



PACING PREDICTOR
ONE MILE AT A TIME...MY RACE, MY PACE

RECAP...

- FIRST 100KM RACE IN 4 WEEKS
- TWICE THE DISTANCE I HAVE EVER RAN BEFORE
- CURRENT GUESS IS A 11-13 HOUR FINISHING TIME.
- CAN I BE MORE PRECISE?



PACING IN ULTRA-MARATHONS IS CRITICAL...



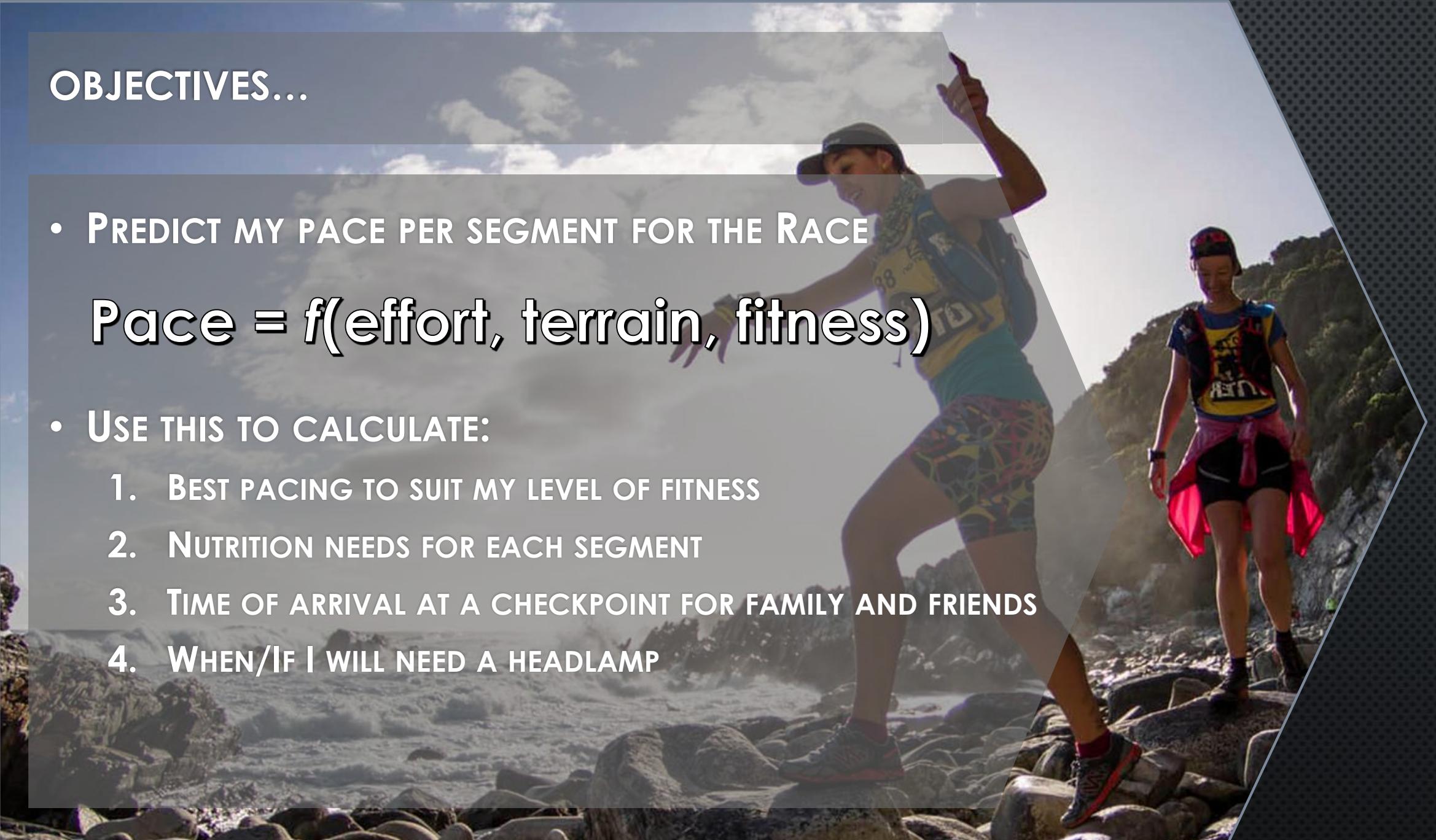
OBJECTIVES...

