



HOE LOOP JE EEN MARATHON?

Of het nu gaat om korte sprintjes of het lopen van marathonaftanden, uiteindelijk wil je eruit halen wat erin zit en optimaal presteren of genieten. Tijdens een marathon moet een loper zijn krachten goed verdelen over de gehele afstand.

Te snel starten kan zorgen voor voortijdige vermoeidheid, terwijl te traag starten niet tot een toptijd zal leiden. Aan het einde van de rit is het fijn de beschikbare energetische reserves optimaal te hebben gebruikt. Dit proces noemen we ook wel 'pacing' (tempo) strategie.

BRYAN VAN INGEN

Een marathon loop je niet zomaar, hier gaat veel training en voorbereiding aan vooraf. Het zou zonde zijn als een loper daarna langzamer dan gewenst, of zelfs helemaal niet kan finishen. Niet iedereen kan een marathon lopen zoals Eluid Kipchoge of Lornah Kiplagat maar elke sporter wil wel bij elke race zoveel mogelijk uit zichzelf halen. Daarbij is het belangrijk dat elke loper – recreant of professional – met een plan aan de start verschijnt. Tijdens dit onderzoek is gekeken naar welke pacingstrategieën het vaakst door marathonlopers worden ingezet.

Hardlopen wordt steeds populairder. In tijden van lockdowns en quarantaines is hardlopen een van de weinige sporten die men nog ongehinderd kan blijven beoefenen. Ook recreatieve hardlopers pakken de trainingen serieus aan. Veel hardlopers hebben uiteindelijk één ultiem doel: de marathon. Steeds meer sporters leggen

zich toe op de magische afstand van 42,2 kilometer. Eén vraag is voor zowel de recreatieve als de professionele sporter heel belangrijk: 'Hoe moet een loper zijn race indelen om zo snel mogelijk de finish te bereiken?'

Data verzamelen

Le Champion, de organisator van de Amsterdamse Marathon, plaatst de uitslagen van de marathon elk jaar online. Sport Data Valley heeft vervolgens in samenwerking met Amsterdam Data Science de data van 2016 tot en met 2019 samengevoegd voor hun Sport Data Challenge 2019. Deze data zijn voor het onderzoek gebruikt.

De data bestaan grotendeels uit tussentijden van de lopers. Van alle hardlopers die meededen aan de Amsterdamse

marathon tussen 2016 en 2019 is bekend wat hun eindtijd is vergeleken met het startschot, en vergeleken met de daadwerkelijke passage van de startlijn. Hierbij zijn tijdsmetingen van de tussentijden van de lopers na elke vijf kilometer ook bekend. Ook weten we van elk individu het geslacht en de leeftijdscategorie waarin de loper valt. Deze leeftijdscategorieën beginnen bij de senioren (18 tot en met 34 jaar oud) en volgen hierna steeds in intervallen van vijf jaar. In totaal zijn de tijden van 49.245 lopers over de vier jaar gebruikt voor het onderzoek, 37.441 mannen en 11.804 vrouwen.

Voordat we de strategieën van de lopers onderzoeken, kijken we eerst naar de eindtijden van de deelnemers. Hier zien we dat, zoals was te verwachten, mannen met een gemiddelde eindtijd van 3 uur en 56 minuten, gemiddeld bijna een half uur sneller lopen dan vrouwen. Vrouwen hadden gemiddeld 4 uur en 21 minuten nodig om hun 42,2 kilometer af te leggen. Bij de leeftijdsgroepen vinden we wel iets opmerkelijks. Vooraf zou men verwachten dat de gemiddelde snelheid af zou nemen naar mate de leeftijd hoger wordt. Echter, wat bleek, de lopers van 35 tot en met 39 jaar deden niet onder voor de jongere lopers. Ze hadden een gemiddelde snelheid van ongeveer 11 kilometer per uur. In de groepen ouder dan 40 vonden we wel steeds langzamere tijden.

De pacingstrategieën

Voor het onderzoek is de pacing van een marathonloper gedefinieerd als het verloop van de gemiddelde snelheden over de intervallen waarvan de tussentijden bekend zijn. Om alle lopers op eenzelfde manier te kunnen analyseren, wordt er voor het bepalen van de strategieën niet gekeken naar de ware snelheden per interval. In plaats hiervan maken we gebruik van het verschil tussen de gemiddelde snelheid van een loper in ieder interval en hun gemiddelde snelheid over de gehele race. Door deze nieuwe variabele te gebruiken, zou er qua strategie geen verschil moeten zijn tussen iemand die constant 20 km per uur loopt en iemand die constant 8 km per uur loopt.

