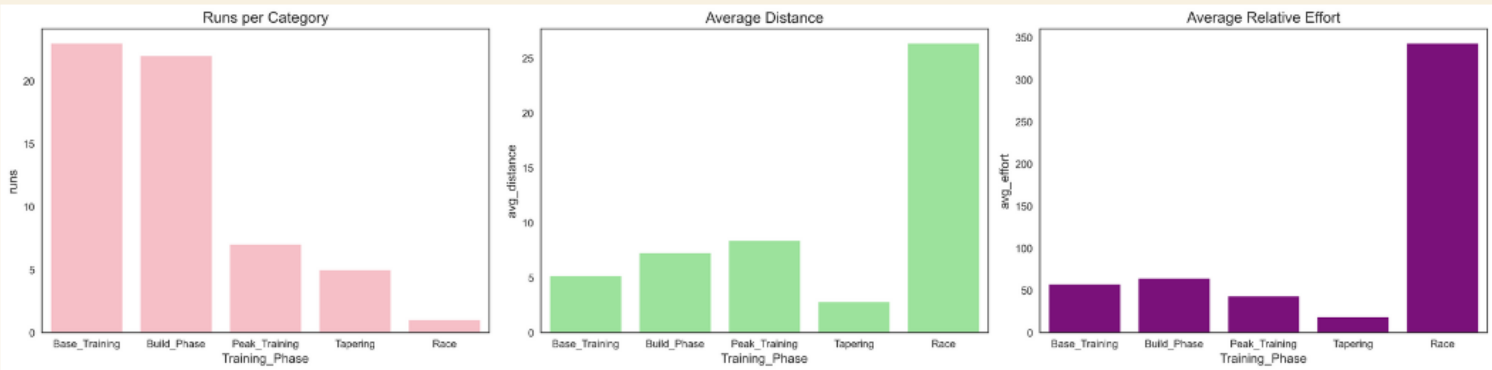
Running the Numbers: Turning Heartbeats into Insights

Part 1: What shaped my training?

A year ago, I completed my second marathon. As a data enthusiast, I thought that it would be interesting to analyze some of my data from my **16-week training cycle**. I analyzed general daily physio data from my Ringconn smart ring as well as running-specific data recorded on my Apple Watch Series 5 imported from Strava. Take a look!

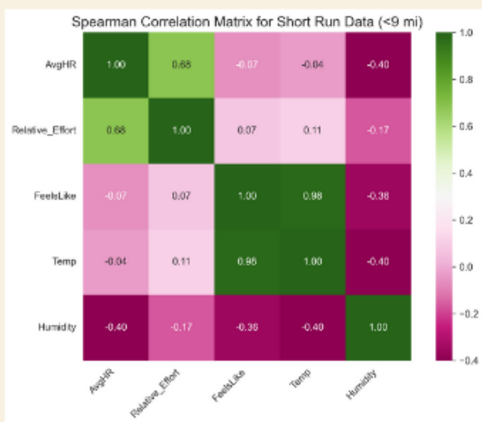
- This data ranged from May 2, 2025 to March 3, 2025.
- I completed **58 runs** totalling **~ 377 total miles**
- I broke down my training runs into 5 subcategories: base training (weeks 1-6) , build phase (weeks 7-12) peak training (weeks 13&14), tapering (weeks 15&16), and race day (last day of week 16)