- PREDICT MY PACE PER SEGMENT FOR THE RACE

Pace = $f(\text{effort, terrain, fitness})$

- USE THIS TO CALCULATE:

1. BEST PACING TO SUIT MY LEVEL OF FITNESS
2. NUTRITION NEEDS FOR EACH SEGMENT
3. TIME OF ARRIVAL AT A CHECKPOINT FOR FAMILY AND FRIENDS
4. WHEN/IF I WILL NEED A HEADLAMP



TWO APPROACHES TO PREDICTING PACE...

Pace = $f(\text{effort, terrain, fitness})$

1. APPROACH 1: PREDICT PACE

- ESTIMATE MY EFFORT WHEN RUNNING 100KM
- UNDERSTAND HOW MY PACE VARIES OVER DIFFERENT TERRAIN
- ADJUST FOR EXPECTED FITNESS LEVEL GOING INTO THE RACE

2. APPROACH 2: PREDICT PLACE

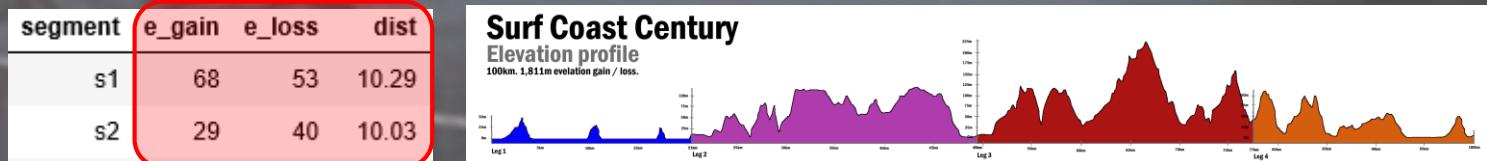
- PREDICT WHAT PLACE I WILL FINISH BASED ON MY PAST RACES
- USE PAST SCC RESULTS TO DETERMINE WHAT PACE PAST PARTICIPANTS IN THAT POSITION COMPLETED THE RACE IN

3 KEY SOURCES OF DATA...

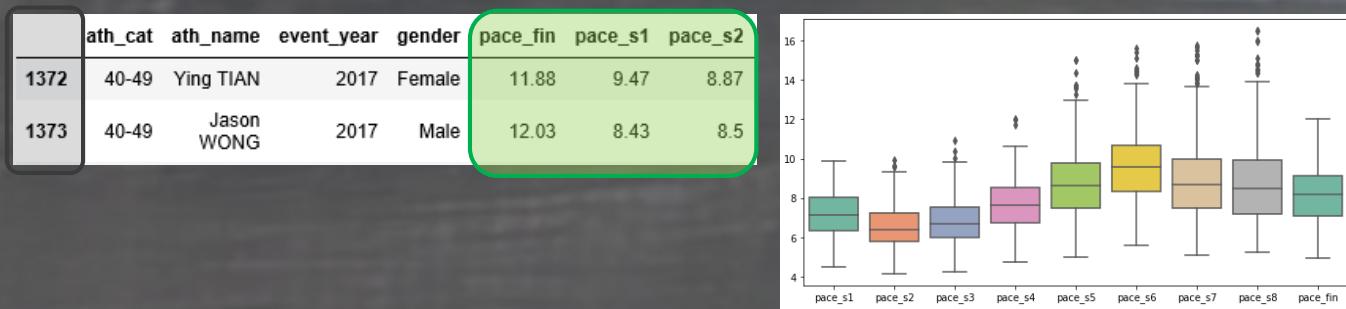
1. MY OWN RACE DATA

Y metric	Effort		Terrain		Fitness			Race Placing					
mins_km	if	hr	distance	gain_km	training_load (ctl)	recovery (tsb)	fitness (v02)	a_per	g_per	o_per	f_per	field_size	
4.82	0.95	169	11	15		68	10	52	0.19	0.21	0.11	0.03	large
4.82	0.88	163	15	11		63	11	57	0.20	0.21	0.13	0.03	medium

