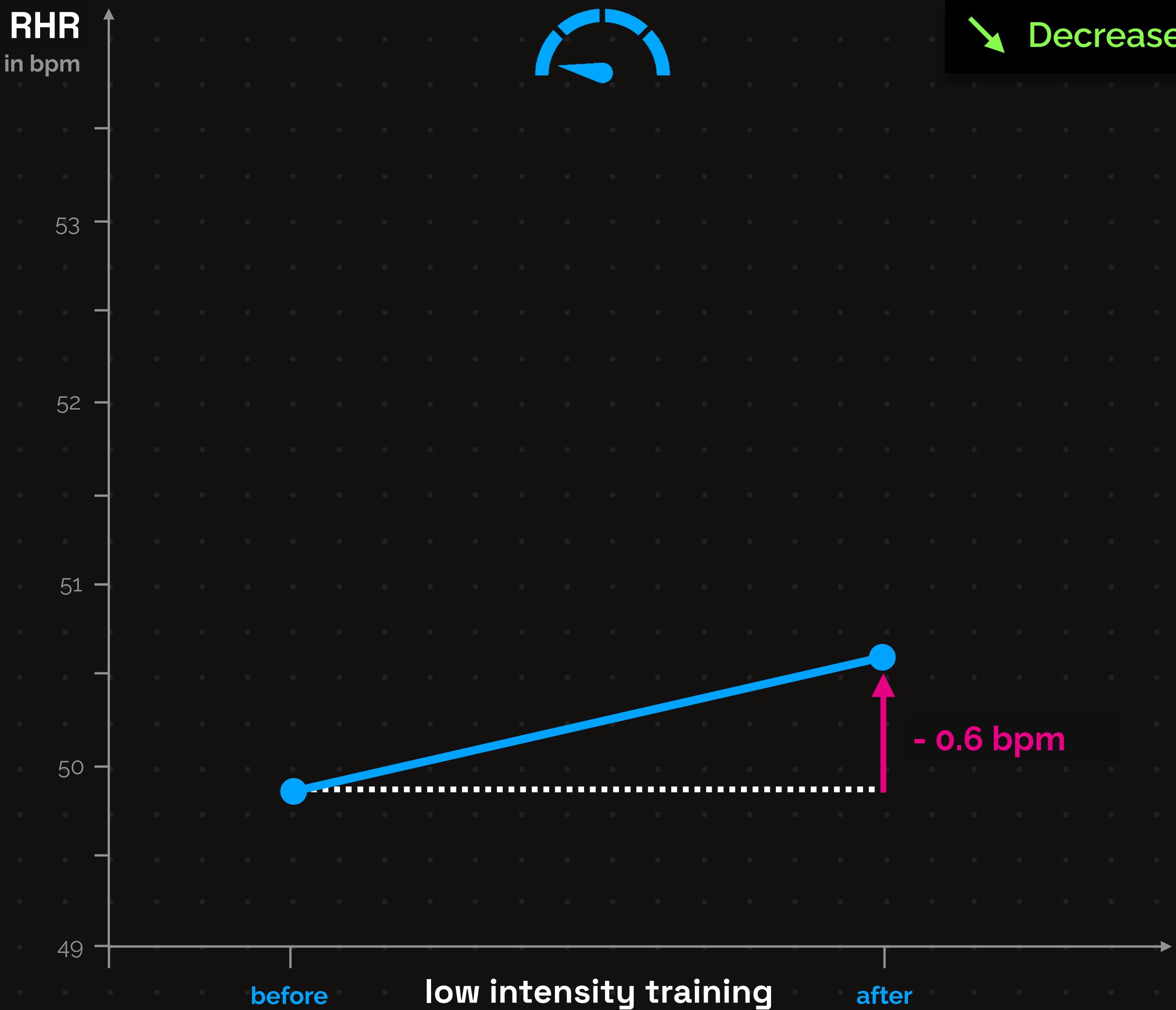
# low intensity

Nono's Data

Nono's Data

# Impact of low intensity training on RHR

For Nono, after the four weeks of low intensity training, we can measure a **negative effect** on RHR, which increased by 0.6 bpm.



n = 62 activities

Nono's Data

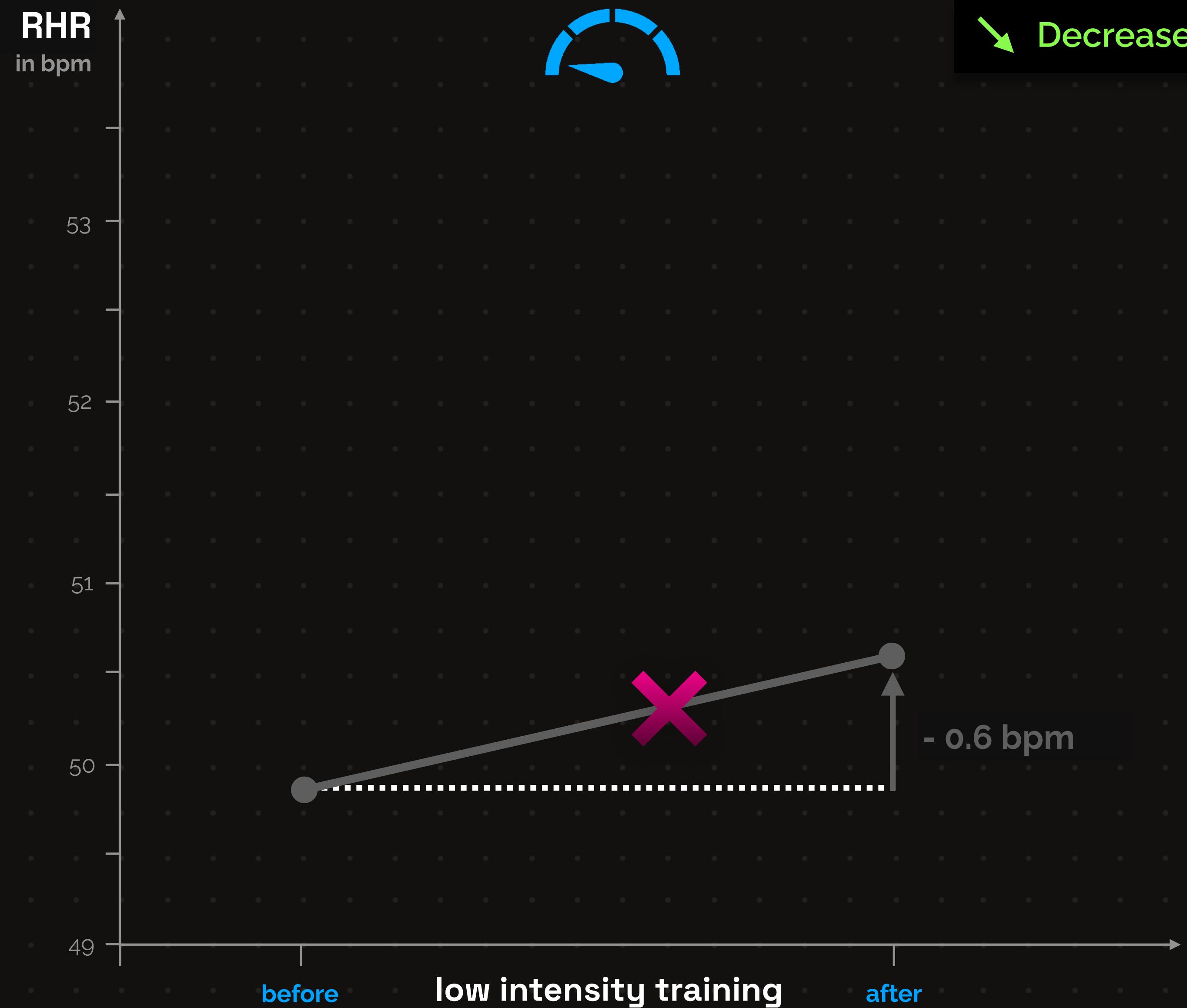
## Impact of low intensity training on RHR

For Nono, after the four weeks of low intensity training, we can measure a **negative effect** on RHR, which increased by 0.6 bpm.



**no significance**  
p-value: 0.466

The result is likely due to chance.  
So we **cannot say** that low intensity is either bad or good.



n = 62 activities

Decrease is good

LEARNING

**High intensity  
increases fitness faster**

especially in moderately active individuals

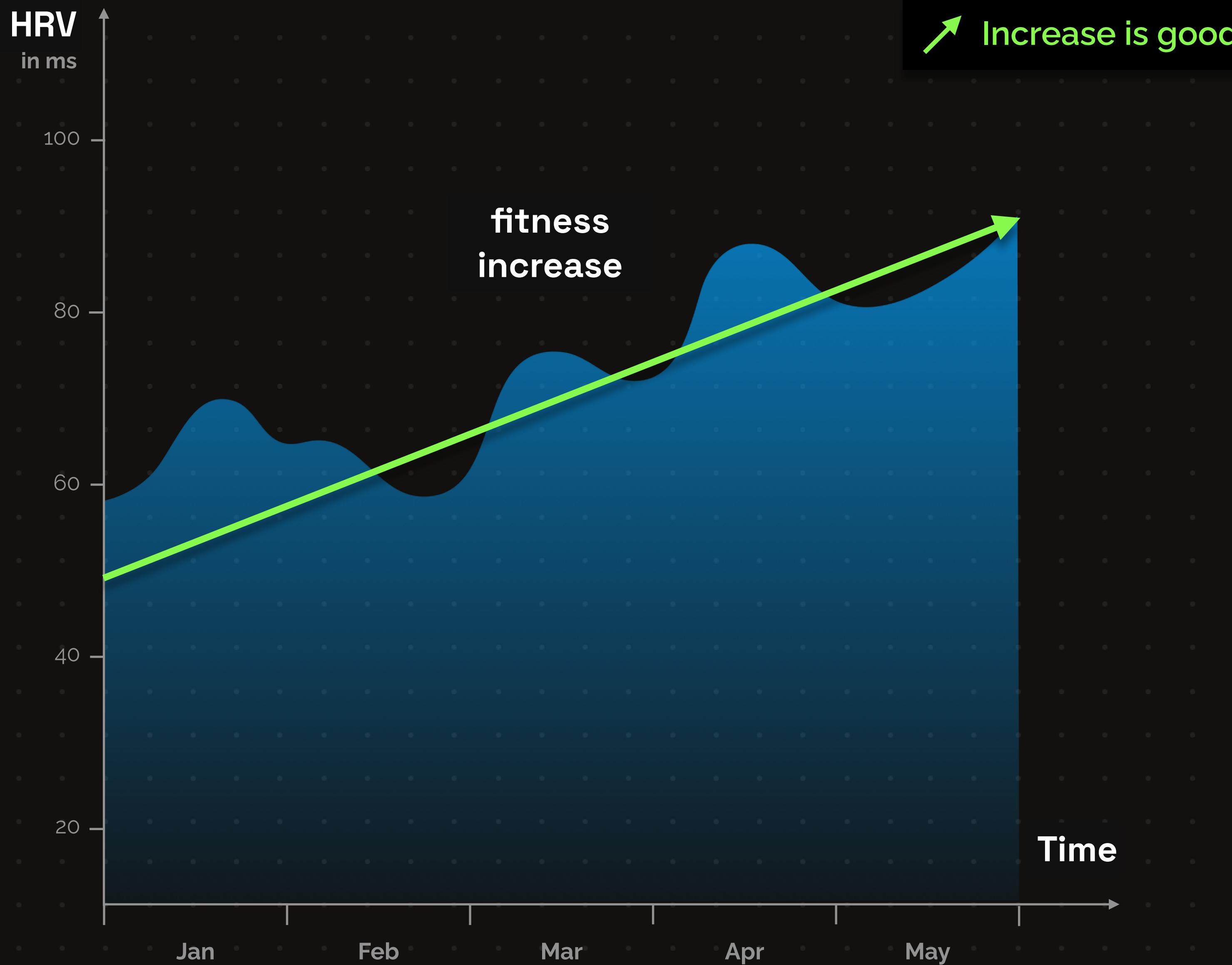
Does **consistent** training  
improve fitness?