De pacingstrategieën worden gevonden aan de hand van een clusteranalyse. Hierbij is er een k-means-algoritme gebruikt op de nieuwe variabelen. Na de data genormaliseerd te hebben, kiest dit algoritme meerdere malen k willekeurige gemiddelden (één voor ieder cluster) en bepaalt vervolgens aan welk cluster een persoon toegewezen moet worden zodat de variantie tussen de variabelen binnen de clusters zo laag mogelijk ligt. Zodra iedere loper aan een cluster is toegewezen, worden k nieuwe gemiddelden bepaald op basis van de gemaakte clusters en

wordt iedere loper opnieuw aan een cluster toegewezen. Dit proces wordt 25 keer herhaald om te verzekeren dat de gevonden clusters een zo laag mogelijke intra-cluster variantie hebben, ongeacht welke gemiddelden gekozen werden bij de initiatie van het algoritme.

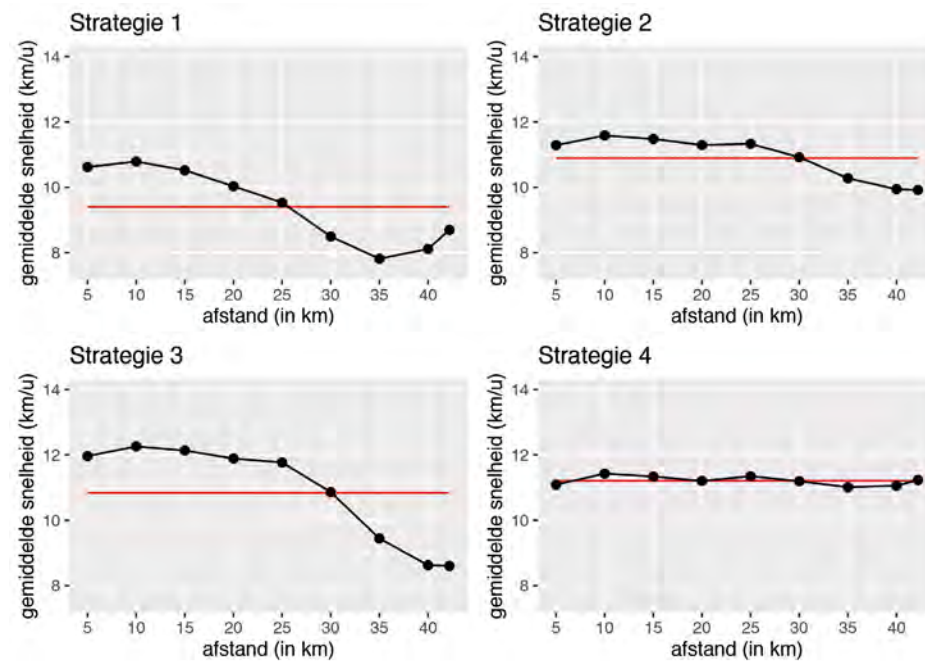
De keuze voor het aantal pacingstrategieën waarmee wij verder werken is afhankelijk van de mate waarin de strategieën onderling van elkaar verschillen. We beginnen met twee strategieën en vervolgens verhogen we dit meermaals met één om de effecten te bekijken. Hierin wordt er gezocht naar grote verschillen tussen de strategieën die gevonden werden. Verschillen zoals lopers in cluster A starten even hard, maar hebben een eindsnelheid van een halve kilometer per uur langzamer dan lopers in cluster B worden niet als onderscheidend genoeg gevonden om mee te nemen. Het blijkt dat als we werken met 4 clusters, en dus 4 strategieën, we het meeste aantal strategieën hebben waar een duidelijk verschil tussen te zien is. De gemiddelde opbouw van de race van lopers in ieder cluster, is weergegeven in figuur 1.

Op basis van wat we in figuur 1 kunnen zien, zijn de vier pacingstrategieën als volgt gedefinieerd:

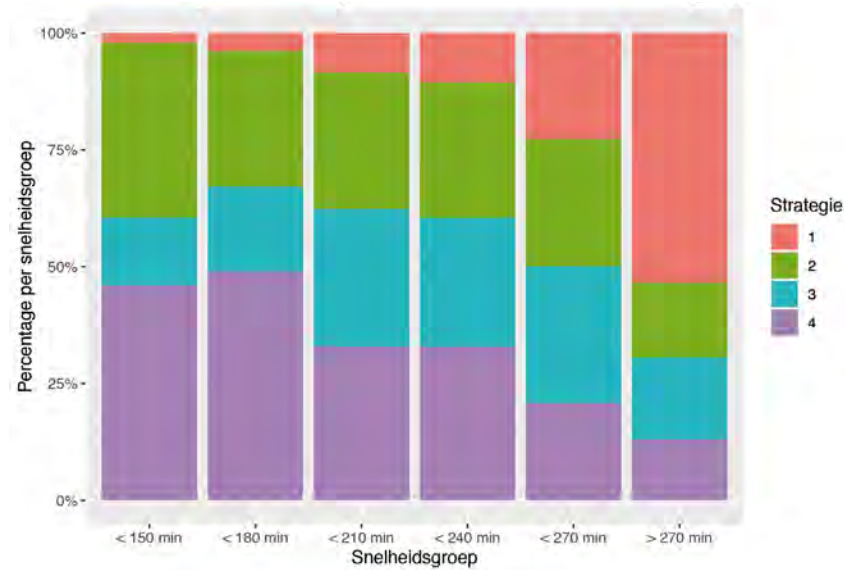
1. De loper start snel, maar gedurende de race verliest hij een groot deel van zijn snelheid. Aan het eind kan de loper er nog wel een korte versnelling uit halen.
2. De loper probeert constant te lopen, de tweede helft zakt het tempo echter wel enigszins in.
3. De loper begint snel, maar in de tweede helft van de race stort de snelheid compleet in.
4. De loper houdt een constante snelheid aan gedurende de race.

In tabel 1 staat weergegeven door hoeveel marathonlopers iedere pacingstrategie gevolgd werd. We zien hier dus dat het grootste deel van de lopers op zijn minst probeert zijn race constant te lopen. Bij de ruim 18.000 lopers die volgens de tweede strategie lopen, zien we dat ze dit helaas niet hebben kunnen volhouden tot de finish.

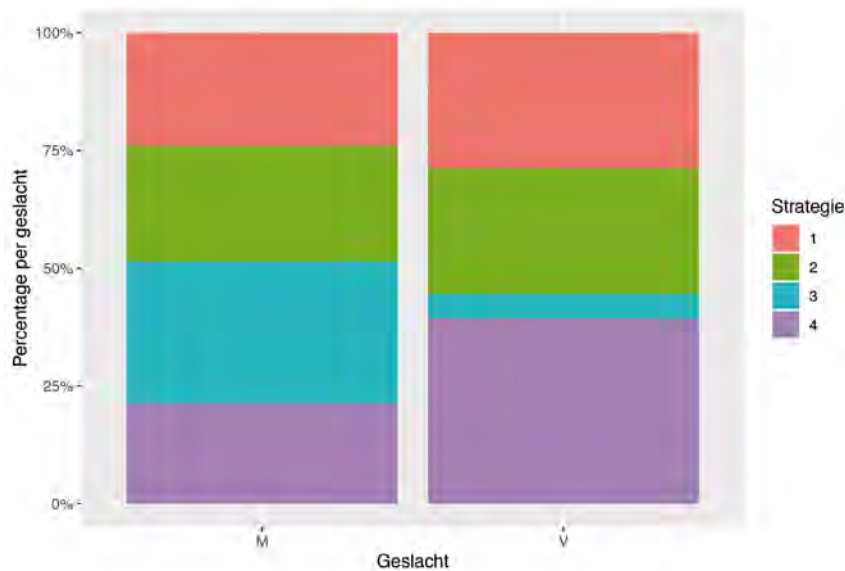
Aangezien we nu onze strategieën hebben, kunnen we kijken hoe de lopers uit verschillende groepen zich verdelen over deze strategieën. Om te kunnen zeggen welke strategie het best lijkt te werken, zullen we de deelnemers eerst in groepen indelen aan de hand van hun eindtijd. Bij deze verdeling beginnen we bij een eindtijd van 2 uur en maken we voor ieder half uur een nieuwe groep. De langzaamste groep zal de groep zijn waarin we iedereen indeelden die langzamer dan viereneenhalf uur deed over de marathon. Natuurlijk zeggen we niet dat dit slechte tijden zijn, maar door de deelnemers op deze manier te verdelen, leggen we de focus bij de snelste groepen en kan er goed gezien worden of daar grote verschillen lig-



Figuur 1. De gemiddelde opbouw van een marathon per pacingstrategie. De rode lijn representeert de gemiddelde snelheid van alle lopers die de strategie volgen



Figuur 2. De verdeling van de pacingstrategieën over de snelheidsgroepen, weergegeven als percentage van elke groep



Figuur 3. De verdeling van de pacingstrategieën per geslacht, weergegeven als percentage van elk geslacht

STRATEGIE	1	2	3	4
AANTAL MARATHONLOPERS	7.580	18.434	5.481	17.750
PROCENTUEEL VAN TOTAAL	15,4%	37,4%	11,1%	36,0%

Tabel 1: Het aantal marathonlopers per pacingstrategie met percentage van totaal.

gen. In figuur 2 zien we per snelheidsgroep de verdeling over de strategieën.