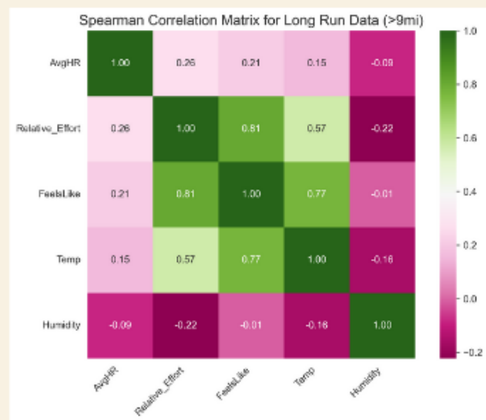


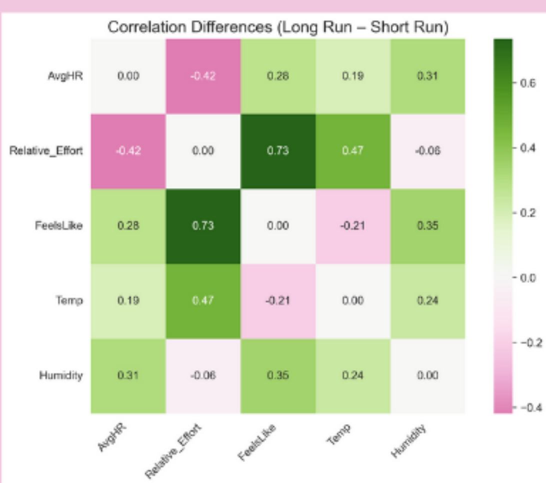
How did weather affect my training efforts?

My training plan ran from early May to early September, so it was hot and humid 🥵

The following plots show correlation matrices between weather factors (humidity, actual & RealFeel temperature) and performance strain (relative effort*, heart rate)







Long runs were significantly more impacted by heat and humidity than short runs

- The link between relative effort and apparent temperature was <u>much</u> stronger in long runs (+0.73 difference).
- The link between relative effort and temperature was stronger in long runs (+0.47 difference).
- The link between relative effort and humidity was stronger in long runs (+0.31 difference).

🌡️ External stressors like heat and humidity **compound over time**, placing greater physiological strain on long runs.

😮‍💨 Short runs are **more insulated from these effects** – likely because heat doesn't have as much time to tax the system

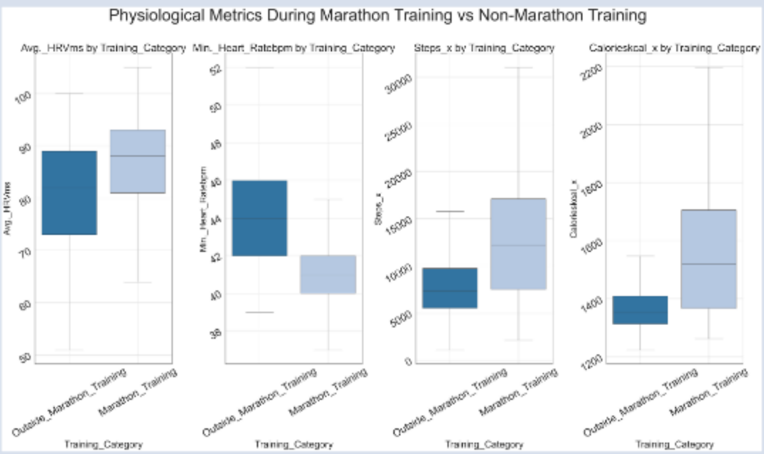
*Strava calculates Relative Effort primarily using heart rate data. It measures the time spent in various heart rate zones and assigns a score based on the intensity of the effort rather than just the total time spent exercising. A short and hard activity can require just as much effort as a long and leisurely one, and Relative Effort makes it so you can compare the two

Part 2: What physiological changes did I experience?

Throughout my training, I sensed internal changes in my cardiovascular system. Though I felt sore physically, I internally felt stronger and less out of breath while I was running. I decided to further investigate the physiological differences that I experienced during marathon training and within different marathon training phases. I focused on analyzing daily measurements of 4 main metrics: **Heart Rate Variability, Min. Heart Rate (Resting HR), Calories burned, and Steps**

Heart Rate Variability (HRV) is a measure of the fluctuations in the time between consecutive heartbeats. It is an indicator of recovery and a measure of your body's ability to adapt to training loads. Elevated HRV between runs suggests good recovery and low stress. A rising HRV trend during the taper phase (while training volume is reduced) is a positive sign of recovery.

Resting Heart Rate (RHR) is also a crucial indicator of fitness. Regular running often results in a lower resting heart rate, indicating improved cardiovascular efficiency.