# Heart Rate Variability HRV

Heart rate variability describes the variance in time between the beats of your heart. The greater this variability is, the more "ready" your body is to execute at a high level.



indicator for cardio-vascular fitness

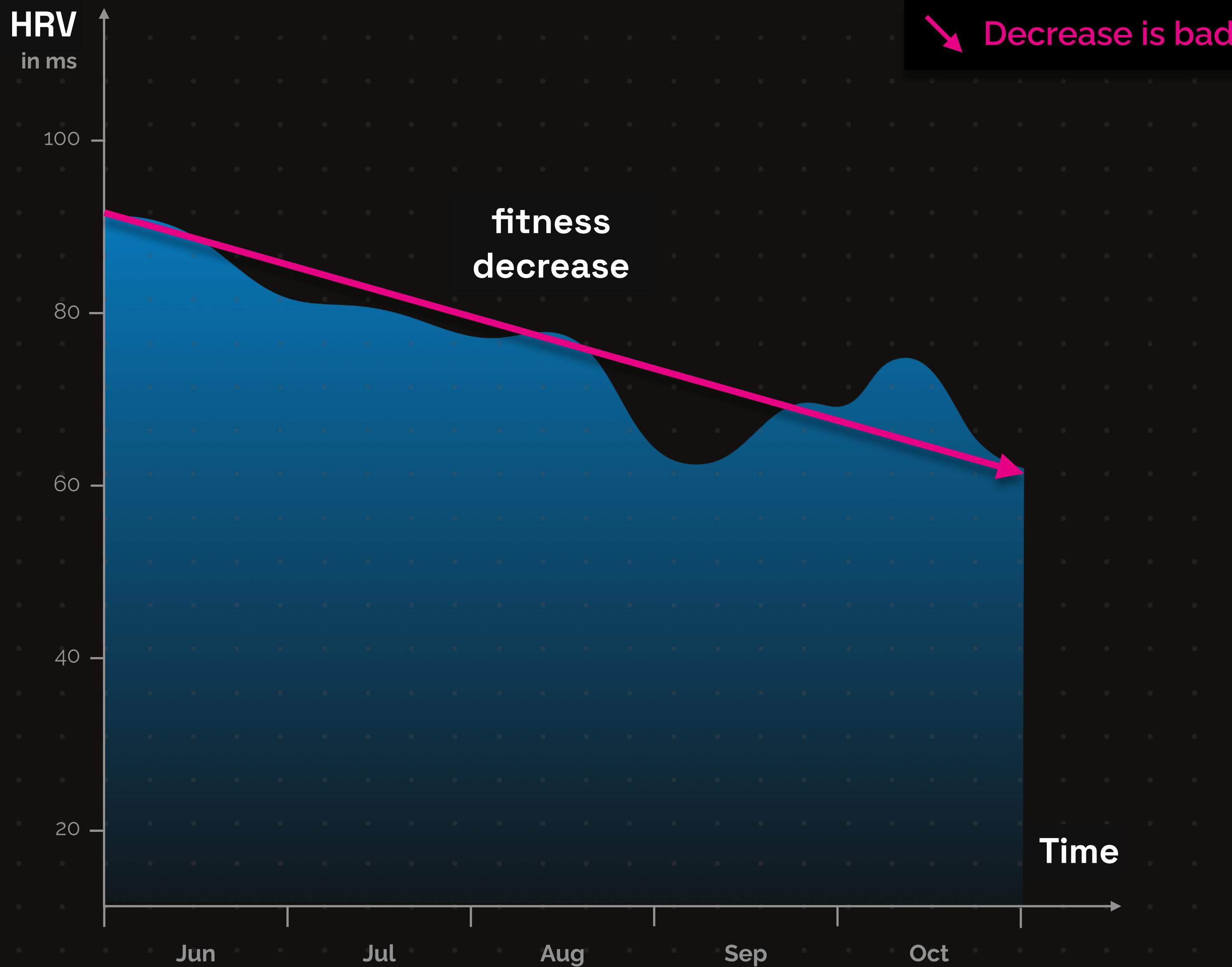


# Heart Rate Variability HRV

Heart rate variability describes the variance in time between the beats of your heart. The greater this variability is, the more "ready" your body is to execute at a high level.



indicator for cardio-vascular fitness





Nono's Data

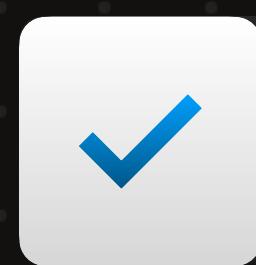
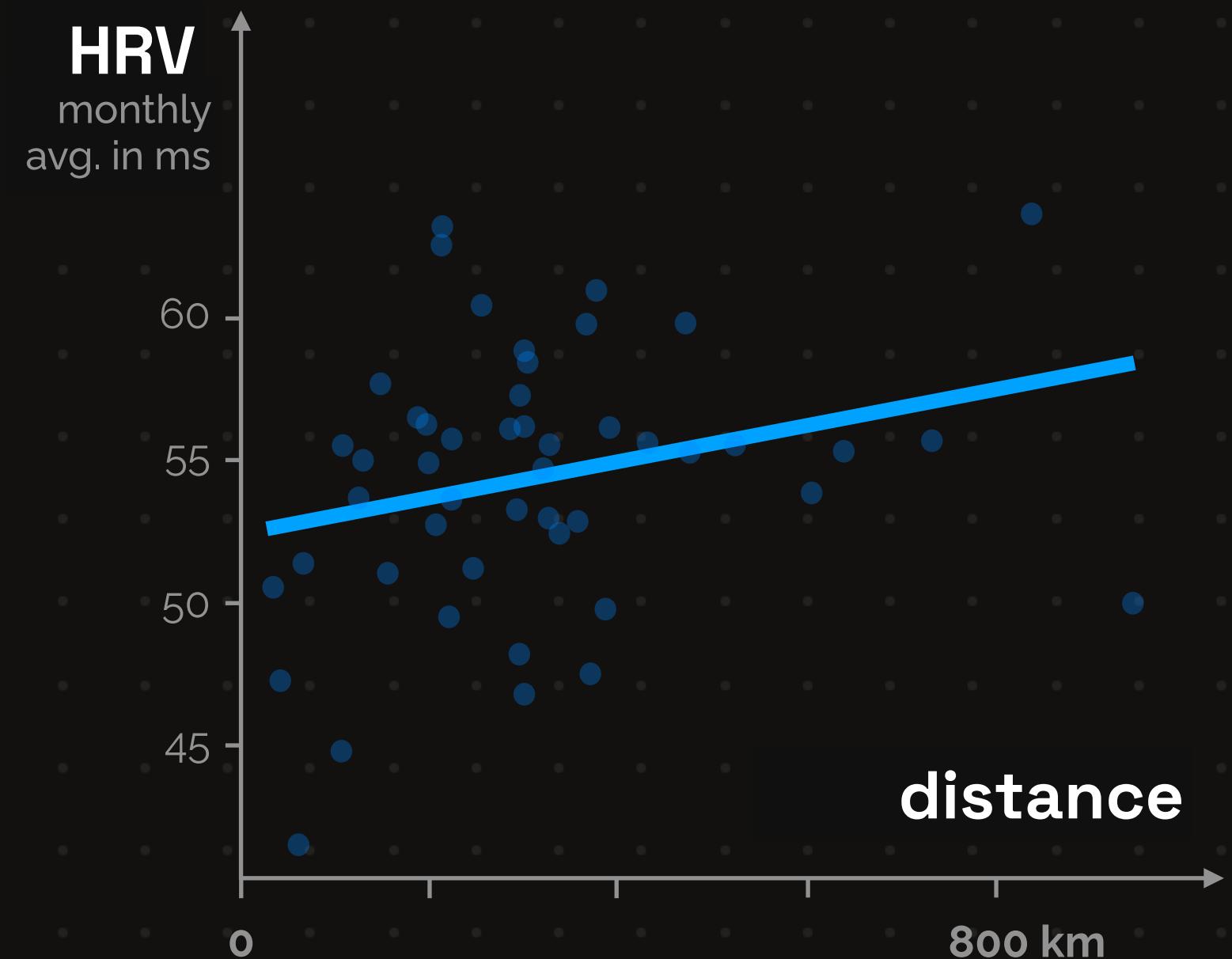
Nono's Data



