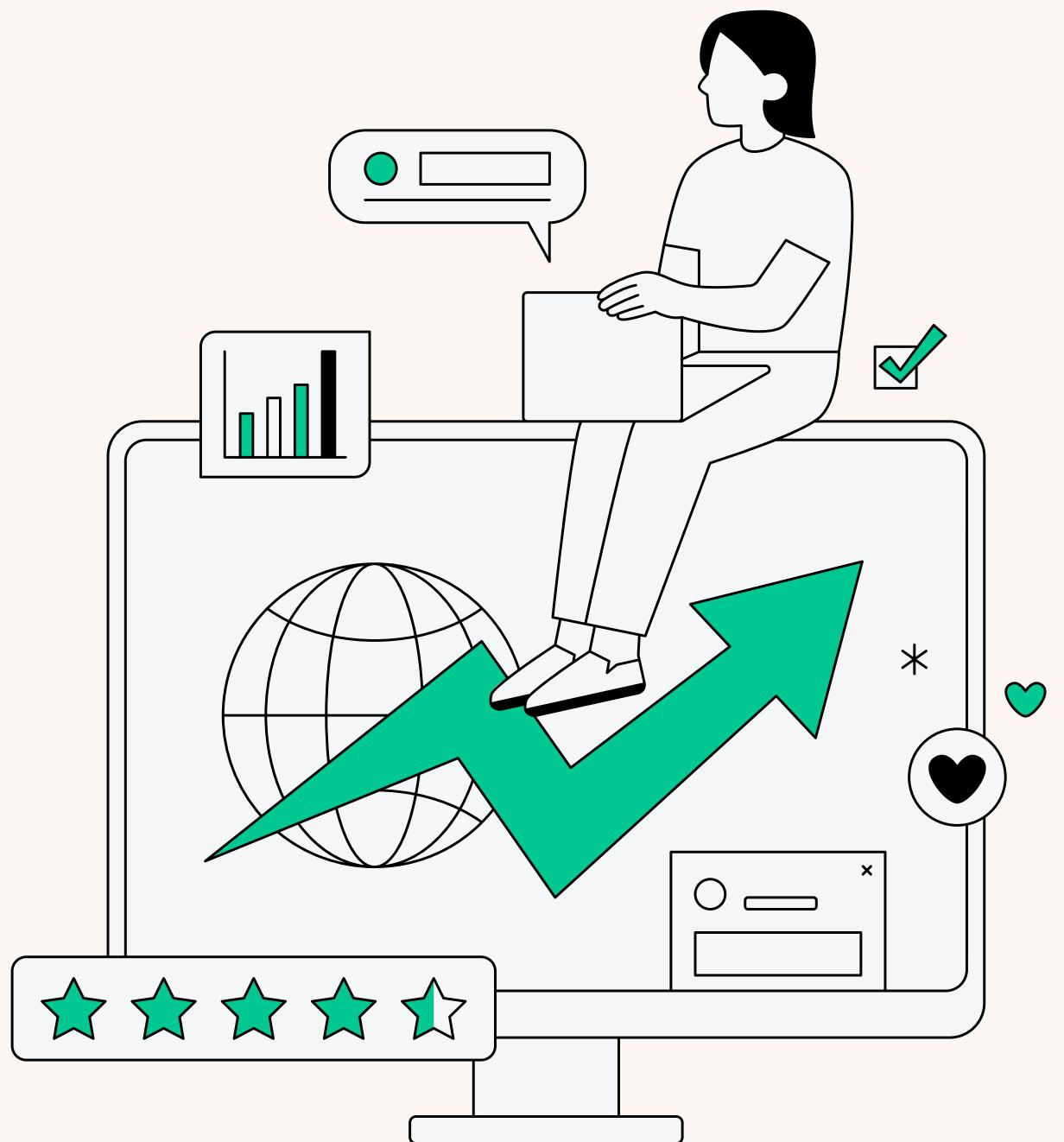


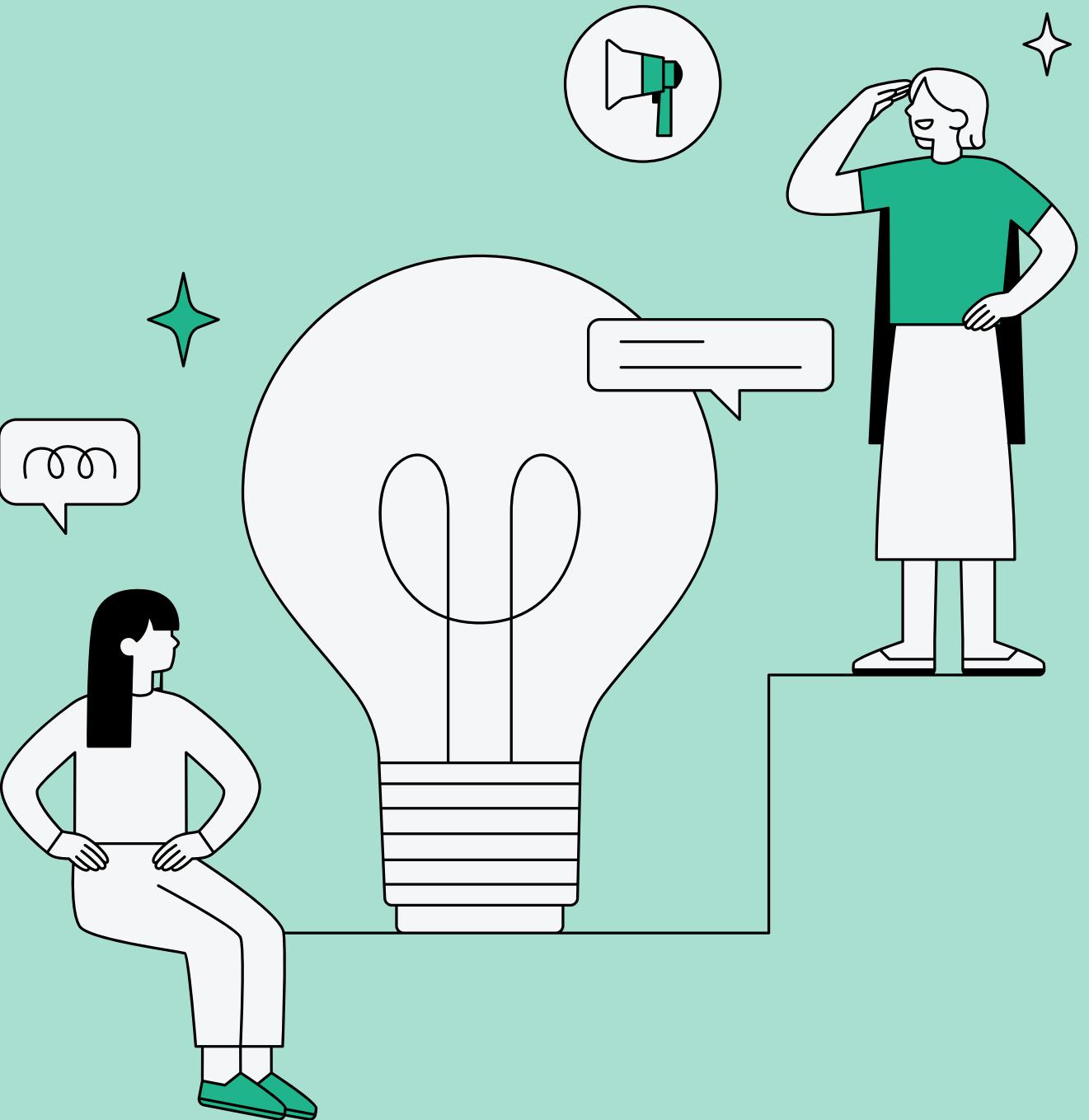
Ultramarathons in the USA: Analyzing Trends and Participation in 2020