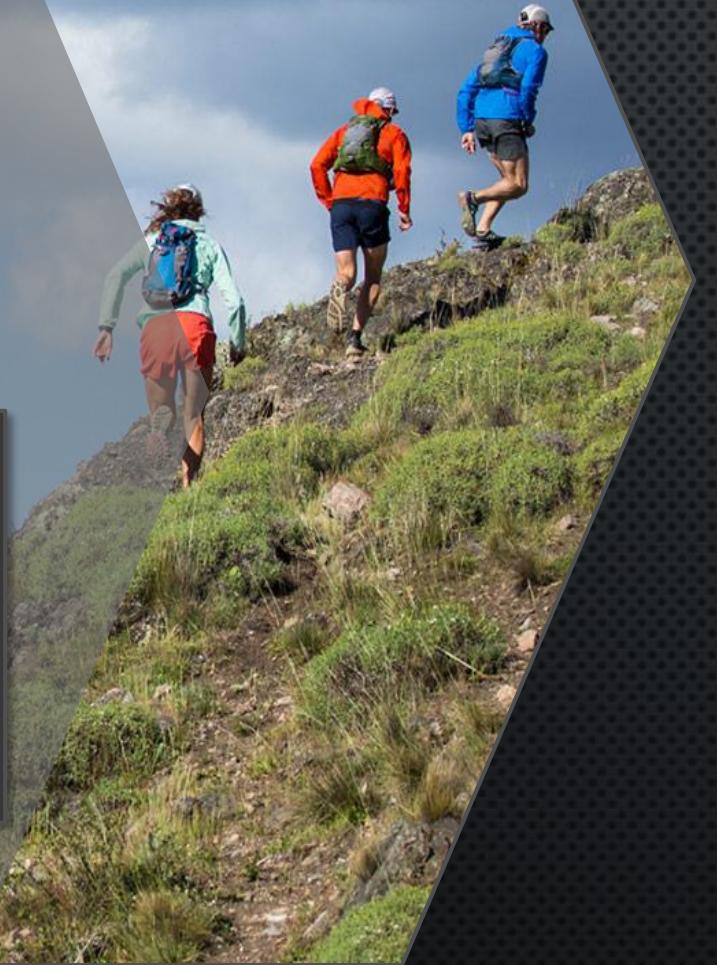
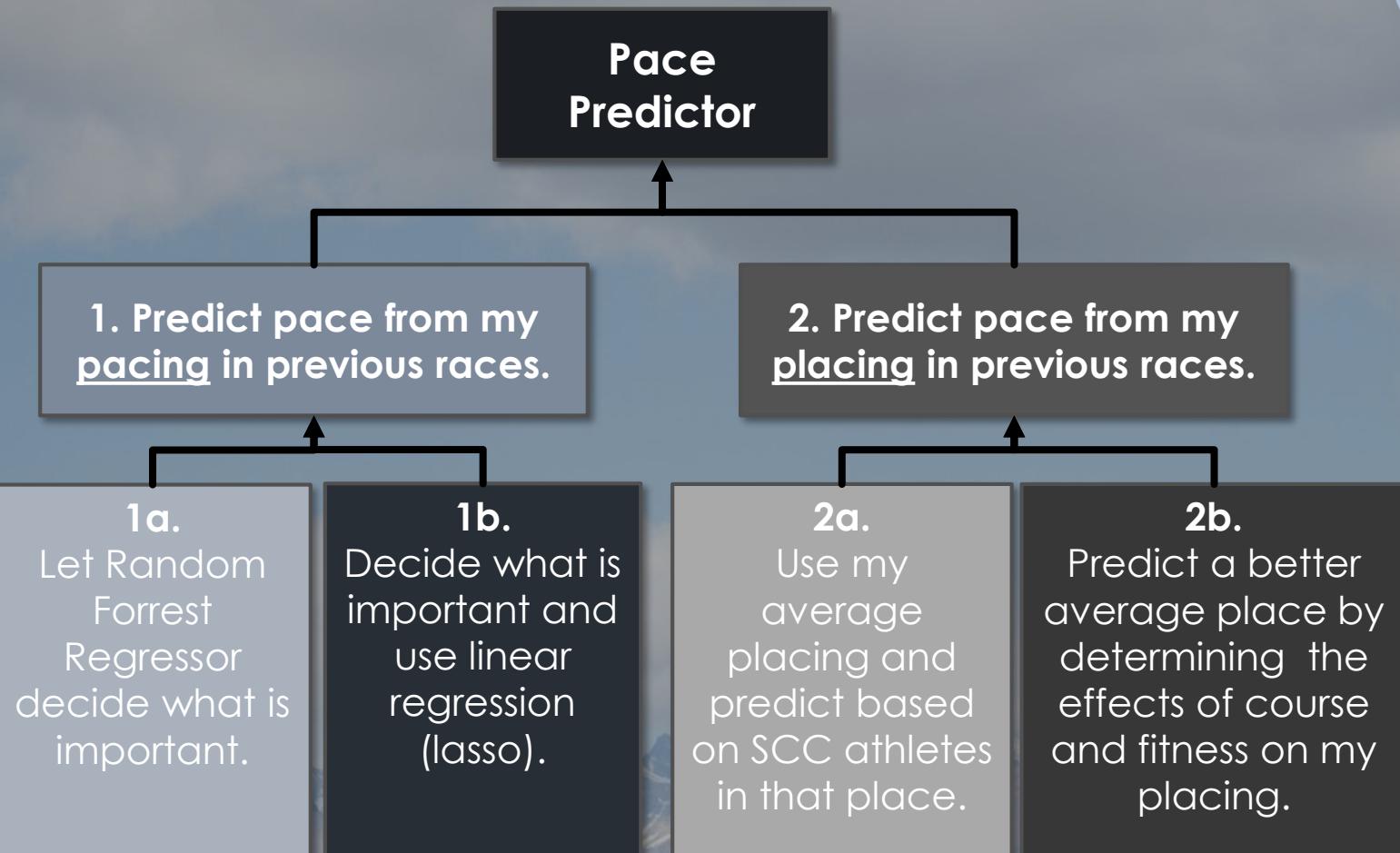
2. SCC COURSE DISTANCE AND ELEVATION