↗ Increase is good

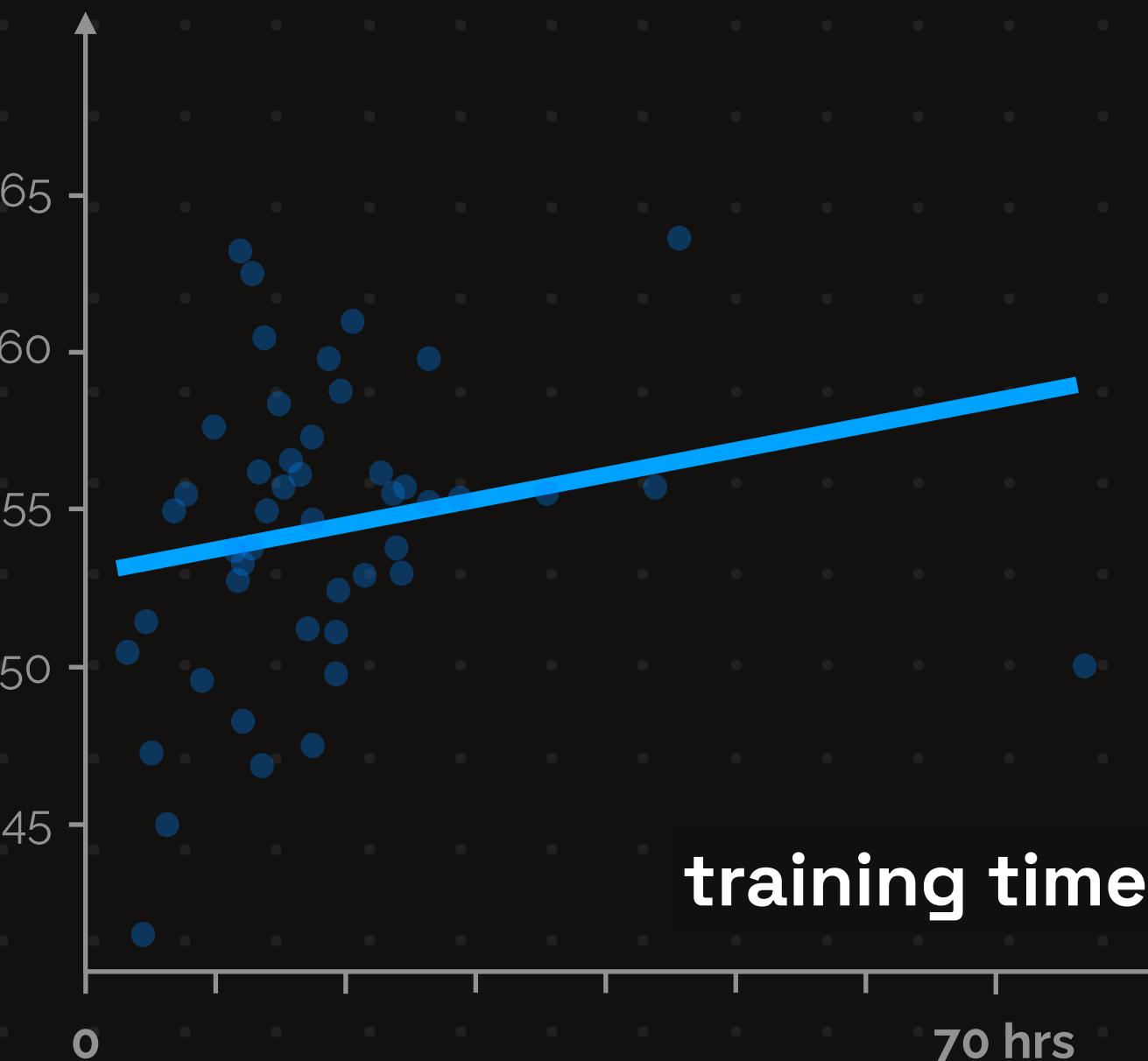
# Factors improving HRV

n = 48 months



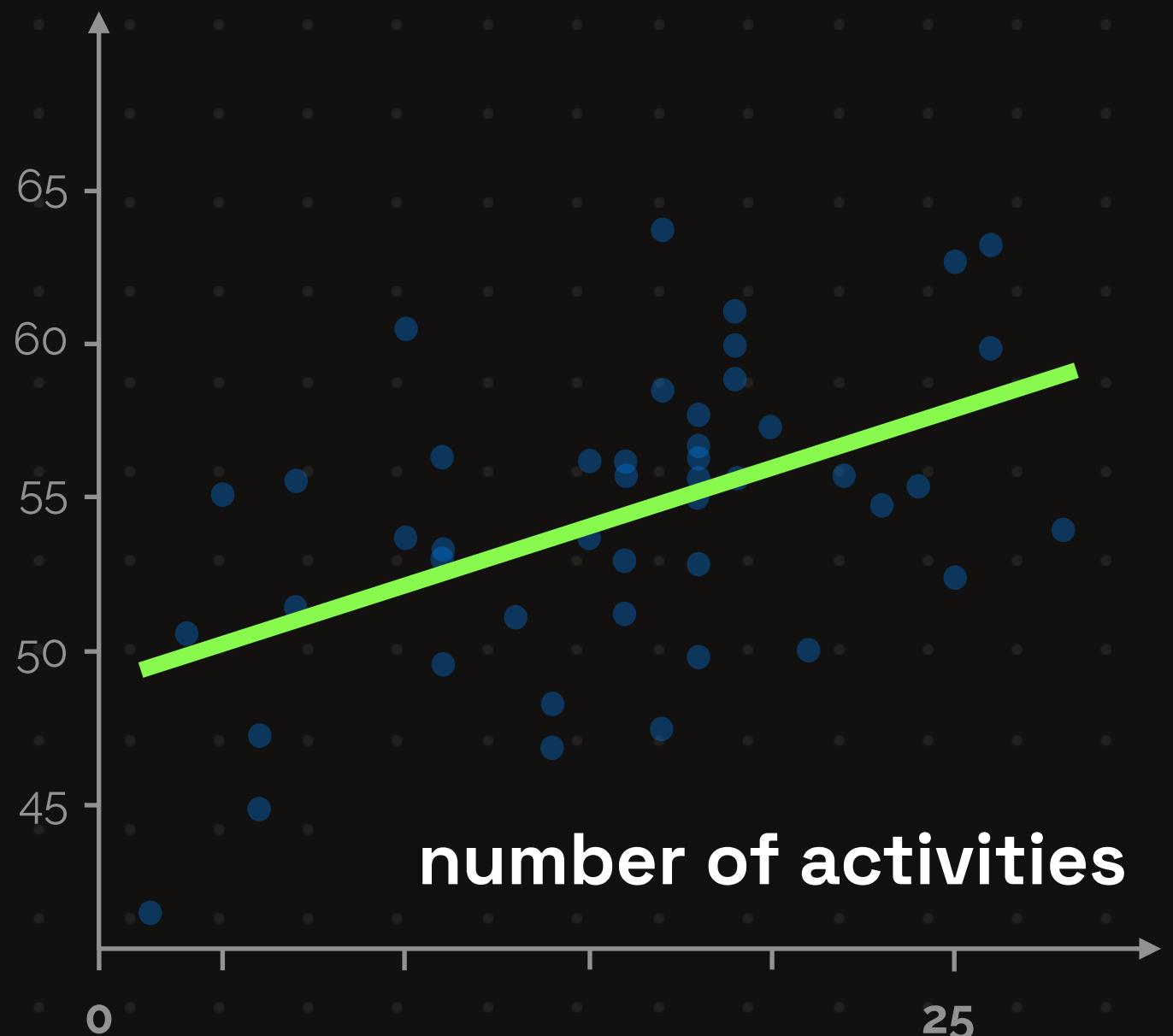
low correlation

correlation coefficient: 0.26



low correlation

correlation coefficient: 0.2



high correlation

correlation coefficient: 0.51

## LEARNING

**Train more often,  
not just harder**

At least in Nono's data we can see that the number of activities has a larger positive impact on his fitness (HRV) than the duration or distance of individual activities.

**What about change?**

# Changes in Training Routine

distance + training time + activities  
variance variance variance

# Changes in Training Routine

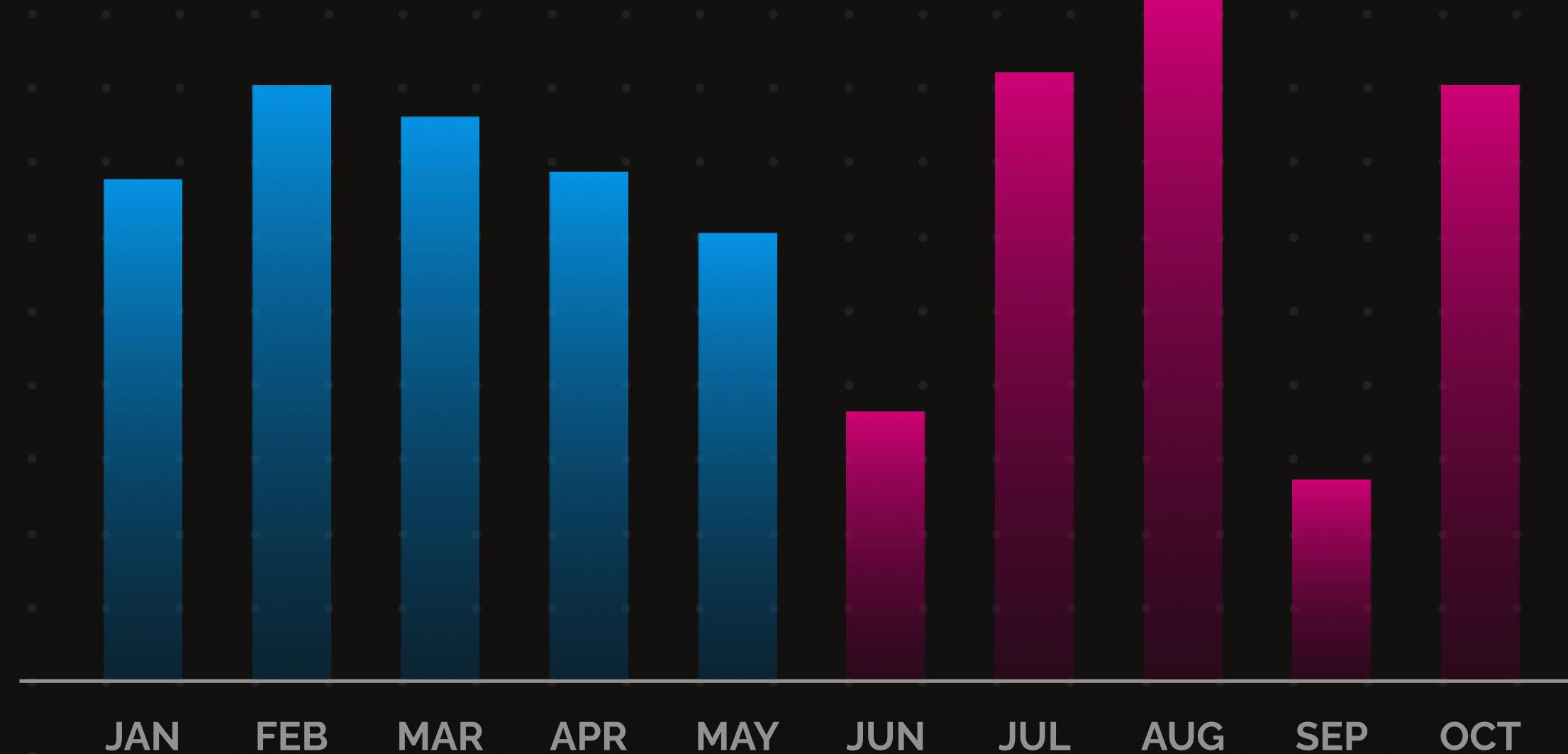
## Consistency

low variation = small change

## Inconsistency

high variation = large change

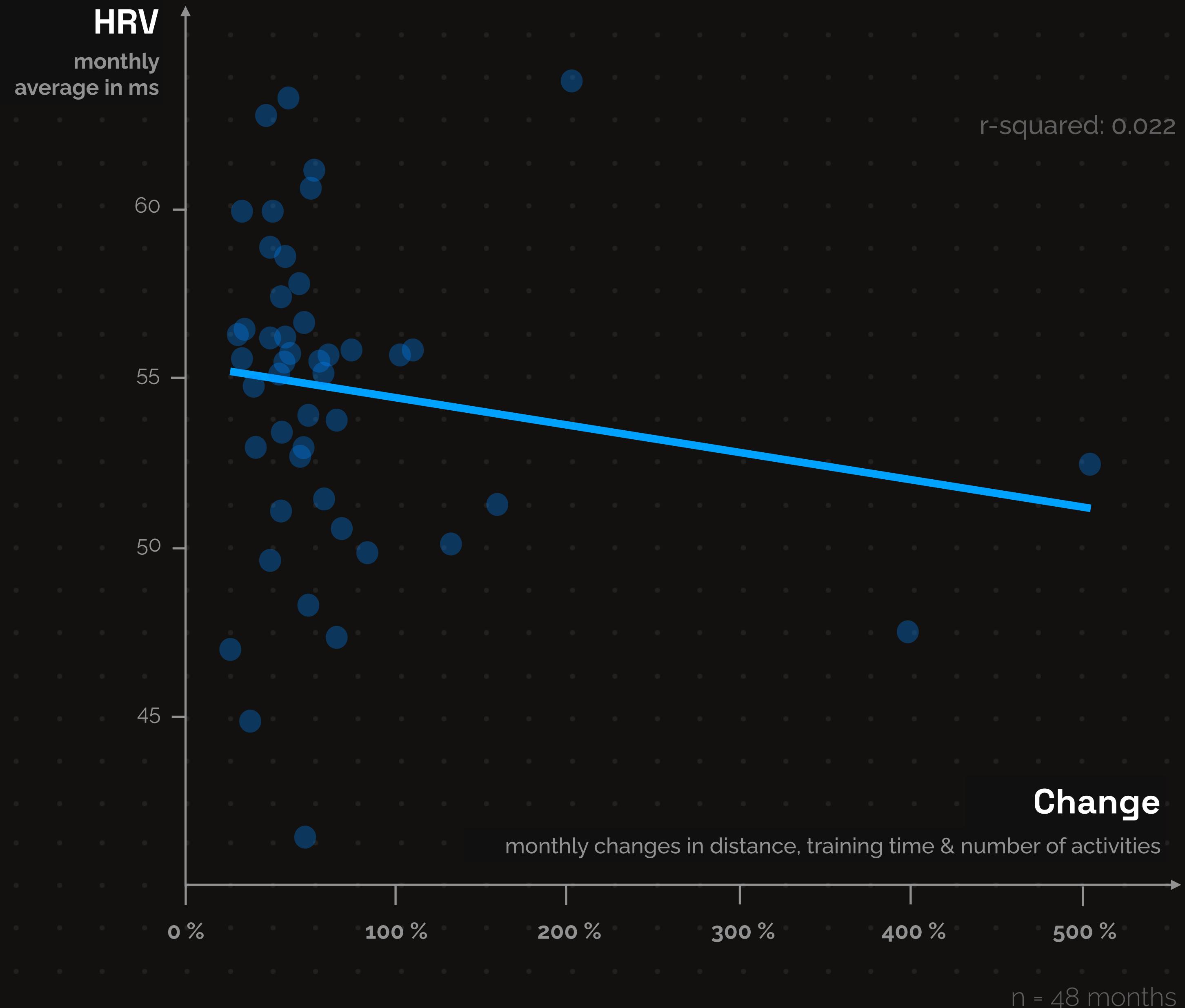
consistent period      inconsistent period