Asher Hagan



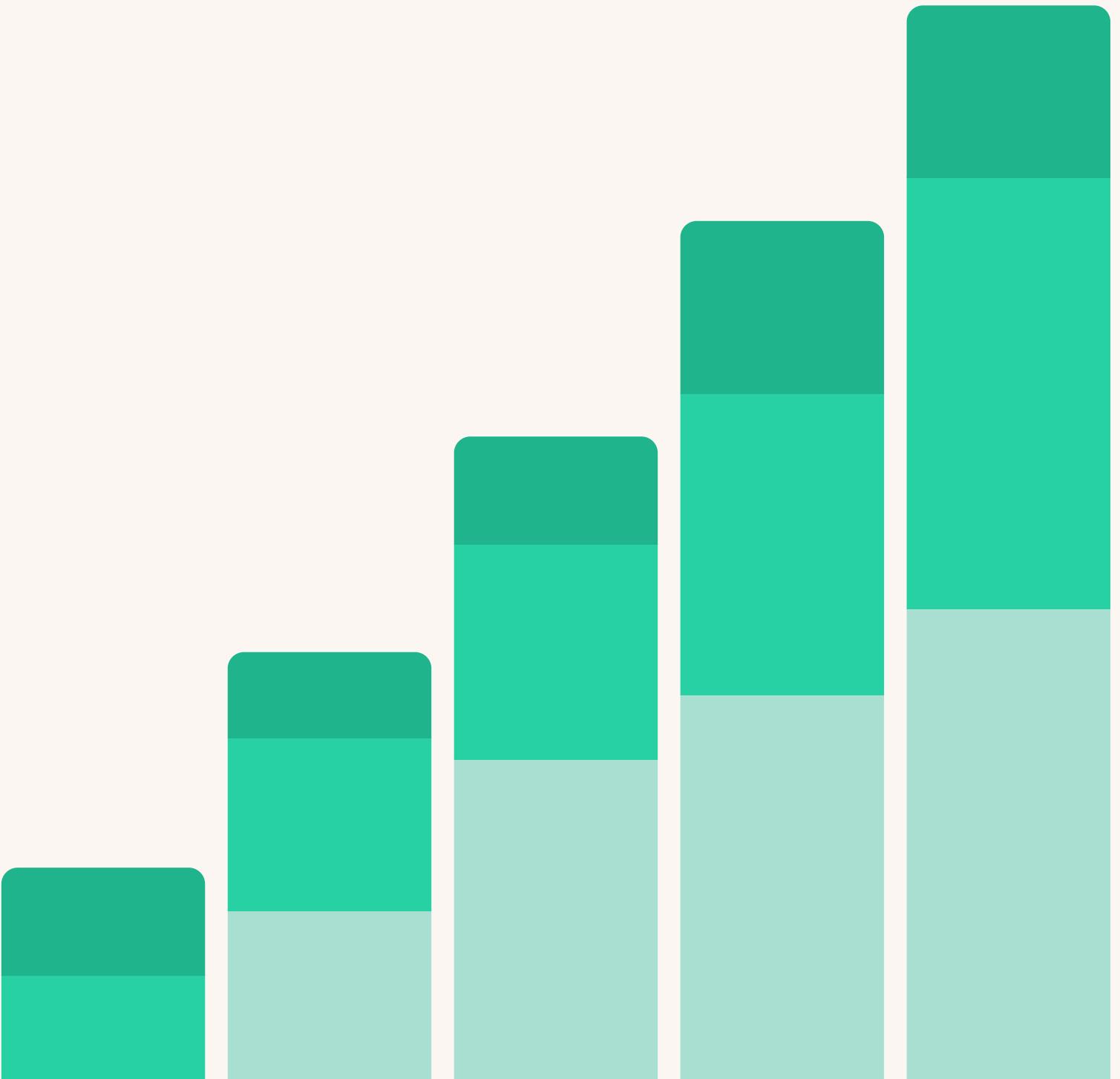
Introduction to results analysis

I wanted to see how different aspects of runners effect their performance during ultra marathons. This report hopes to answer that question. I got my data from a Kaggle data set that had running entries from all the way back to 1995. This is the data set [Kaggle.com](https://www.kaggle.com).



Methodology used in the analysis

The main tools that I used to complete this analysis was python with the pandas library for the data manipulation, seaborn for the data visualizations, and I coded it all up in a Jupyter notebook.



Main challenges identified

01.

The data set was large. The first step I took was to filter the data down to only the 50km and 50mi races, then just races in the USA, and finally filtered down to just races in the 2020's. This brought the data set from 746195 rows to 25857 rows.

02.

The data set had some unnecessary information that made the data set look worst. First I took the USA off of the location because all of the races were in the USA. I also drop the unnecessary columns and empty rows.

03.

The data had column labels that did fit what was in the column. I renamed them by taking the space out to make it easier to type and reordered them to have the most important information up front.