In de figuur is duidelijk te zien dat naarmate de eindtijden sneller worden, de lopers vaker strategie 2 of 4 lijken te volgen, terwijl bij langzamere lopers strategie 1 juist steeds dominant wordt. Hieruit kan geconcludeerd worden dat snelle lopers vaker constanter lopen. Het is wel belangrijk dat we ons hierbij realiseren dat dit onderzoek niet zegt wat de oorzaak en wat het gevolg is in dit verband. We kunnen niet op basis van deze resultaten zeggen of mensen die een eindsprint proberen te doen, daardoor langzamer lopen, of dat mensen die langzamer lopen vaker nog energie over hebben voor een eindsprint.

In figuren 3 en 4, zien we op dezelfde manier de verdelingen per geslacht en leeftijdscategorie respectievelijk. Hetgeen dat het meest opvalt in figuur 3, is dat bij mannen strategie 3 dominant is, terwijl vrouwen vaker strategie 4 volgen. Mannen mogen dan dus wel sneller lopen, ze schatten zichzelf ook vaak iets te goed in waardoor ze instorten, terwijl vrouwen op een constant tempo naar de finish gaan.

Bij de verschillende leeftijdsgroepen valt op dat strategie 4 steeds minder vaak voorkomt bij hogere leeftijden. Het lijkt er dus op dat oudere deelnemers steeds meer moeite krijgen met het aanhouden van een constante snelheid. Dit gaat gepaard met een groei bij strategie 2. Oudere lopers hebben dus steeds meer moeite met het vasthouden van een constant tempo, maar ze storten niet

compleet in. Het lukt lijkt de ouderen nog steeds goed te lukken om hun snelheid vast te houden, ze verzwakken alleen licht in de tweede helft van de marathon.

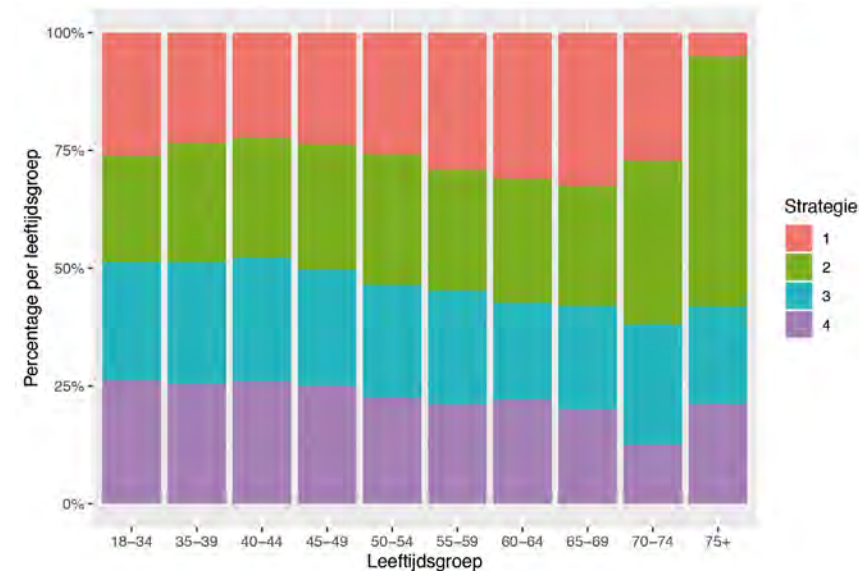
Conclusie

Geslacht, leeftijd en de gemiddelde snelheid van een loper hebben alle drie samenhang met de pacingstrategieën van marathonlopers. Over het algemeen kunnen we zien dat lopers met een snellere eindtijd, en dus een hogere gemiddelde snelheid, vaker een constante pacing aanhouden. Bij verschillende leeftijden zien we dat oudere deelnemers vaker moeite hebben het vasthouden van hun snelheid en daardoor gedurende de tweede helft van de race vertragen. Bij de geslachten zien we dat vrouwen vaker constanter lopen dan mannen, die een neiging hebben om te hard te beginnen en vervolgens instorten.

BRONNEN

Dataset: Sport Data Valley, (2019) , Github repository. <https://github.com/sportdatavalley/sport-data-challenge-2019>

BRYAN VAN INGEN studeert Toegepaste Wiskunde aan de Hogeschool van Amsterdam. Met drie studiegenoten analyseerde hij in opdracht van Sport Data Valley, data van de Amsterdamse Marathon. Dit artikel is een samenvatting van het onderzoek. Zie voor een kort filmpje <https://bit.ly/3qC7LNH>. E-mail: Bryan@famingen.nl



Figuur 4. De verdeling van de pacingstrategieën over de leeftijdscategorieën, weergegeven als percentage van elke categorie