Mission Analytics Data Exercise

Solution Report Authored by Kirtan Jani

Data sanitization overview

- Issues in female data (tab delimited error) fixed using RegEx in terminal

- Noted age issues in male and female data (for future use)

- Combined male and female data

- Created new columns (in seconds) to ease simplicity of calculations

Question 1 - Solution

Solutions given in HH:MM:SS format.	Males	Females
Mean	00:52:08	00:58:28
Median	00:51:23	00:57:50
Mode	00:35:28	00:48:06
Range	01:11:51	01:13:51

Brief Analysis:
- Males are faster than females
- Greater spread in female data
- Positive skewness (mean>mode) in both male & female net times
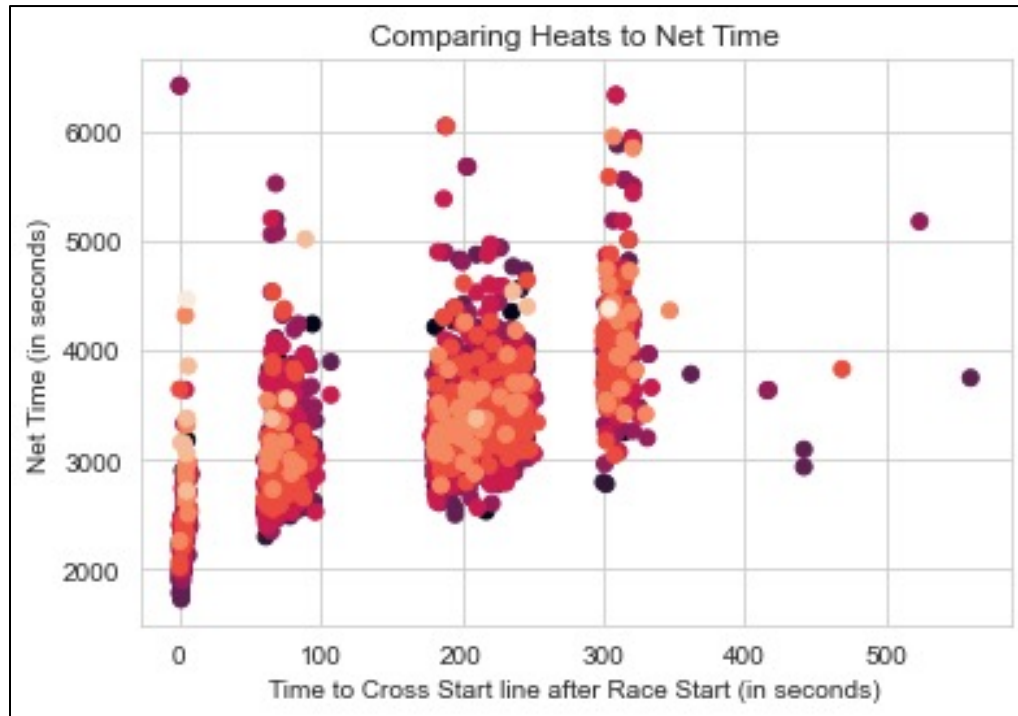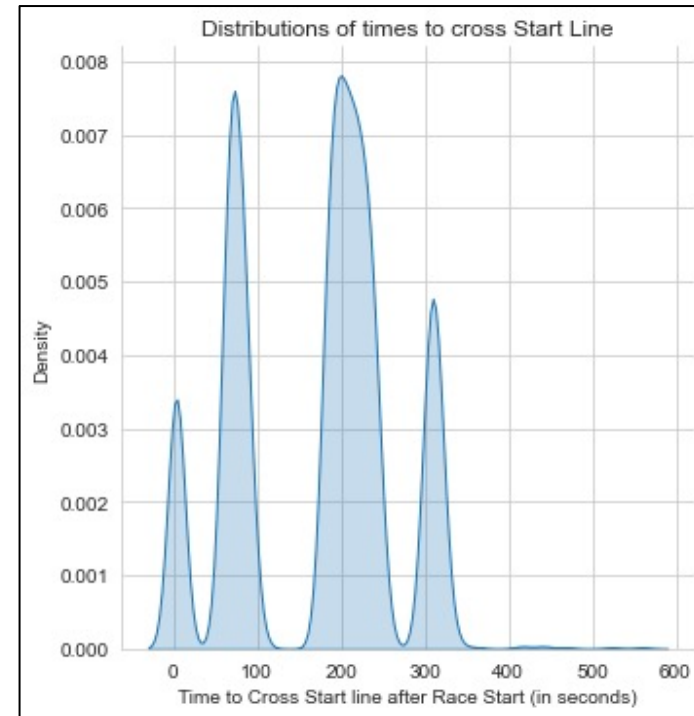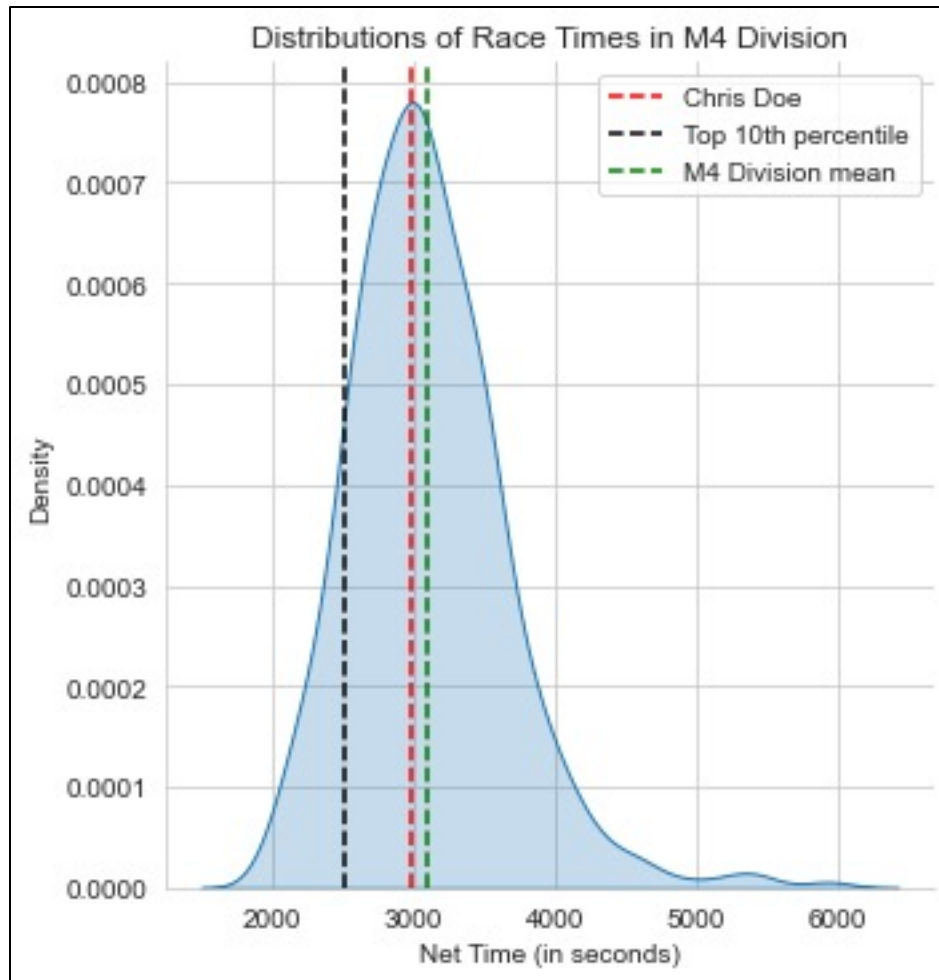
Question 2 - Solution