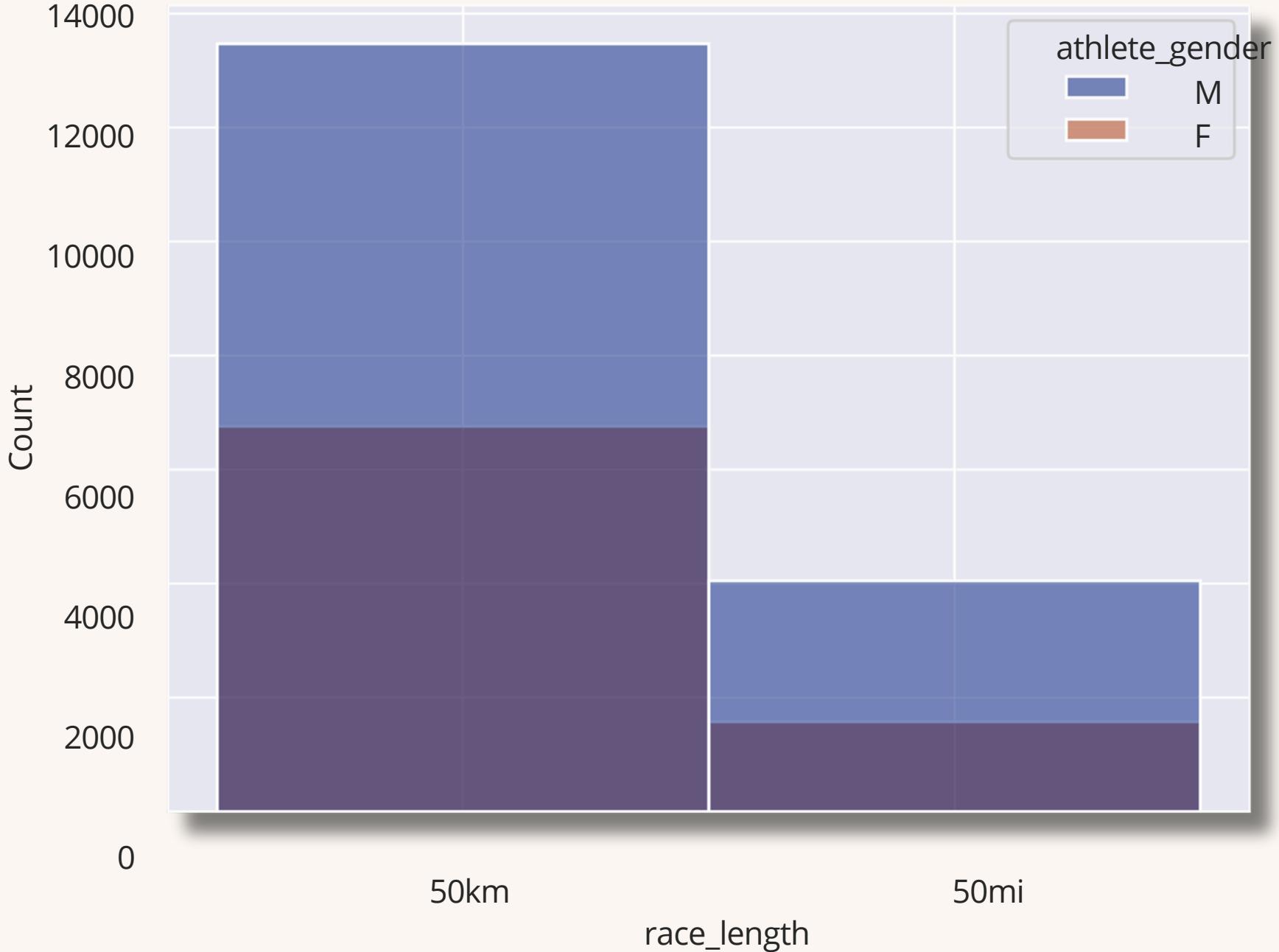


Question from the data set

I want to answer four main things for the data, what is the difference between male and female runners, what is the best age for runners, what are the worst age group for running, if someone wanted to start running what season is the best for their first race?