General Change



# Relationship between Change & HRV





# Relationship between Change & HRV



**no correlation**

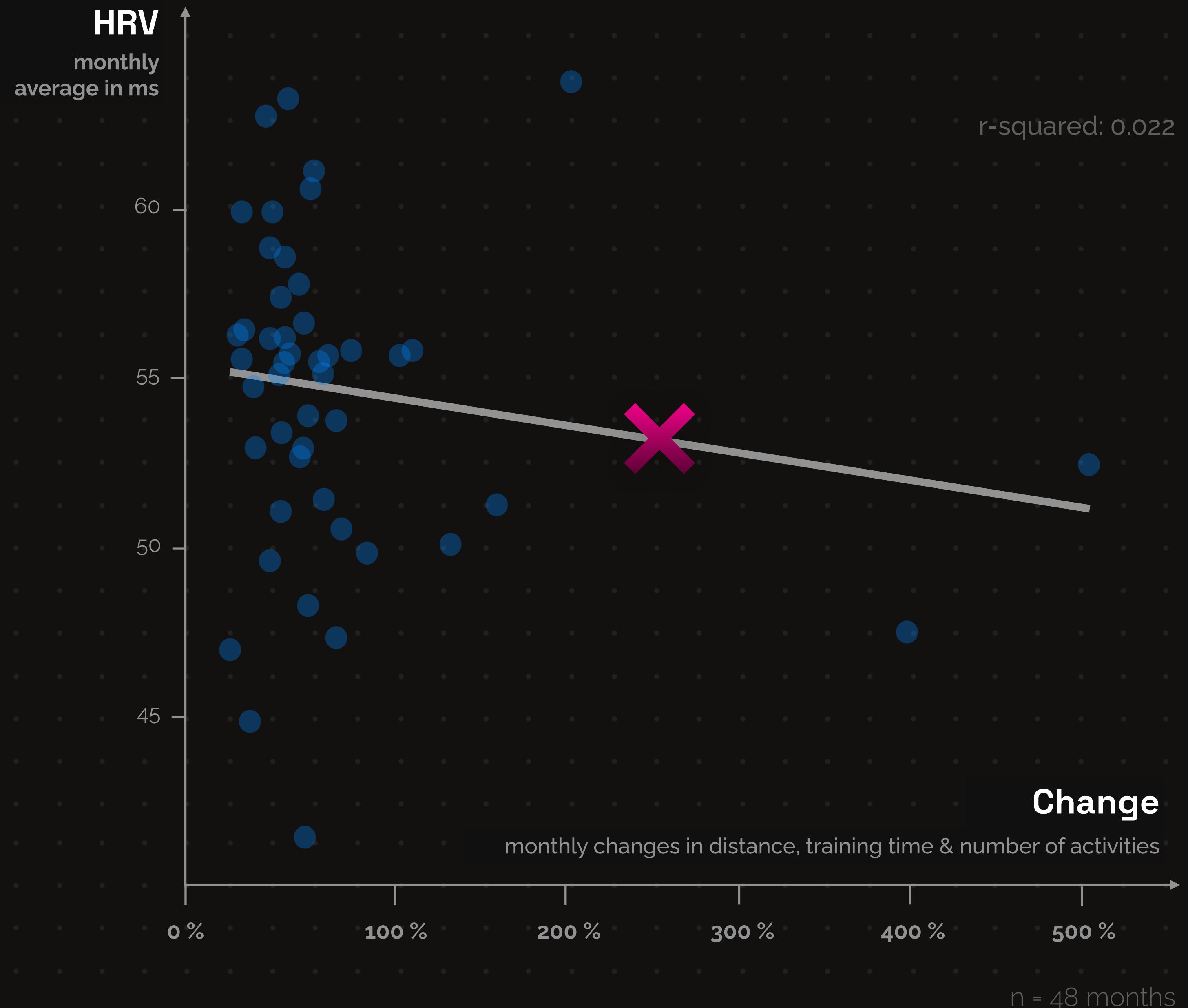
correlation coefficient: -0.15



**no significance**

p-value: 0.16

We can't completely rule out a relationship, but if one exists, it's likely very weak.



**but ...**



# Relationship between Change & HRV



**strong correlation**

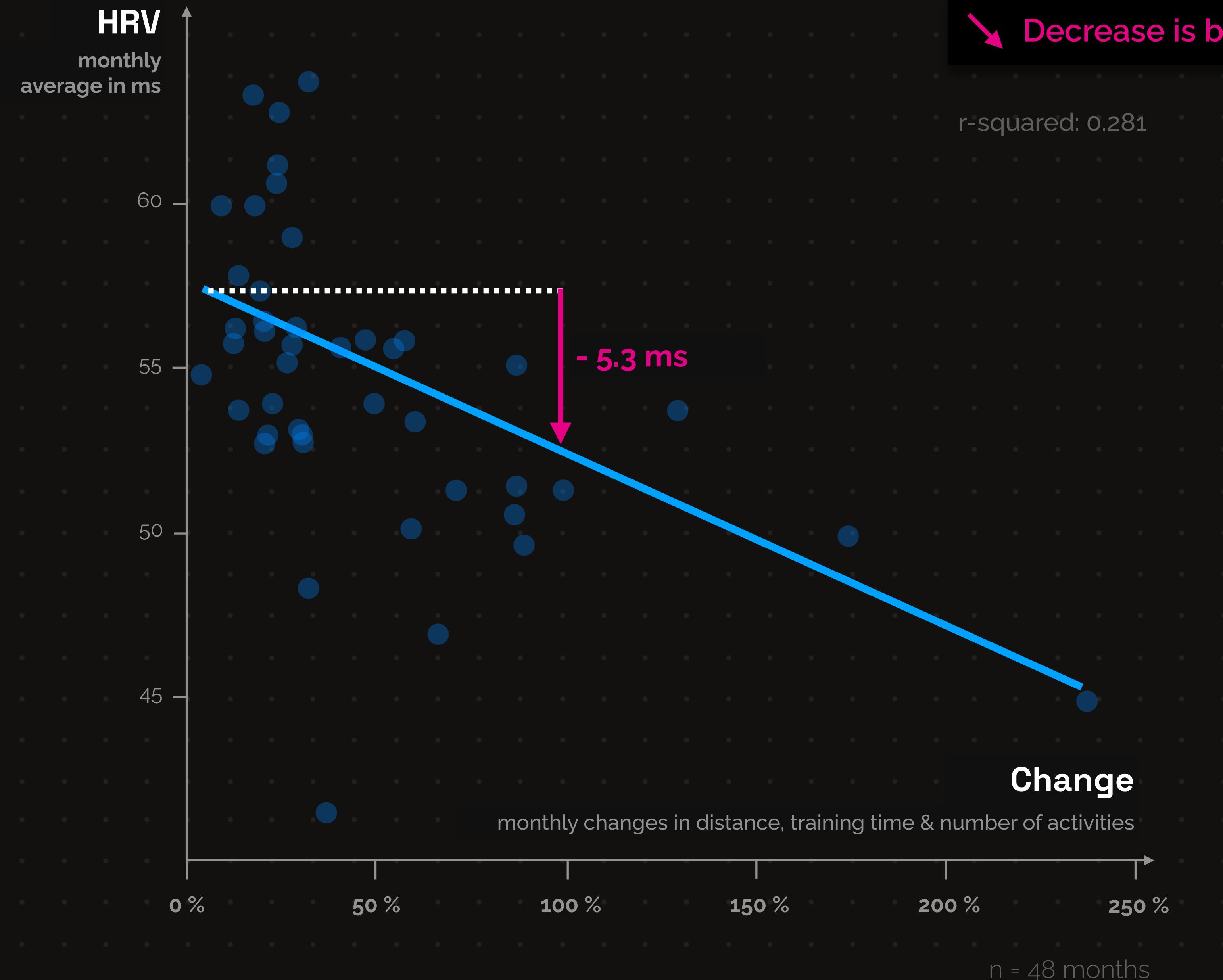
correlation coefficient: -0.53



**significant**

p-value: 0.0

In running, as **changes in training routine** increase, HRV tends to **decrease** in a meaningful way.



LEARNING

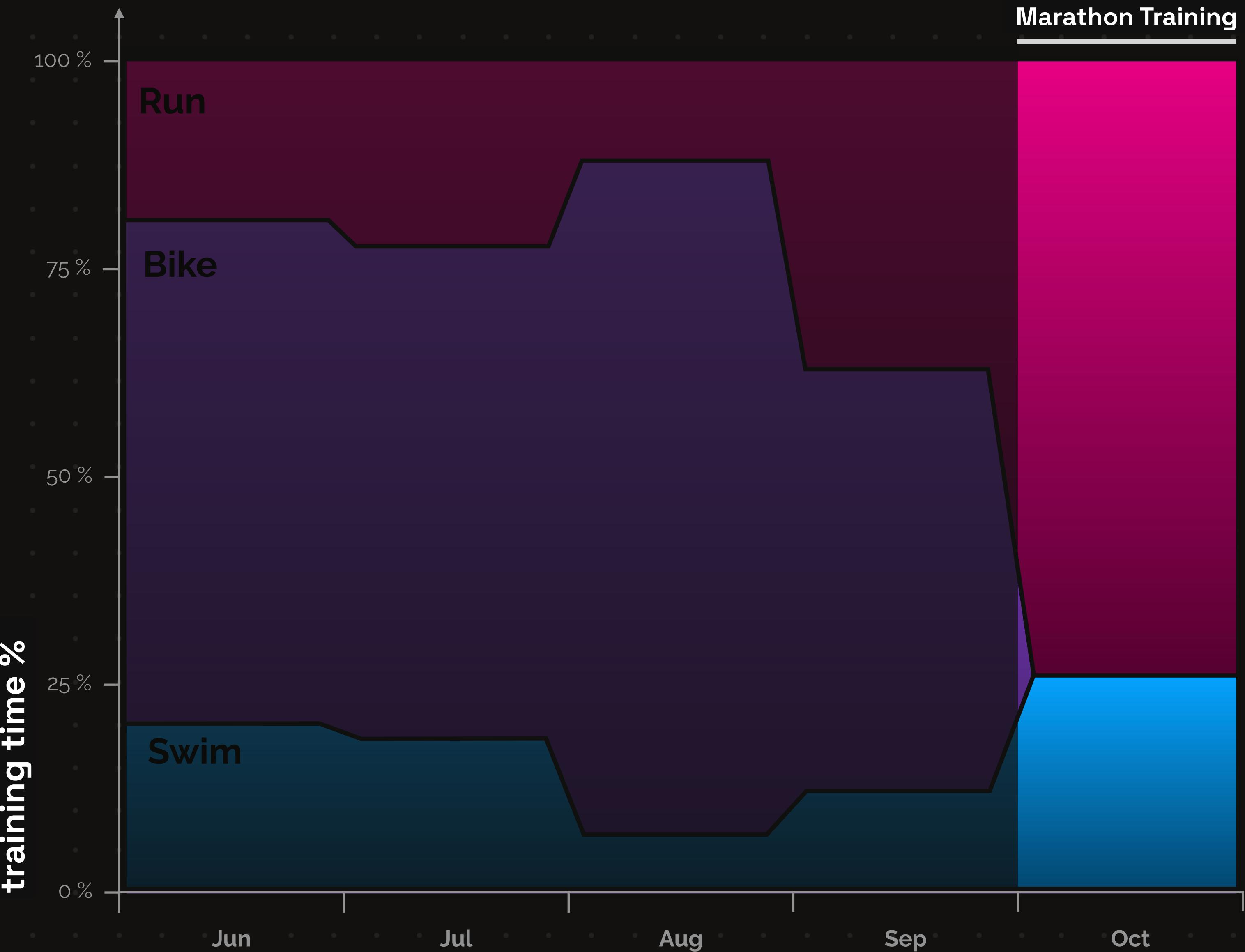
Don't change your  
training routine drastically