Physiological Metrics During Different Phases of Marathon Training

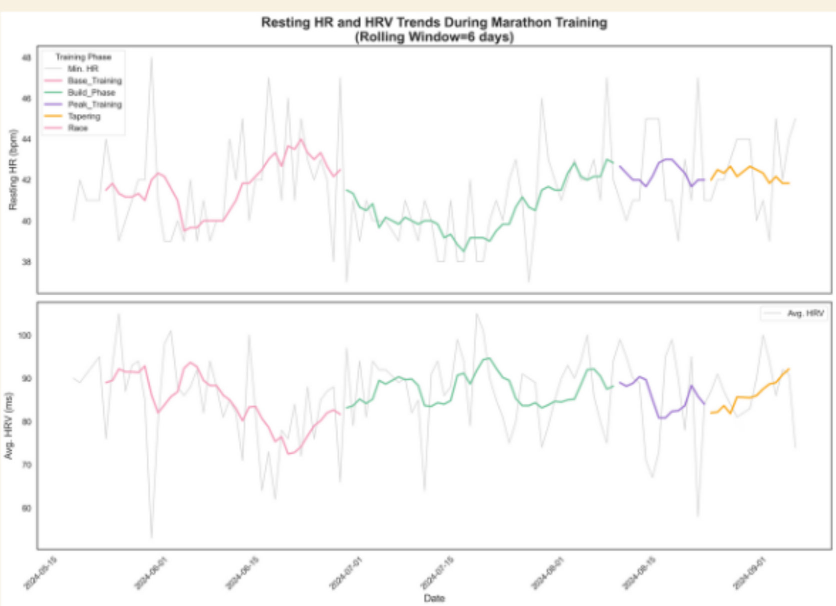
Daily Average HRV: On average, my **HRV was 7.6% higher** while marathon training, indicating improved recovery and resilience during training

Daily Resting HR: On average, my **RHR was 6.1% lower** while marathon training, indicating a more efficient cardiovascular system

Daily Steps: On average, I walked **65.1% more steps** while marathon training

Daily Calories Burned: On average, I burned **13.3% more daily calories** while marathon training

*All differences were highly statistically significant based on p-values (All ~0.0001)
*All effects were moderate or large based on Cohen's d values (0.54, -1.01,.88, and .99, respectively)



As expected, my RHR and HRV fluctuated with the intensity of the training cycle:

- Base phase: RHR dropped as my fitness improved and then slightly increased as training volume ramped up. HRV initially dropped as my body adapted to more cardio, then gradually rebounded as my aerobic fitness improved
- Build phase: RHR remained relatively stable but trended upwards slightly, reflecting accumulating training stress. HRV fluctuated, reflecting the added training stress and varying workout intensity of this period.
- Peak phase: RHR peaked as expected due to the high intensity and cumulative fatigue. HRV dipped noticably, as expected as the high training load placed greater strain on recovery.
- Taper phase: RHR stabilized and gradually declined as recovery took priority and training volume was reduced. HRV consistently rose as my body recovered in preparation for race day.

Part 3: What obstacles affected my training?

During this marathon training cycle, I was determined to maintain the 80/20 method, which involves running 80% of my runs at a very comfortable pace. My training plan recommended that I avoid focusing on pace at all when it comes to my long runs and instead focus solely on completing the distance at any pace comfortable to me.

Did I stick to my training goal of "undertraining" on my long runs?

To assess whether I successfully ran my long runs at a comfortable pace, I built a multiple linear regression model to predict relative effort on long runs, based on factors that influenced my effort during short runs (less than 9 miles). This approach used my short-run data—the majority of my training—to represent my typical running patterns outside of long runs. I then applied the model to my long runs: **if the actual relative effort was lower than predicted, it would indicate that I successfully ran those long runs at an easier pace, achieving my goal of undertraining.**

Objective: to use a predictive model to see if I successfully ran my long runs at an easier effort than expected ("undertrained") based on my short run patterns.

Model: Ridge regression with 5-fold cross-validation

Predictors: cumulative recovery (of the 3 days prior), time asleep the night before, distance of run, temperature difference (apparent temperature-actual), humidity, maximum speed, elevation gain, grade adjusted pace (pace normalized for hills).

