MARATHON PROJECT BRIEF

Presented by Misonh Vladyslav

QUESTION 1: DATA IMPORT AND FORMATTING

1. IMPORT DATA FROM DATASET

2. FORMATTING DATA

Collect and Store Data

QUESTION 2: FASTEST/SLOWEST, MEDIAN, MEAN TIME

Question 2-a				
Name of marathons	Fastest Time	Slowest Time	Median Time	Mean Time
2016_half_marathons	1:11:15	4:54:41	2:26:18	2:31:57
2016_marathons	2:25:42	7:00:00	4:45:19	4:47:48
2017_half_marathons	1:10:58	6:18:24	2:39:34	2:43:59
2017_marathons	2:40:25	6:38:19	4:58:14	4:54:26
2018_half_marathons	1:09:25	4:48:06	2:25:32	2:31:46
2018_marathons	2:28:16	6:34:41	4:48:21	4:48:23
2019_half_marathons	1:10:03	5:03:35	2:26:07	2:32:07
2019_marathons	2:34:59	6:43:55	4:41:57	4:42:09

Question 2-b	
Median for 4 years marathon	Mean for 4 years marathon
4:48:53	4:48:27

Question 2-c	
Median for 4 years half-marathon	Mean for 4 years half-marathon
2:29:38	2:35:15

1. **Comparing Half Marathons and Marathons:**
   a. **The marathon time is about twice as long as the half marathon, which is expected since the marathon distance is twice as long**
2. **Stability between years:**
   a. **The relatively small fluctuations in the mean and median times between years indicate a stability in the level of participants and the conditions of the competition.**

QUESTION 3: OPRAH WINFREY'S TIME

Question 3-a		Question 3-b				Oprah Winfrey's time
Name of marathons	Number of runners who beat Oprah's time	Percentage of runners	Number of all runners			4:29:20
2016_marathons	1099	37.23%	2952			
2017_marathons	615	24.94%	2466			
2018_marathons	756	36.00%	2100			
2019_marathons	793	39.20%	2023			

1. **Difference between years:**
   a. **The number and percentage of runners who beat Oprah's time were the lowest in 2017. This could be due to external factors such as weather conditions or fewer experienced participants.**
2. **Overall trend:**
   a. **While the number of runners has been declining from 2016 to 2019, the percentage of runners who beat Oprah's time has been gradually increasing, indicating an improvement in the average level of participants or a change in the profile of participants.**

QUESTION 4: QUARTILE ANALYSIS

Question 4			
Name of marathons	First quartile	Second quartile	Third quartile
2016_half_marathons	2:08:17	2:26:18	2:50:11
2017_half_marathons	2:18:17	2:39:34	3:05:10
2018_half_marathons	2:07:20	2:25:32	2:51:16
2019_half_marathons	2:07:04	2:26:07	2:51:30

1. **First Quartile (Q1):**
   a. **The first quartile (Q1) value for each half marathon ranges from 2:07:04 (2019) to 2:18:17 (2017).**
   b. **The highest Q1 in 2017 indicates that the bottom 25% of participants had slightly slower times compared to other years. This may indicate that less fit participants competed in 2017.**
2. **Second Quartile (Q2 or Median):**
   a. **The median time ranges from 2:25:32 (2018) to 2:39:34 (2017).**
   b. **This data confirms that 2017 had the slowest average time, as in previous challenges.**
3. **Third Quartile (Q3):**
   a. **The Q3 value ranges from 2:50:11 (2016) to 3:05:10 (2017).**
   b. **The third quartile in 2017 is the highest, indicating that even the top 25% of participants ran slower than in other years.**

QUESTION 5: THE SLOWEST YEAR

Question 5	
The slowest year (half-marathon)	2017_half_marathons
The slowest year (marathon)	2017_marathons
Hypothesis	For me, this data set don't have enough info about any conditions that could affect the marathon. But I did a research and in period of marathon in 2017 was the hottest weather.

Name of marathons	Mean Time
2016_marathons	4:47:48
2017_marathons	4:54:26
2018_marathons	4:48:23
2019_marathons	4:42:09
2016_half_marathons	2:31:57
2017_half_marathons	2:43:59
2018_half_marathons	2:31:46
2019_half_marathons	2:32:07

1. **Finding the slowest year:**
   a. **For half marathons, the slowest year is 2017, with an mean time of 2:43:59.**
   b. **For marathons, the slowest year is also 2017, with an mean time of 4:54:26.**
   These results are consistent with previous analysis, where 2017 performed worse across all metrics.