especially in running



# Marathon Training & HRV

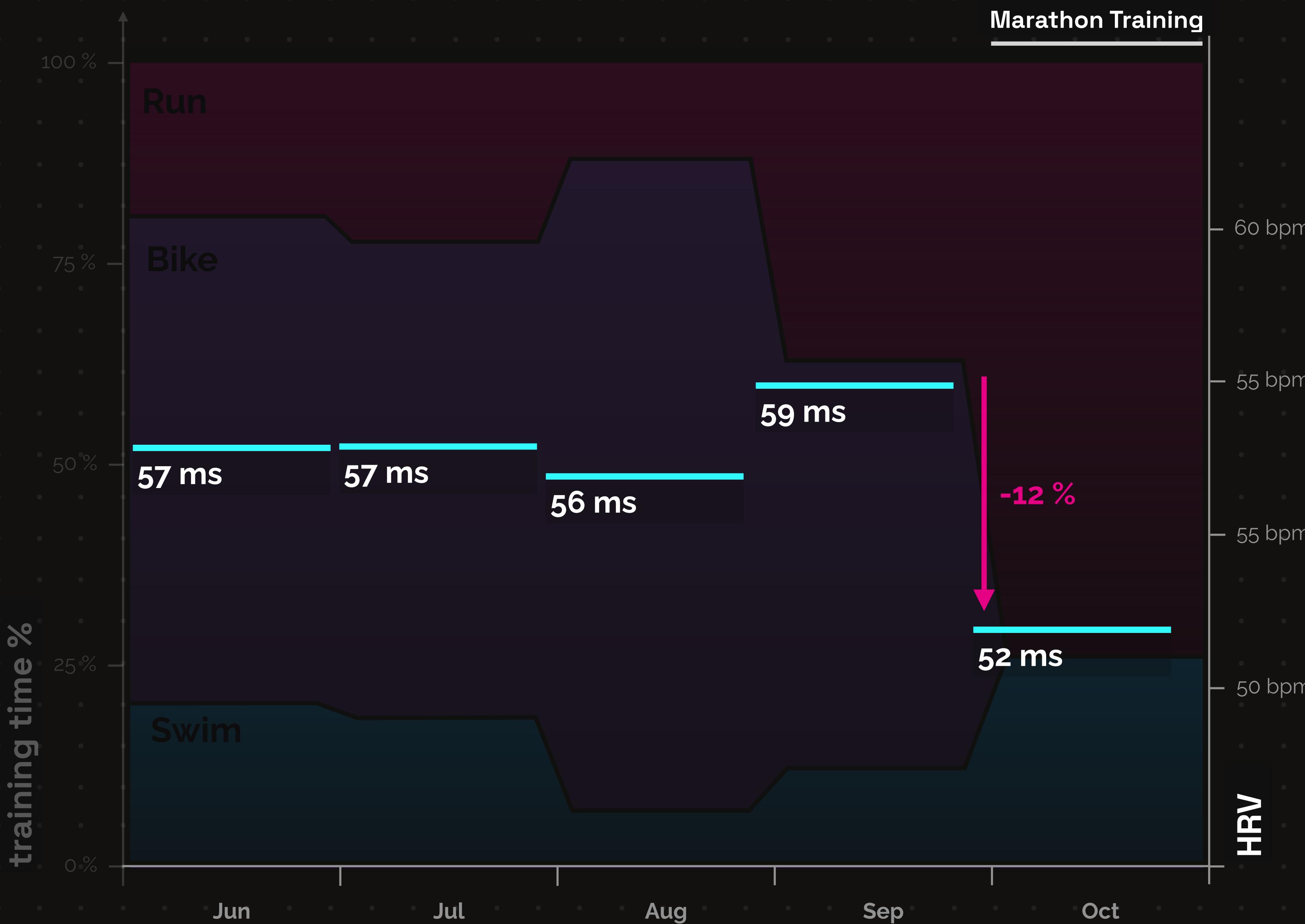
Looking back, the four months (Jun-Sep) before Nono was specifically training for his first marathon the distribution of the three sports in regard to training time was relatively consistent. However, the month before the marathon (October) he panic-trained and increased his amount of running drastically ...





# Marathon Training & HRV

Looking back, the four months (Jun-Sep) before Nono was specifically training for his first marathon the distribution of the three sports in regard to training time was relatively consistent. However, the month before the marathon (October) he panic-trained and increased his amount of running drastically which accordingly lowered his HRV drastically.



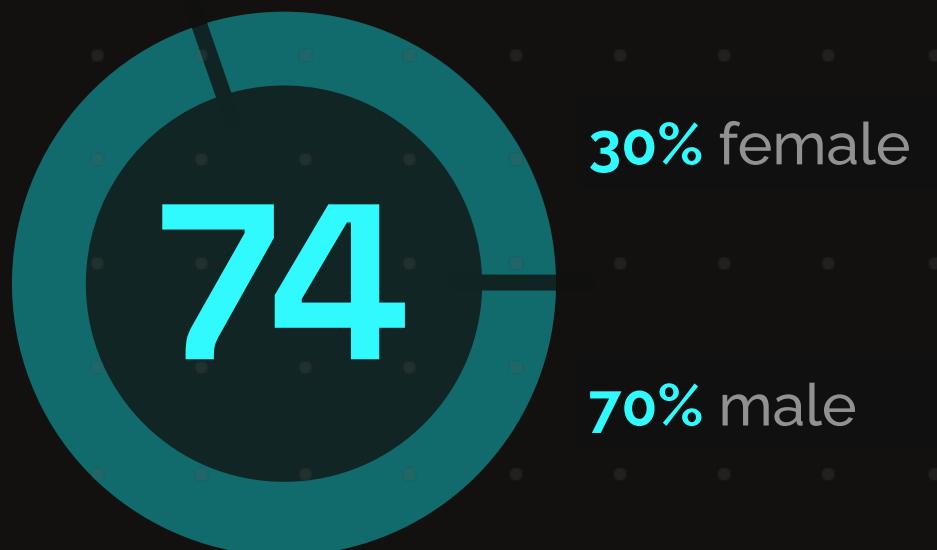
What factors  
increase injury risk?

## Study Data

# Run Injury Dataset

To better understand which factors lead to run injuries, we looked at an open source dataset of middle and long distances runners which accumulates 583 run injuries over a time period of 7 years.

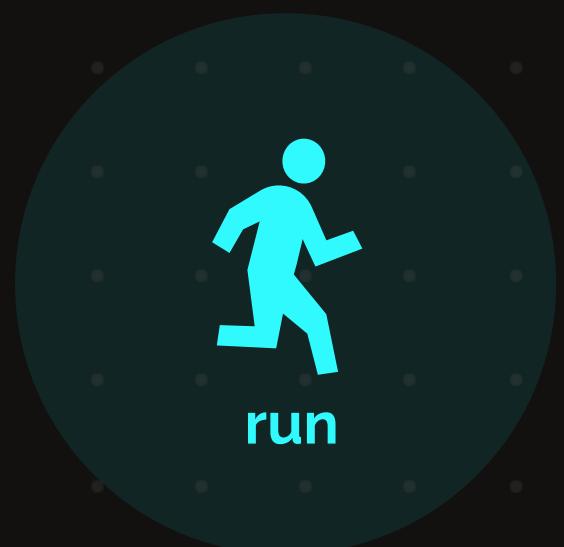
participants



time period



type of activity



number of injuries

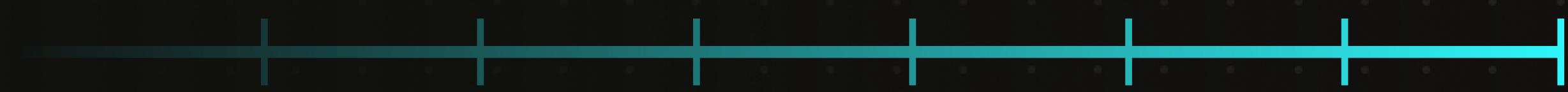


## SOURCE

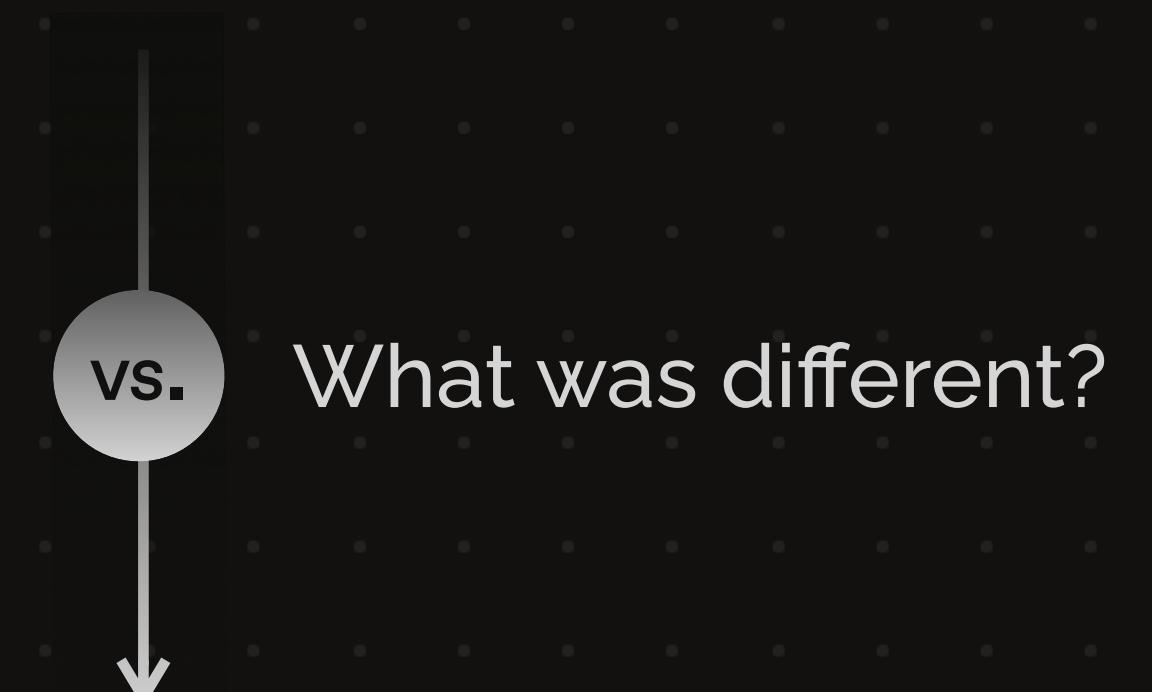
University of Groningen: „Injury Prediction in Competitive Runners with Machine Learning“ Lövdal, Safe, den Hartigh, Ruud, Azzopardi, George

# Logic of the Analysis

By dividing the dataset into two groups (normal weeks vs pre-injury-week) we can search for deviations in training patterns before an injury.



Normal Week



Pre-Injury Week



# Causes for Injury

Whilst a variety of factors may influence injuries, we are focussing only on four common metrics, which are often referred to as the most influential risk factors.

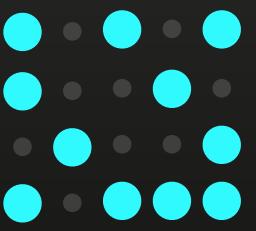
distance  
change in km



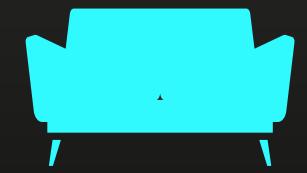
high intensity  
change in km



activities  
sum



rest days  
sum



Do these factors increase the risk of

Injury

Open Source Data



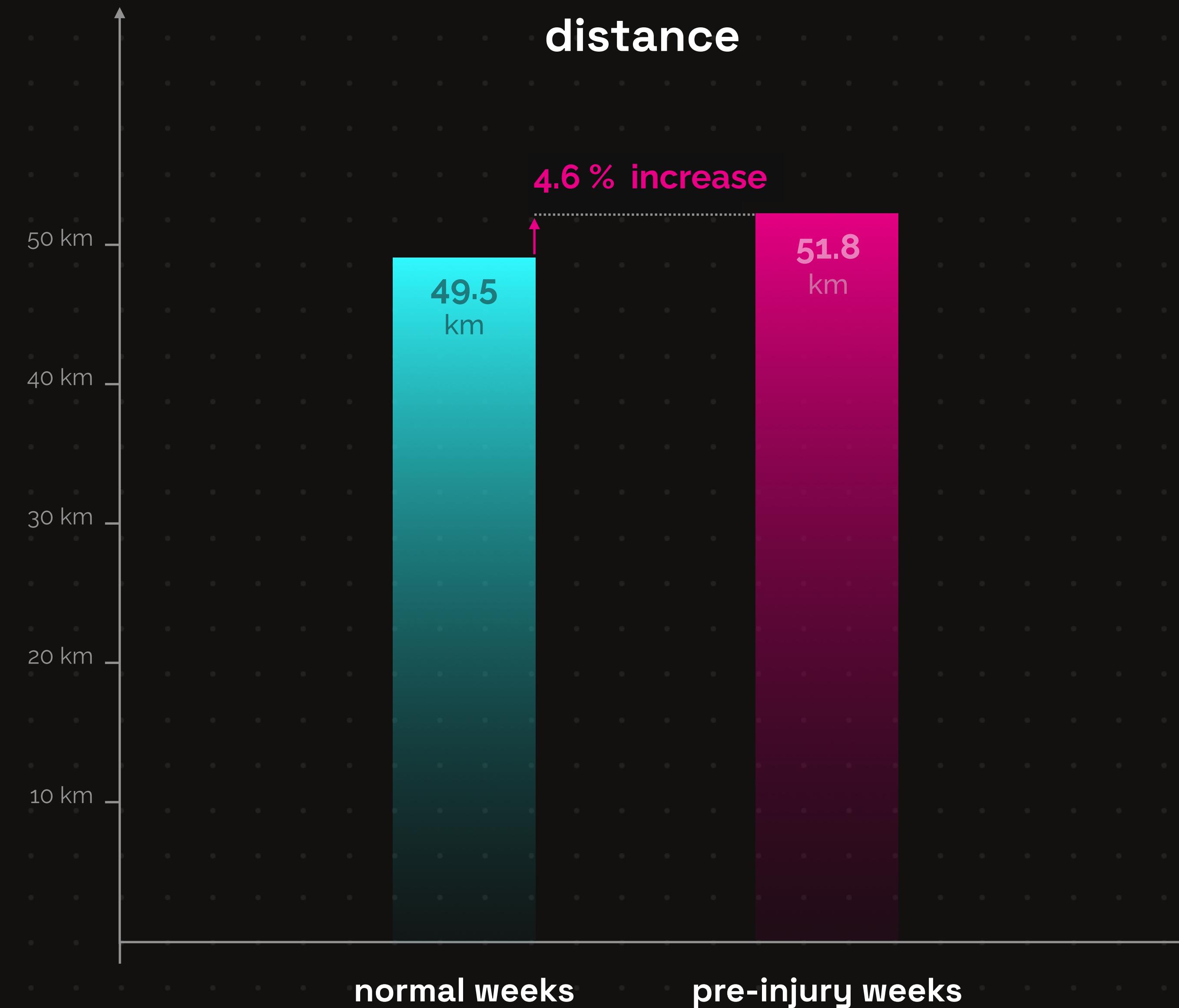
# Distance & Injury

For the average mileage a statistically significant increase by 4.6% in pre-injury weeks can be a warning signal for a higher injury risk.