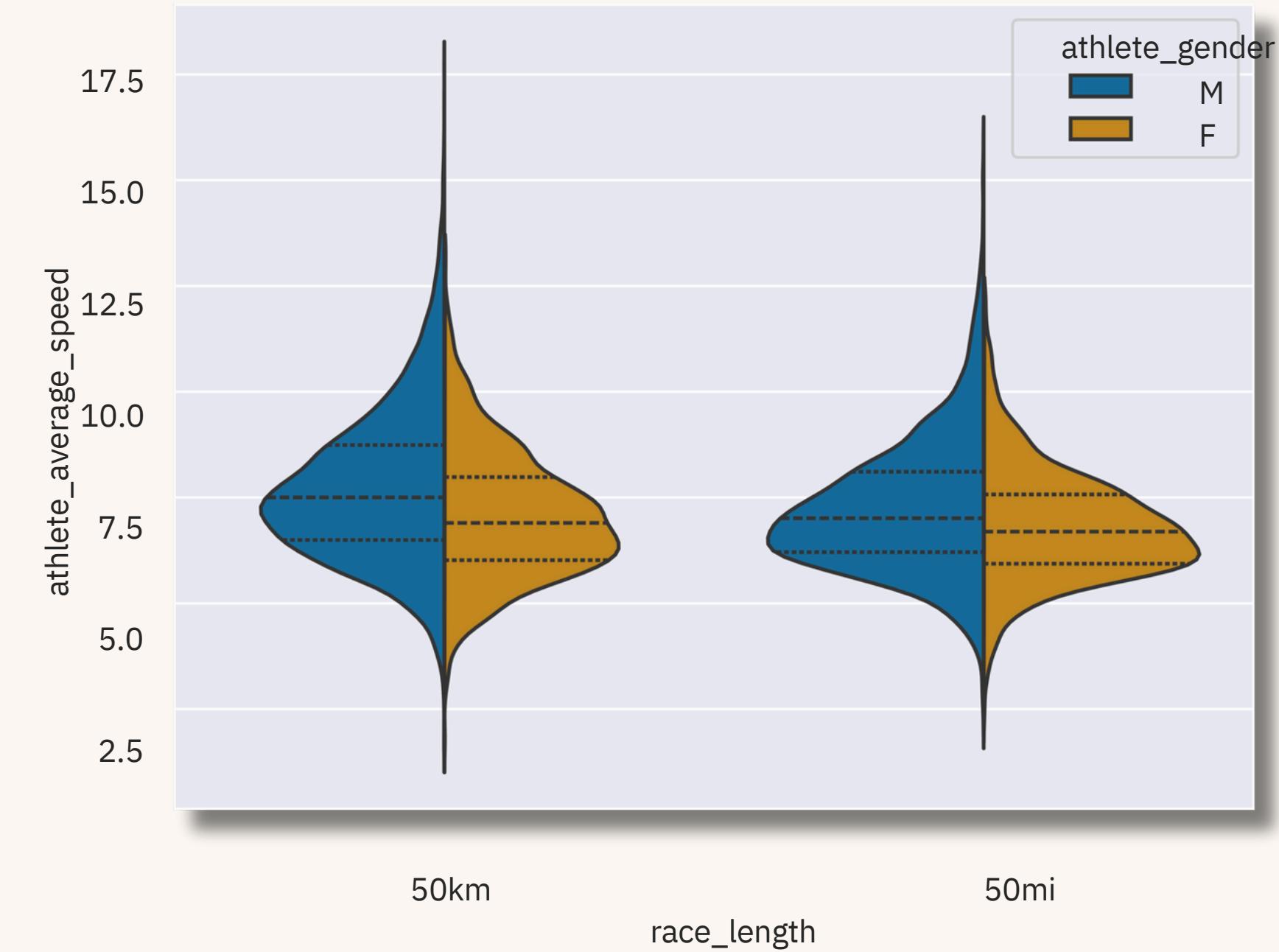
Male vs Female Count on 50mi vs 50KM



50km is more popular than 50 mi

Looking at the two different lengths of races the 50km is split close to half males and half females. The 50mi starts to shift more towards male athletes.

Average Speed of Male and Female Runners



Speed gets Closer as Races get Longer

During the shorter races the difference between males and females in the 50km is significant, but as the races get longer difference between males and females gets smaller and smaller

Comparison

	mean	count
athlete_age		
29	7.902252	135
23	7.779800	55
28	7.575252	107
30	7.569204	157
25	7.540923	91
31	7.451638	138
38	7.430022	231
35	7.422359	195
36	7.403854	185
26	7.379800	75

The age group with the quickest average time is runners age 29.

This makes since because the runners are usually not new to running. They have had time to work on not only the running part but also the metal part of just running that far.

	mean	count
athlete_age		
70	5.470667	12
65	5.934786	14
68	5.957500	10
67	6.114909	11
66	6.194300	10
60	6.261788	33
62	6.272730	37
61	6.358355	31
63	6.514806	31
58	6.582328	67

The age group with the slowest average time is runners age 70.