Figure 1

Figure 2

Brief Analysis:
- Runners start the race in 4 distinct groups* after the formal race start
- The earlier groups have a faster avg net time.
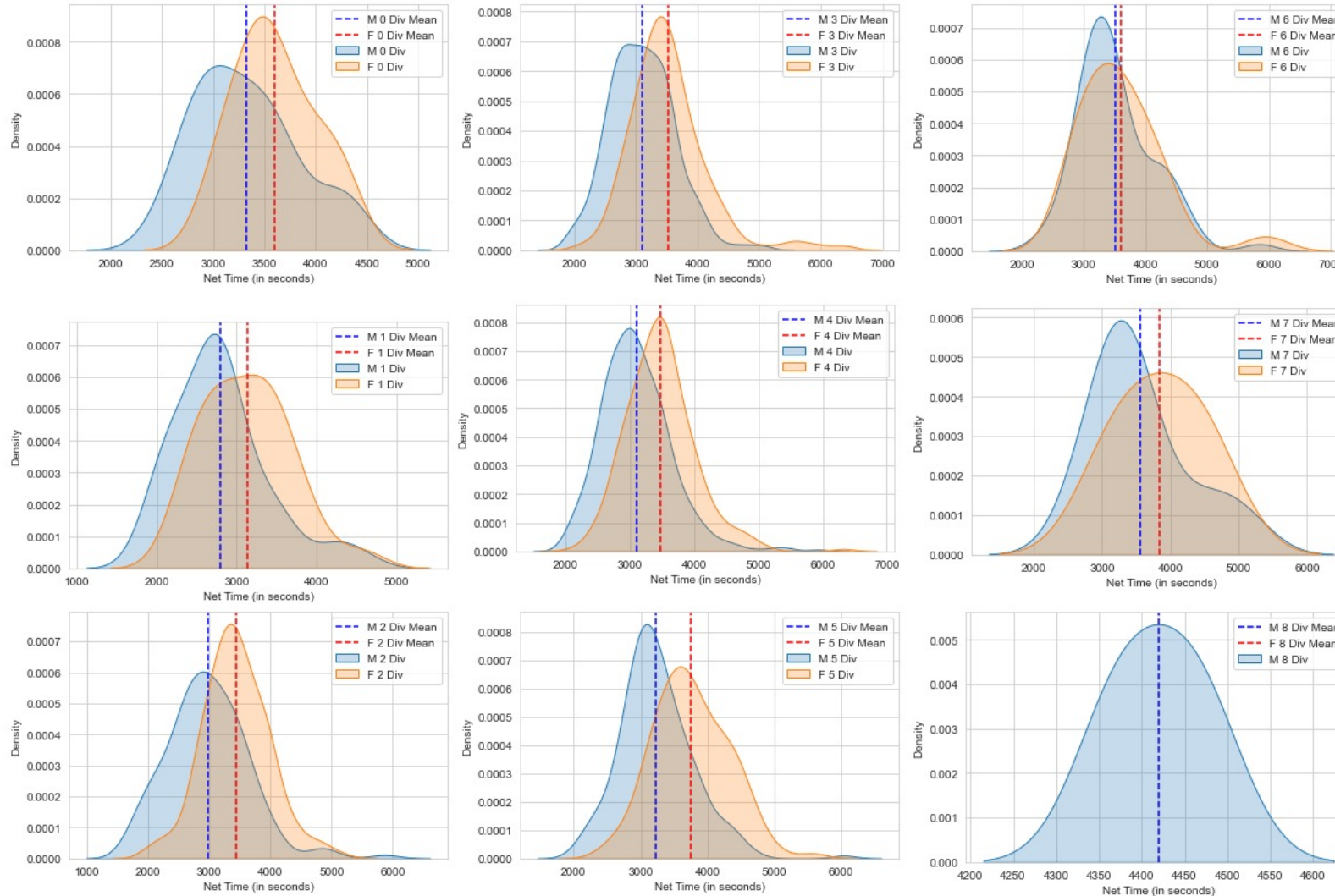
*(0 min, 1 min, 3 min, 5 min)

Question 3 - Solution

Brief Analysis:
- Chris Doe ran a 00:49:43
- The top 10th percentile min is 00:41:41

- To break into the top 10th percentile Chris Doe needs to run **00:08:02 FASTER**

Question 4 - Solution

Brief Analysis:
- Male divisions have faster net times
- Fastest avg divisions* are:
  - **M1** and F1 (ages 15-19)

* M1 - 00:46:39 | F1 - 00:52:08