**significant**  
p-value: 0.02

↗ more distance increases injury risk



n = 42 798 periods

Open Source Data



# High Intensity & Injury

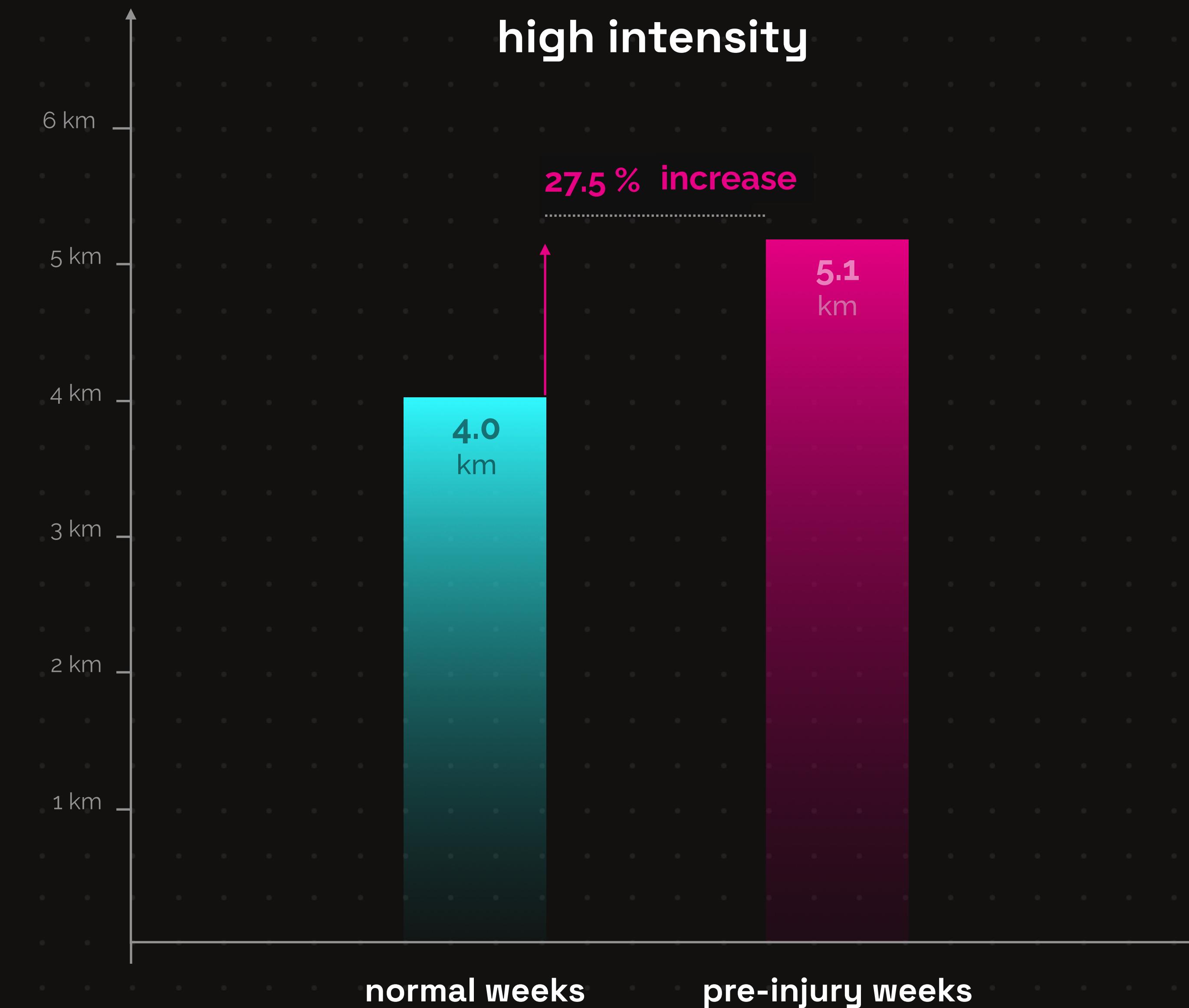
The significant increase by 27.5% for the average distance at high-intensity before an injury can also be considered to be a warning signal.



**significant**

p-value: 0.02

↗ more intensity increases injury risk



n = 42 798 periods

Open Source Data



# Number of Activities & Injury

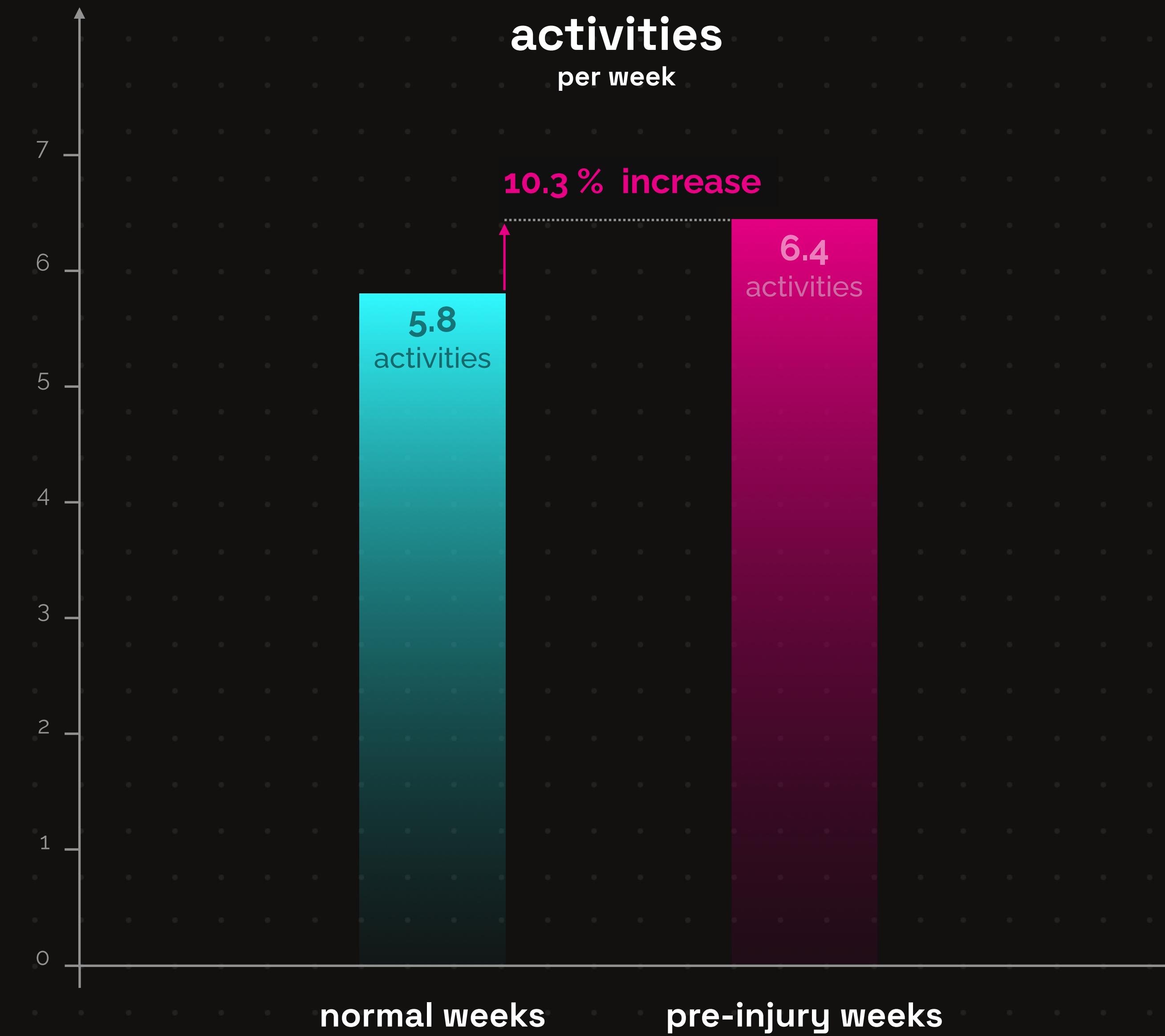
The rise of the average number of activities per week by 10.3% is a significant difference between normal and pre-injury weeks. Another warning signal.



**significant**

p-value: 0.02

↗ more activities  
increase injury risk



n = 42 798 periods



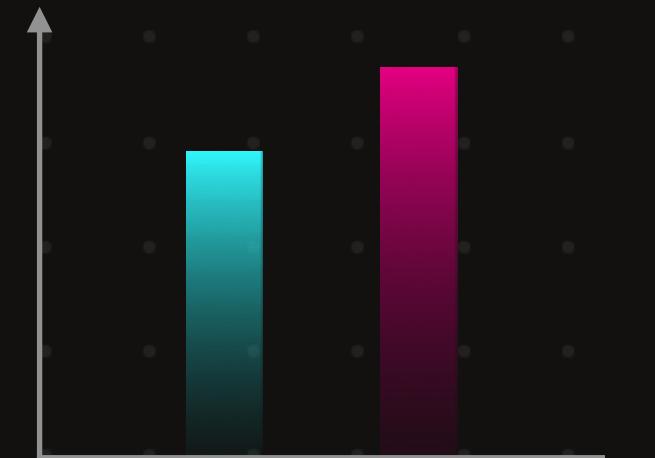
# Causes for Injury

In Nono's data we can only see a significant difference regarding the number of activities. In pre-injury weeks he increased his number of activities by 53% – which makes this metric a relevant injury risk factor, too.