3. PAST SCC RESULTS FROM 2015-2017



MODELLING APPROACH TREE...



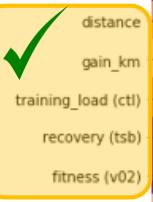
PREDICT PACE FROM MY PACING DATA...

- 1A. LET RANDOM FOREST REGRESSOR DECIDE WHAT IS IMPORTANT.

Features	Importance Score
gain_km	0.683306
distance	0.147672
recovery (tsb)	0.077060
fitness (v02)	0.047703
training_load (ctl)	0.044258

- 1B. DECIDE WHAT IS IMPORTANT AND USE LINEAR REGRESSION (LASSO).

- ID VARIABLES WITH A STRONG RELATIONSHIP TO MINS/KM
- ID WHICH VARIABLES THEY ARE ALSO RELATED TO AND CHOOSE 1.



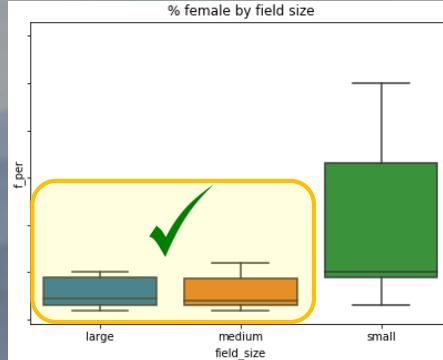
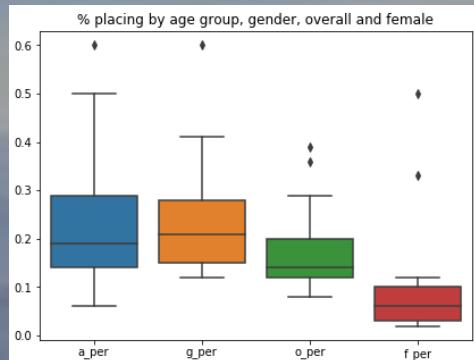
mins_km	1	-0.78	-0.79	0.56	0.72	0.44	-0.13	-0.11	0.72	0.72
if	-0.78	1	0.89	-0.79	-0.34	-0.17	0.13	-0.021	-0.33	-0.33
hr	-0.79	0.89	1	-0.83	-0.3	-0.35	0.097	-0.055	-0.31	-0.31
distance	0.56	-0.79	-0.83	1	0.035	0.36	0.17	0.09	0.027	0.033