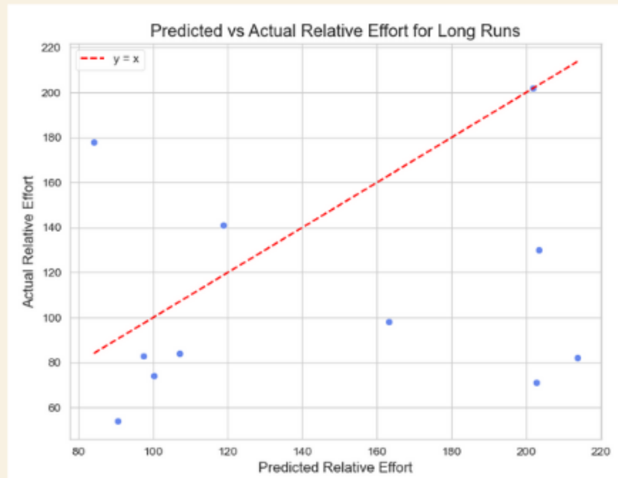
Results:
The model explains 81% of variance in Relative Effort on short runs ($R^2 = 0.807$) → strong predictive performance on short run dataset.

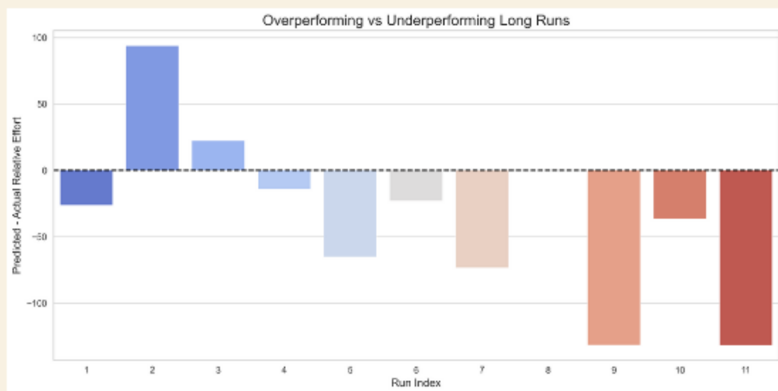
Applied to long runs:
If the predicted relative effort was greater than the actual, this would suggest that I successfully undertrained on my long runs.

- **9 of my 11 long runs were undertrained**
  - Overtrained runs: 2nd long run (significantly), 3rd long run (slightly)

Limitations

- Some predictors have high coefficients (>10) which suggests potential collinearity.
- Weather affects long runs differently than it does short runs. The model may underestimate the impact of weather on long runs.

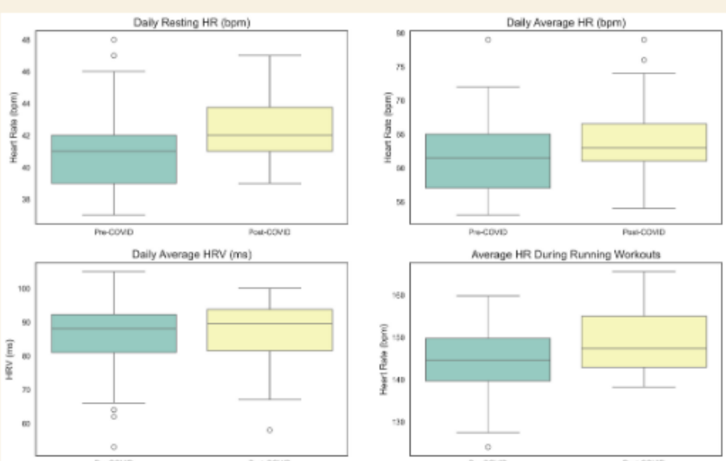




The COVID Saga

Unfortunately, 5 weeks before Race Day, I tested positive for COVID. After taking about a week off to recover, I could definitely feel my post-covid body struggling. I felt fatigued quickly and had a really hard time keeping my HR down on runs, even at a slow pace.

I was interested in comparing my pre and post-covid physio stats. looked at daily resting HR, average HR, and average HRV:



I also ran some t-tests to see if, at a statistically significant level, my physio stats changed after having COVID.

Daily Resting HR: My post-COVID average daily resting HR was **3.64% higher**. This was at a statistically significant level. ($p=0.007$).

Daily Average HR: My post-COVID daily average HR was **2.43% higher**.

Average HRV: My daily average HRV stayed about the same (0.77% higher).

Average Running HR: My average HR during runs was **3.08% higher** after recovering from COVID. ($p=0.077$).