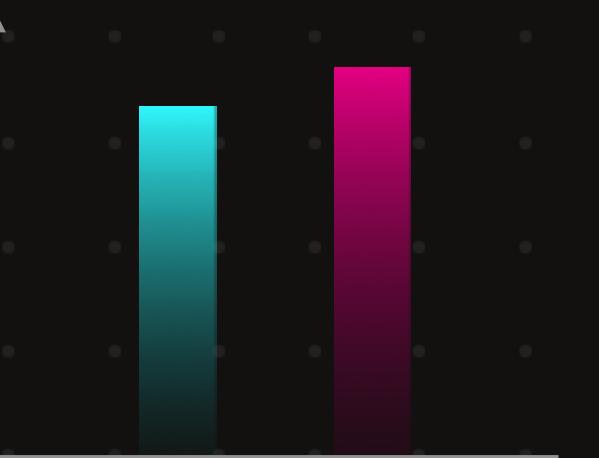
Open Source  
Data



more distance  
increases injury risk



more high intensity  
increases injury risk



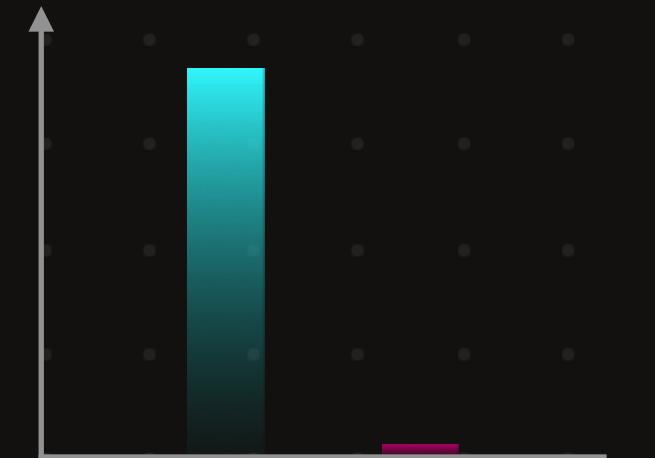
more activities  
increases injury risk

n = 42 798 periods

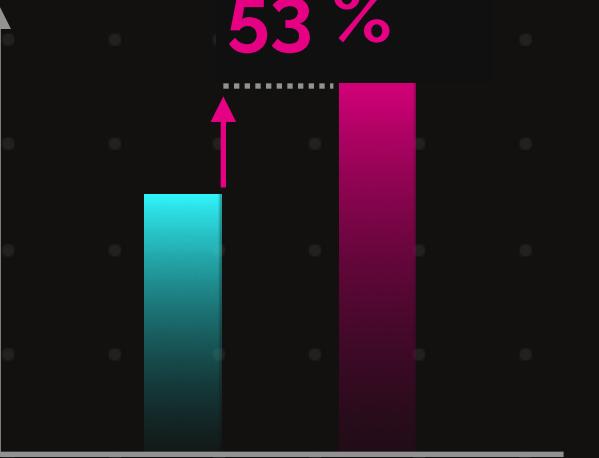
Nono's Data



no conclusion  
possible  
no significance



no conclusion  
possible  
not enough data



more activities  
increases injury risk

significant  
p-value: 0.045

n = 2 309 periods

# Lack of Rest Days & Injury

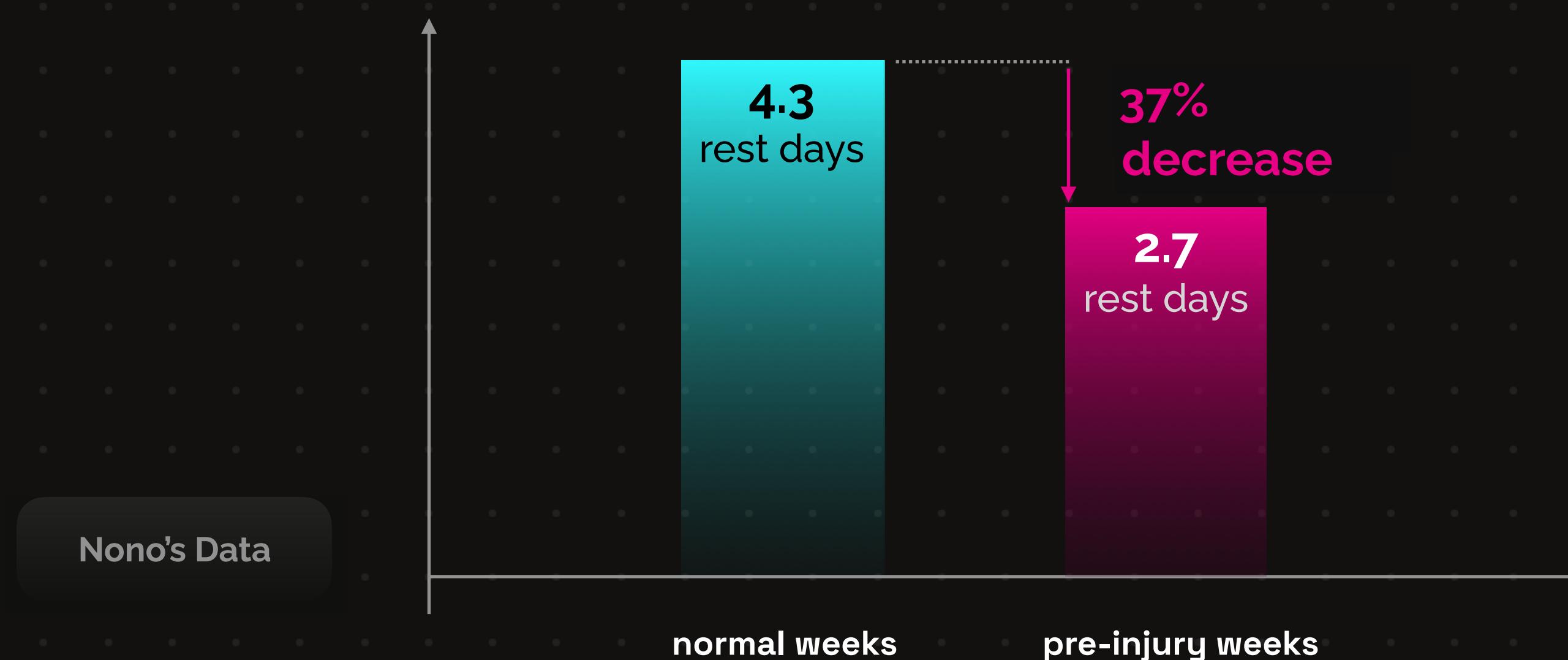
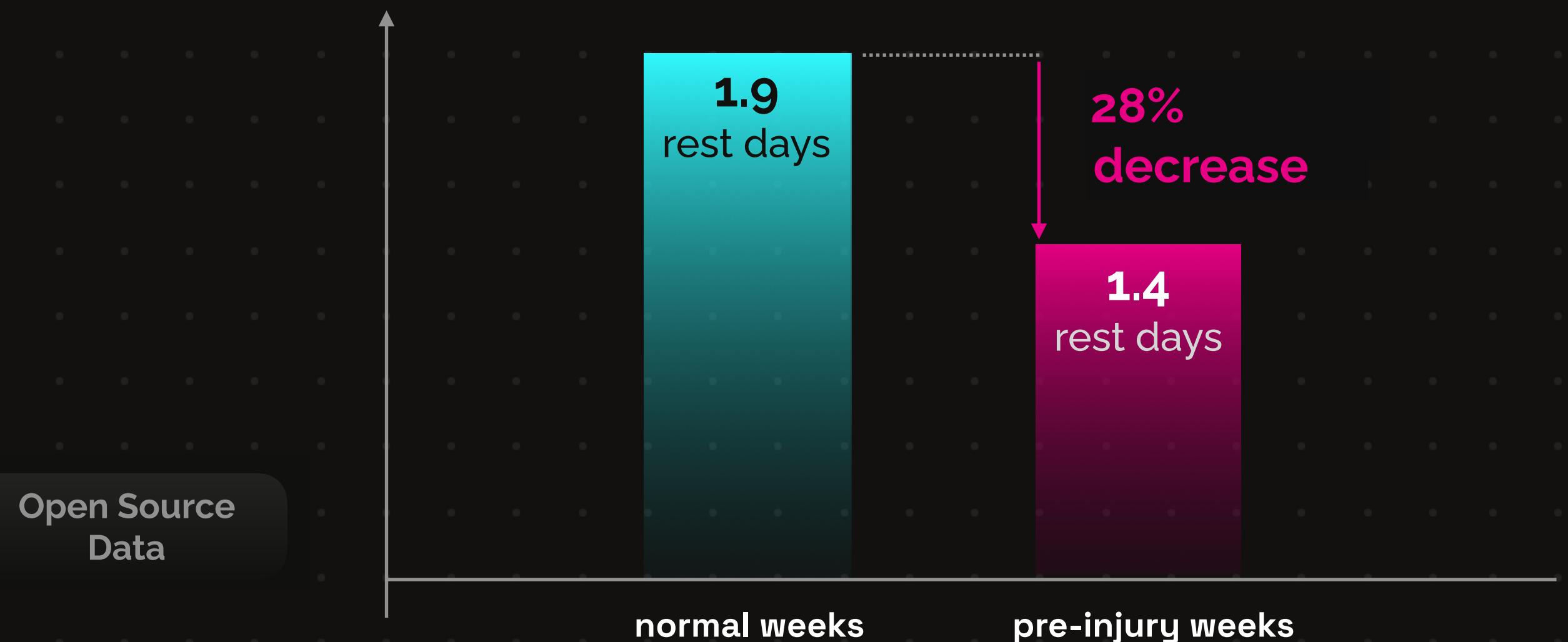
For rest days, we can see a significant decrease before injuries in both datasets. Hence, a lack of rest days can be considered as another injury risk factor.



**both significant**

p-value: 0.00

↗ **less rest days increases injury risk**



Open Source Data

Nono's Data

# Injury Patterns

Nono's data didn't show the same pattern as the average of the Open Source Dataset. However, by looking at each individual athlete in the open source data, we can see a variation of injury risk profiles for each athlete, too. Thus, whilst all four factors are risk factors in general, not all of them might be for every athlete.

different injury patterns  
for each athlete

distance

high int.

activities

rest days

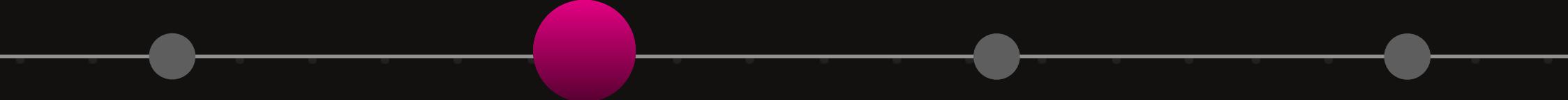
NONO



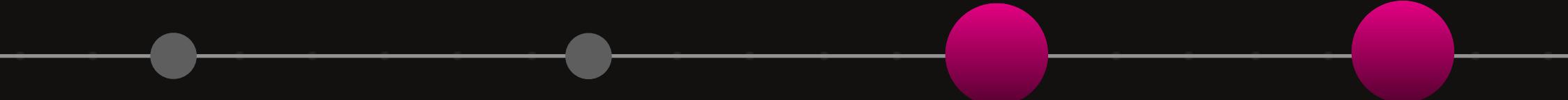
Athlete 1



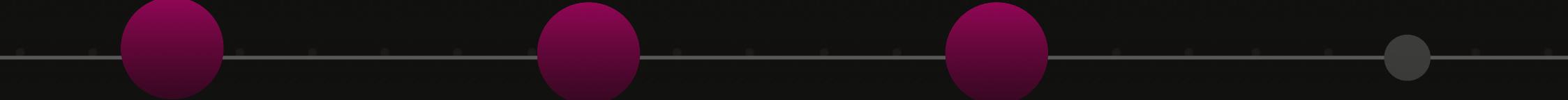
Athlete 2



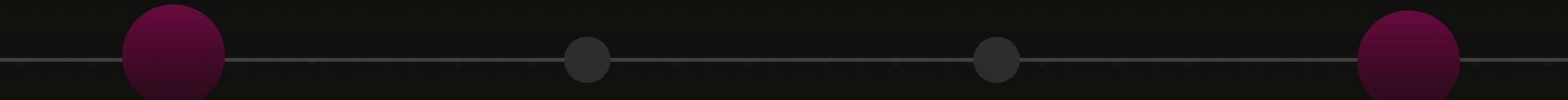
Athlete 3



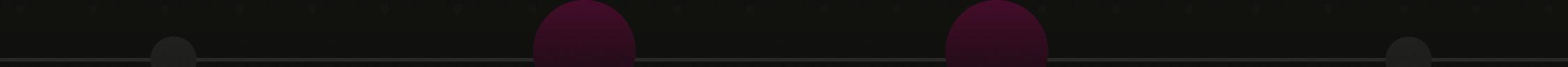
Athlete 4



Athlete 5



Athlete 6



Injury risk is  
**athlete-dependant**  
and is a combination of  
**multiple factors**

# Summary

# How to improve endurance sports performance?

## Incorporate high & low intensity

high intensity has a larger impact over a short period of time  
– but gains plateau the fitter you are

## Train consistently

by changing distance, training time and number of activities gradually  
– not drastically