gain_km	0.72	-0.34	-0.3	0.035	1	0.36	-0.19	-0.3	1	1
training_load (ctl)	0.44	-0.17	-0.35	0.36	0.36	1	-0.053	0.17	0.36	0.37
recovery (tsb)	-0.13	0.13	0.097	0.17	-0.19	-0.053	1	0.14	-0.21	-0.2
fitness (v02)	-0.11	-0.021	-0.055	0.09	-0.3	0.17	0.14	1	-0.31	-0.31
loss_km	0.72	-0.33	-0.31	0.027	1	0.36	-0.21	-0.31	1	1
change_km	0.72	-0.33	-0.31	0.033	1	0.37	-0.2	-0.31	1	1
mins_km		if	hr	distance	gain_km	training_load (ctl)	recovery (tsb)	fitness (v02)	loss_km	change_km



PREDICT PACE FROM MY PLACING DATA...

• 2A. USE MY AVERAGE PLACING

- % FEMALE PLACING IN MED/LARGE FIELDS
- FITNESS, EFFORT AND TERRAIN DATA TO EXPLAIN REMAINING VARIATION



• 2B. PREDICT A BETTER AVERAGE PLACE

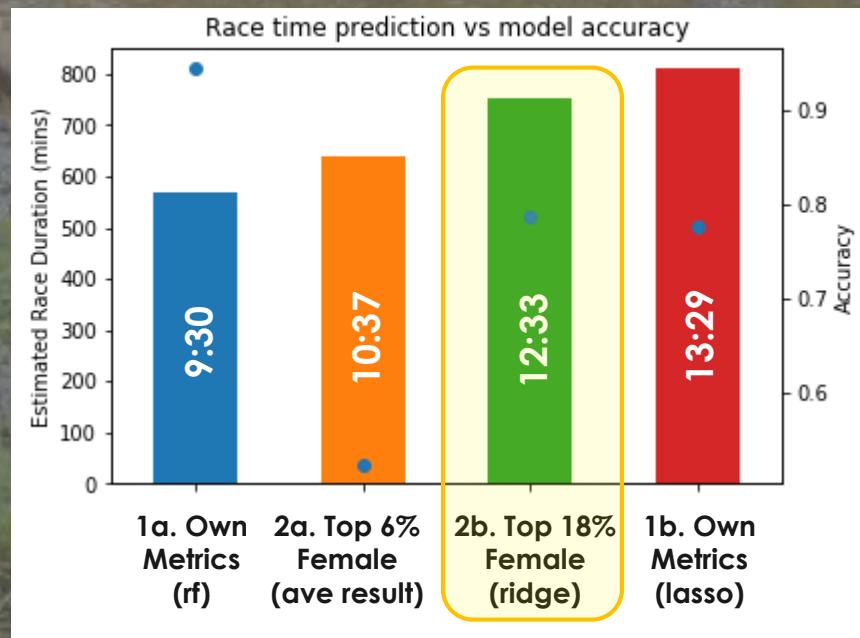
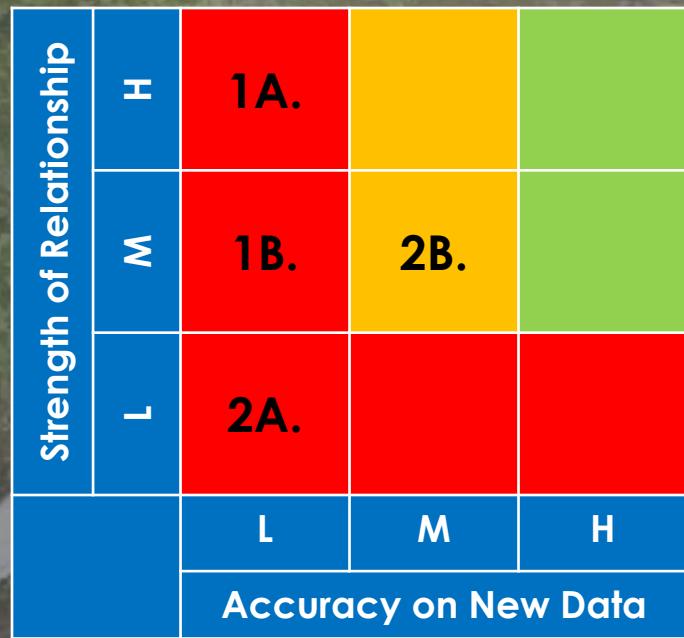
- DETERMINING THE EFFECTS OF COURSE AND FITNESS ON MY PLACING.