This makes since because the runners have been runner for a while but there body is no longer in the prime running years.

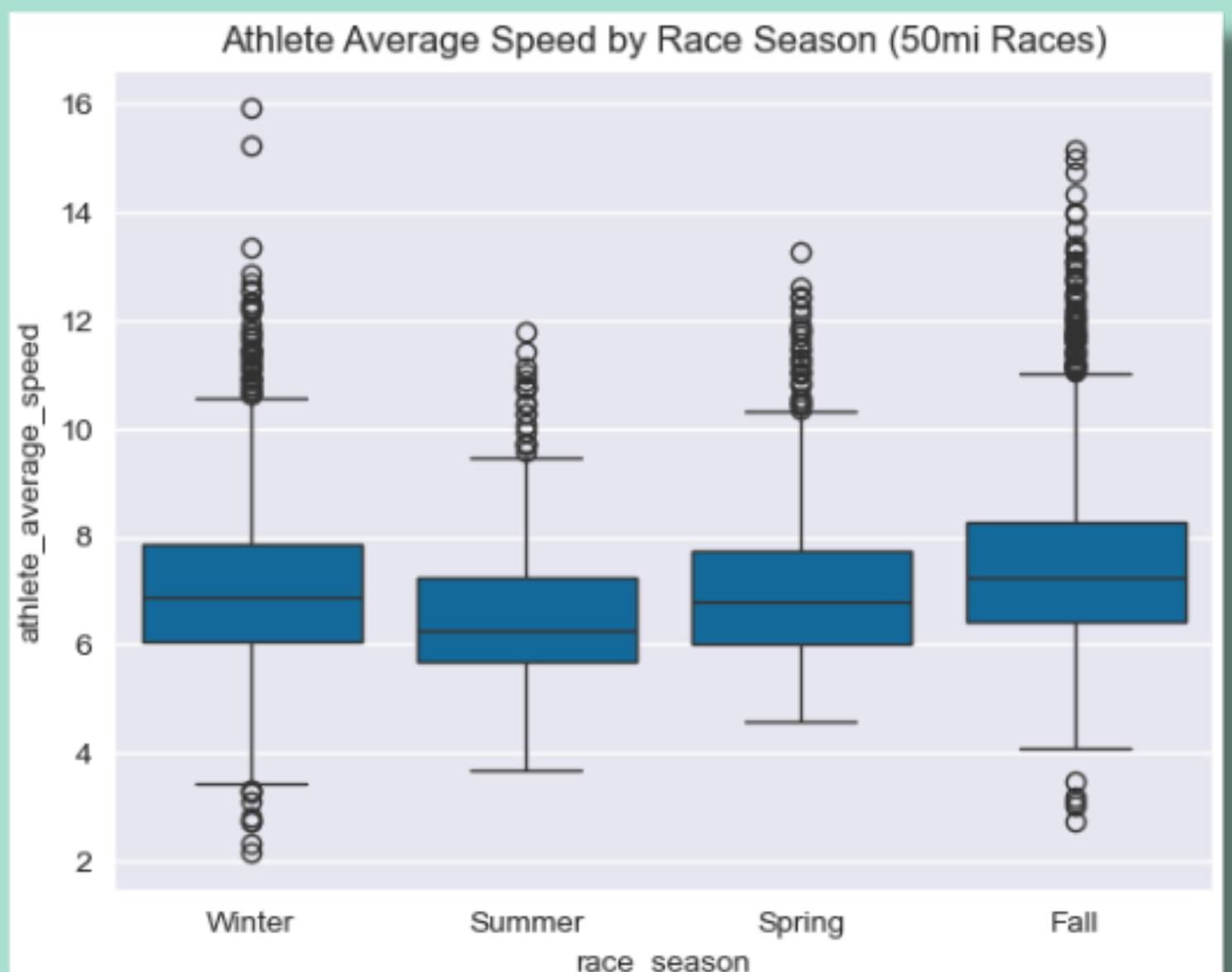
What is the best season to start running?

Looking at the average speed of runners during the 50 mile race the quickest runners are on average running in the winter

	mean	count
race_season		
Fall	7.511585	1997
Spring	7.082557	836
Winter	7.048442	1977
Summer	6.505776	817

The only issue with this data set is there are only half the amount of data entries in the spring and fall

Start in the winter!



Presented by Asher Hagan

Thank you very much!