## Injury patterns are individual

Injuries are athlete-dependant and can have a variety of factors

Intensity

increases

Consistency

decreases

Injury risk

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

←

# What Nono will incorporate in his future training

Incorporate high & low intensity

Prioritise low-intensity workouts during build-up

ca. 70% low int

# What Nono will incorporate in his future training

**Incorporate high & low intensity**

Prioritise high-intensity workouts just before races

ca. **70%** high int

**Train consistently**

Maximum 20% changes from activity to activity

max. **20%** change

**Injury patterns are individual**

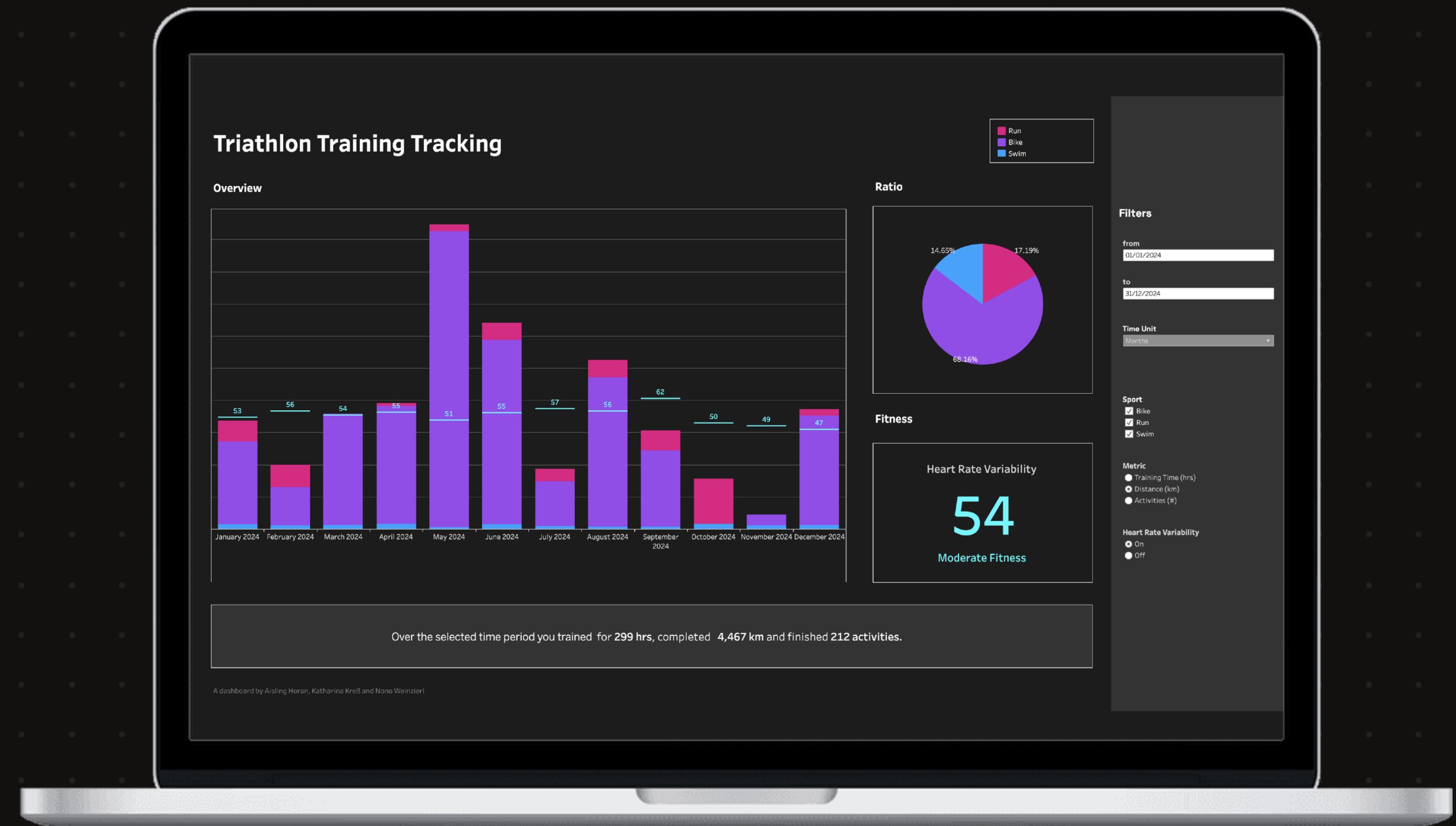
Minimum 2 rest days per week

min. **2** rest days

# Check out our Dashboard

Browse through more than 6 years of training data and find interesting patterns and extract learnings by yourself.

Visit Dashboard



Many thanks for  
your attention

# Appendix

# Let's connect on LinkedIn



**Aisling Horan**

Data Analysis

[linkedin.com/in/  
aisling-h/](https://linkedin.com/in/aisling-h/)



**Katharina Kreß**

Sales  
Data Analysis

[linkedin.com/in/  
kathkre/](https://linkedin.com/in/kathkre/)



**Nono Weinzierl**

UX & Service Design  
Data Analysis

[linkedin.com/in/  
nonow/](https://linkedin.com/in/nonow/)

# Project Roadmap



Final Presentation

# Tooling & Tech Stack

## Organisation & Administration

project management



communication



## Problem Definition



## Data Collection

personal data sources



public data sources



## Storing & Exchanging

storing



exchanging data & code



## Cleaning & Analysis

code editors



languages

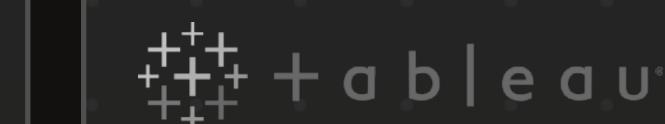


packages



## Data Visualisation

tools



languages



packages

seaborn

matplotlib

GeoPandas

## Presentation & Dashboard

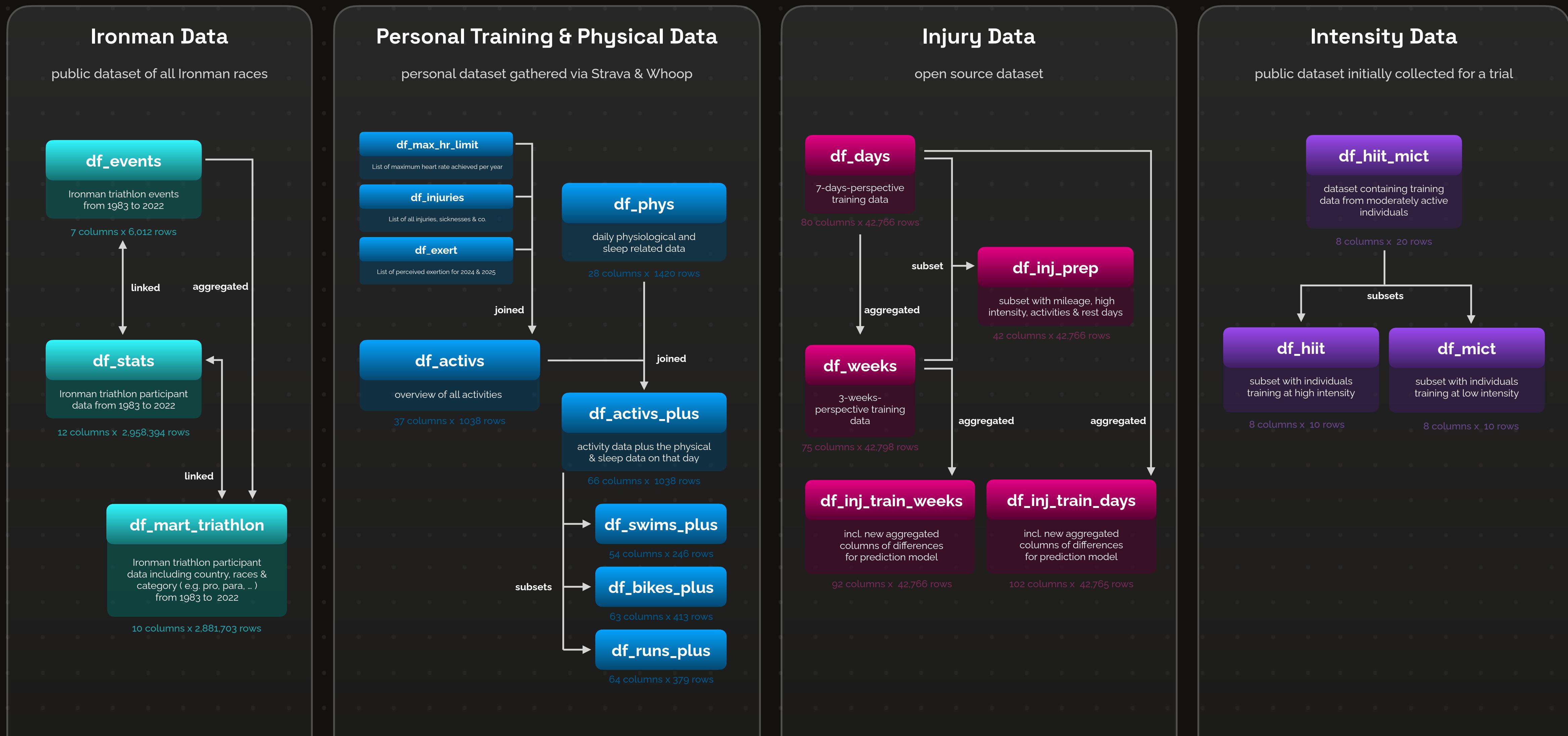
presentation



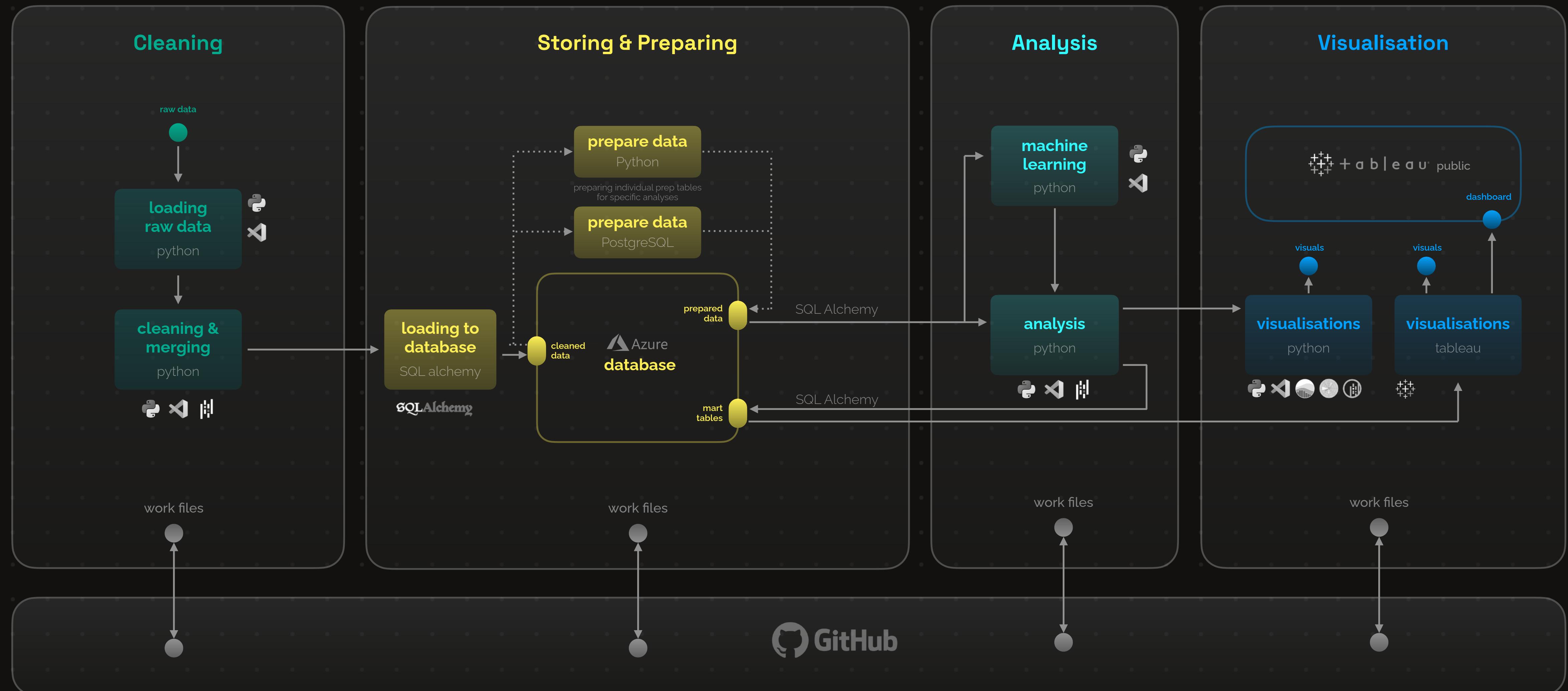
dashboard



# Datasets & Relations



# Data Architecture



# More Information & Material

## Presentation Recording

An engaging recording as YouTube Video of us presenting the results of our analysis

## Triathlon Dashboard

A visual overview over Nono's training data over the last six years

## GitHub Documentation

More information and material on our process, methods and outcomes

## Bootcamp Information

More insights on the three-months bootcamp by Neue Fische that we completed with this project