training_load (ctl)	-0.05	0.34	1	-0.0026	-0.054	-0.26	0.27	0.36	0.11
recovery (tsb)	-0.14	0.16	-0.0026	1	0.099	0.065	0.094	-0.044	0.29
if	-0.64	-0.77	-0.054	0.099	1	0.85	-0.19	-0.43	-0.31
hr	-0.75	-0.82	-0.26	0.065	0.85	1	-0.28	-0.43	-0.33
fitness (v02)	-0.21	0.23	0.27	0.094	-0.19	-0.28	1	-0.29	0.22
gain_km	0.34	0.24	0.36	-0.044	-0.43	-0.43	-0.29	1	0.14
o_tot	0.18	0.34	0.11	0.29	-0.31	-0.33	0.22	0.14	1
f_per									



RESULTS...

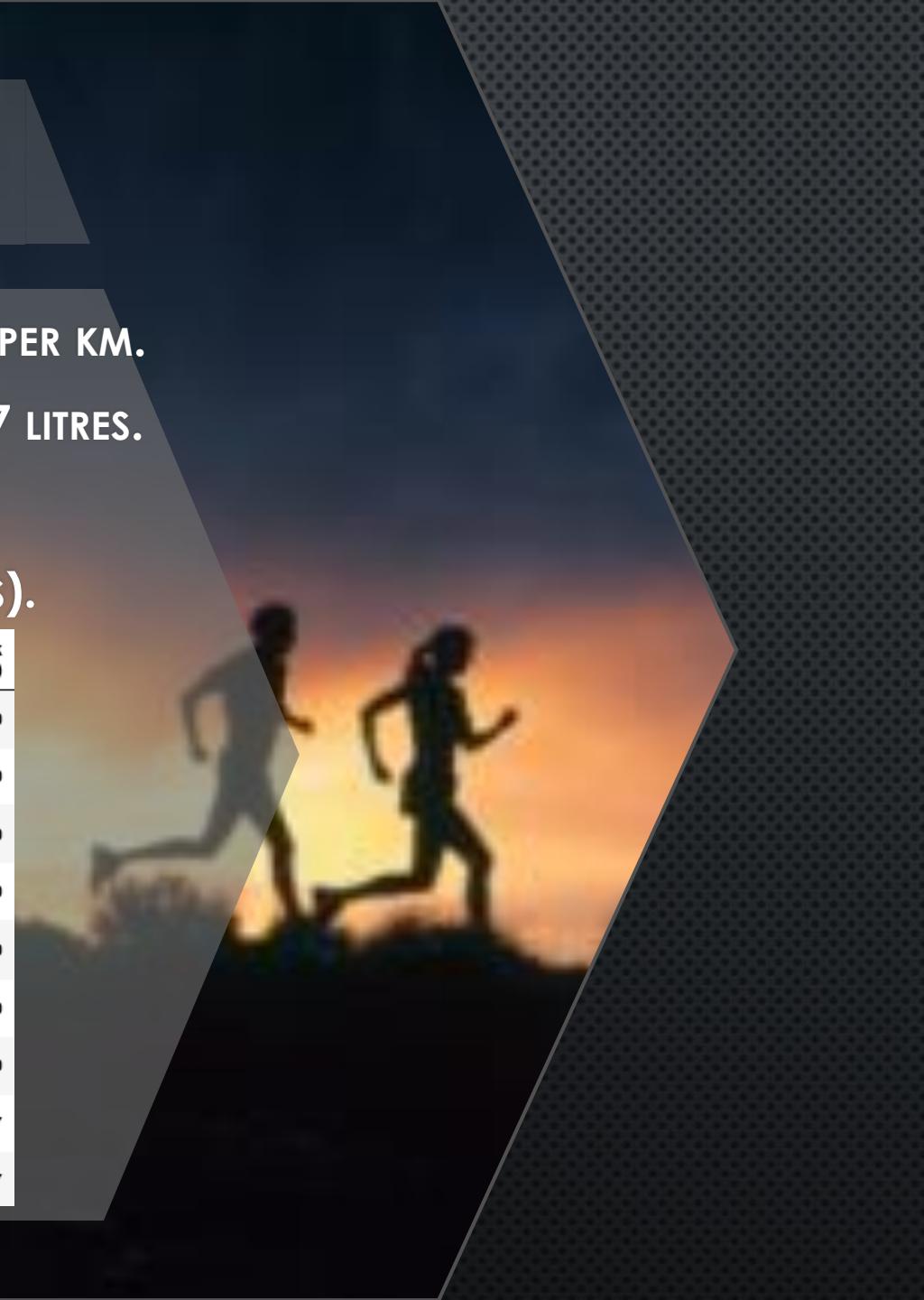
- MODEL SUCCESS METRICS
 - MAXIMISE R² (RELATIONSHIP) AND MINIMISE MSE (PREDICTION ERROR)
 - HAVE LOW VARIATION IN MSE DURING CROSS VALIDATION (ACCURACY ON NEW DATA)



THE RACE PLAN...

1. SCC WILL TAKE ME 12.5HRS, RANGING BETWEEN 6:30 – 8:45 PER KM.
2. OVER THE 100KM, I WILL CONSUME 13 GELS, 13 TABS AND 8.7 LITRES.
3. I WILL SEE MY FAMILY AT 1:15PM (50KM) AND 8PM (100KM)
4. I WILL NEED MY HEADLAMP FOR THE LAST SEGMENT (87 MINUTES).

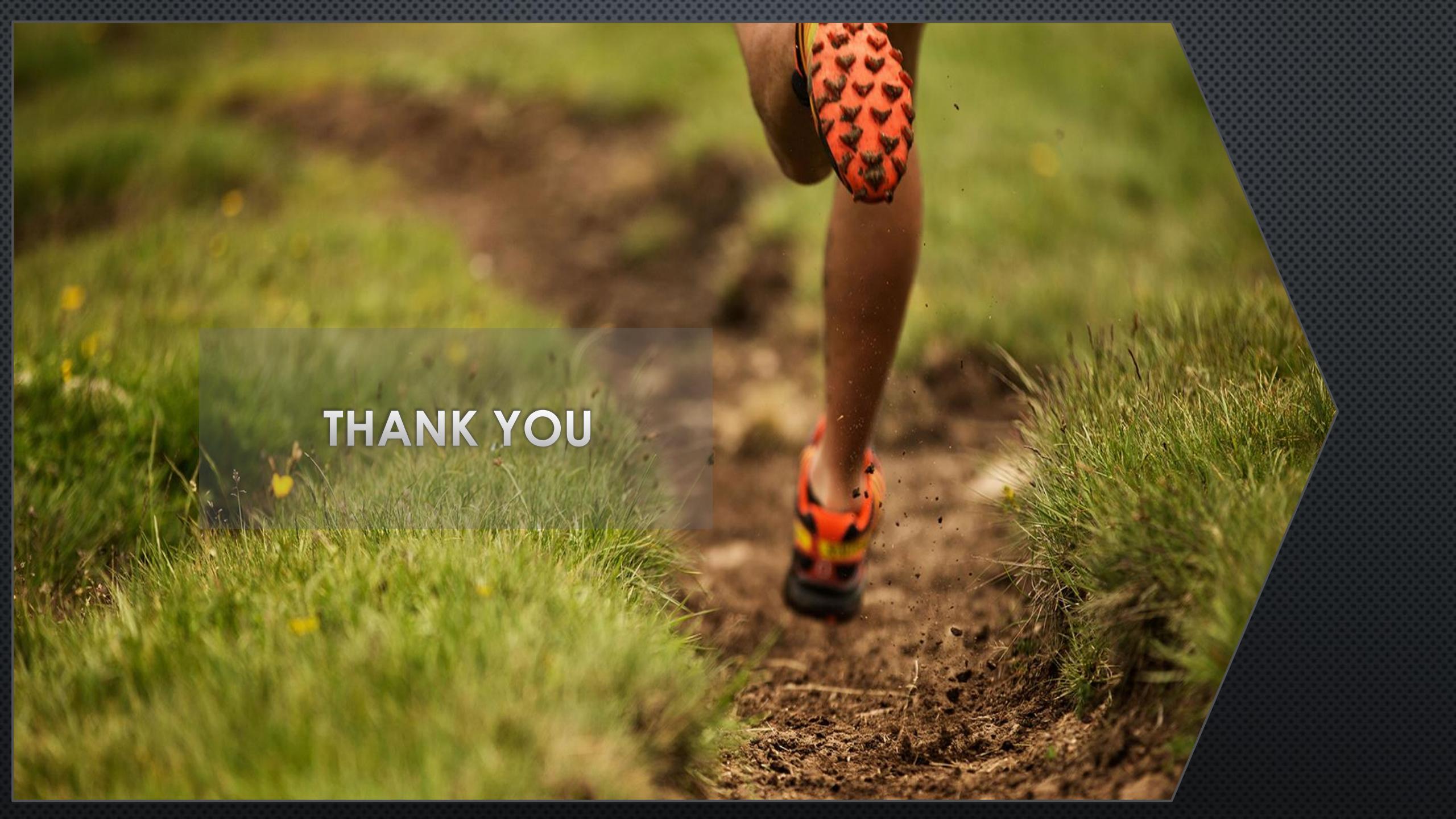
Segment	Distance	Pace (mins/km)	Duration (mins)	Race Time	Arrival Time	Gels Taken	Perpetuam (1 tab)	Fluids Taken (ml)	Dark (mins)
s1	10.29	7.496667	77.140700	1h 17m	1900-01-01 08:47:08	1.285678	1.285678	899.974833	0
s2	10.03	6.590000	66.097700	2h 23m	1900-01-01 09:53:14	1.101628	1.101628	771.139833	0
s3	10.56	6.610000	69.801600	3h 33m	1900-01-01 11:03:02	1.163360	1.163360	814.352000	0
s4	18.48	7.170000	132.501600	5h 45m	1900-01-01 13:15:32	2.208360	2.208360	1545.852000	0
s5	20.46	7.976667	163.202600	8h 28m	1900-01-01 15:58:44	2.720043	2.720043	1904.030333	0
s6	7.66	8.796667	67.382467	9h 36m	1900-01-01 17:06:07	1.123041	1.123041	786.128778	0
s7	8.39	8.186667	68.686133	10h 44m	1900-01-01 18:14:48	1.144769	1.144769	801.338222	0
s8	14.13	7.686667	108.612600	12h 33m	1900-01-01 20:03:25	1.810210	1.810210	1267.147000	87.4167
fin	100.00	7.534254	753.425400	12h 33m	1900-01-01 20:03:25	12.557090	12.557090	8789.963000	87.4167



NEXT STEPS...

- CONFIDENCE LEVELS: CREATE A RANGE OF PACES/TIMES
- INCREASE SAMPLE SIZE: INTRODUCE TRAINING DATA
- ENVIRONMENTAL DATA: WIND, ALTITUDE, TEMP AND HUMIDITY.
- COURSE ELEVATION: PER 10M RATHER THAN 10KM
- COURSE TERRAIN: SAND, STAIRS, RIVER, ROAD, TRAIL
- EQUIPMENT: TREKKING POLES, CHANGE OF SHOES, CLOTHES





THANK YOU