Reasons for the slow results in 2017 may include the following factors:
1. **Weather conditions:**
   a. **Temperatures in 2017 ranged from 73°F to 91°F, making it the hottest year on record.**
   b. **High temperatures negatively impact a runner's endurance and performance, especially over long distances.**

Year	Date	Day of the Week	Temperature during the day (°F)	Humidity (%)	Pressure ("Hg)	Weather type
2016	4/30/2016	Saturday	64-75	73-88	29.87-29.92	Cloudy/Rain
2017	4/29/2017	Saturday	73-91	44-76	29.93-29.99	Cloudy
2018	4/28/2018	Saturday	46-72	32-85	29.9-30.13	Sunny
2019	4/27/2019	Saturday	45-75	28-79	29.79-30.03	Cloudy
	Source:		https://www.timeanddate.com/weather/usa/nashville/historic?month=4&year=2016			

QUESTION 6: SCOTT WIETECHA

Name of runner		Error message	
Scott Wietecha		Did not participate	
	Question 6		
Name of marathons	Wietecha's time	Next runner time	Time difference
2016_half_marathons	Did not participate	Did not participate	Did not participate
2016_marathons	2:25:42	2:34:43	0:09:01
2017_half_marathons	Did not participate	Did not participate	Did not participate
2017_marathons	2:40:25	2:56:28	0:16:03
2018_half_marathons	Did not participate	Did not participate	Did not participate
2018_marathons	2:28:16	2:29:52	0:01:36
2019_half_marathons	Did not participate	Did not participate	Did not participate
2019_marathons	2:34:59	2:35:24	0:00:25

Same data but in hours for better chart creating			
Name of marathons	Wietecha's time	Next runner time	Time difference
2016_half_marathons	0.00	0.00	0.00
2016_marathons	2.43	2.58	0.15
2017_half_marathons	0.00	0.00	0.00
2017_marathons	2.67	2.94	0.27
2018_half_marathons	0.00	0.00	0.00
2018_marathons	2.47	2.50	0.03
2019_half_marathons	0.00	0.00	0.00
2019_marathons	2.58	2.59	0.01
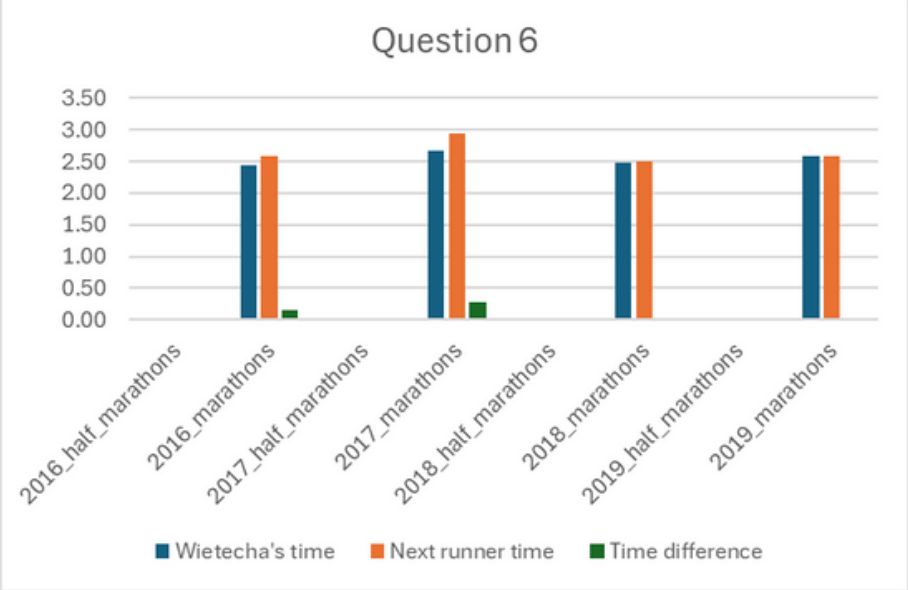
Marathon results:
- **2016:** The difference between Scott Wieteka's time (2:25:42) and the next runner (2:34:43) was 9 minutes and 1 second.
- **2017:** The difference increased to 16 minutes and 3 seconds (2:40:25 vs. 2:56:28), indicating a significant advantage.
- **2018:** The difference decreased to 1 minute and 36 seconds (2:28:16 vs. 2:29:52), indicating a more fierce competition.
- **2019:** The difference was the smallest in all years — just 25 seconds (2:34:59 vs. 2:35:24).

QUESTION BONUS: TOP-3 RUNNERS

BONUS Question							
	2016_marathons		2017_marathons		2018_marathons		2019_marathons	
	Runner name	Runner time	Runner name	Runner time	Runner name	Runner time	Runner name	Runner time
First place	Scott Wietecha	2:25:42	Scott Wietecha	2:40:25	Scott Wietecha	2:28:16	Scott Wietecha	2:34:59
Second place	Brian Shelton	2:34:43	Ryan Regnier	2:56:28	Garang Madut	2:29:52	Jordan Wilson	2:35:24
Third place	Christopher Capps	2:38:43	Daniel Everett	3:00:55	Kevin Fink	2:38:59	Steelton Flynn	2:39:59

Unique names by top 3 runners of each year	How many marathons finished by all years
Scott Wietecha	4
Brian Shelton	1
Christopher Capps	1
Ryan Regnier	1
Daniel Everett	1
Garang Madut	1
Kevin Fink	1
Jordan Wilson	1
Steelton Flynn	1

Unique runners	9

1. **Unique Runners:**
   a. **There were 9 unique runners in the top 3 list over all 4 years.**
   b. **Of these, only one runner, Scott Wietecha, has been in the top 3 all four years in a row, demonstrating his consistency and skill.**
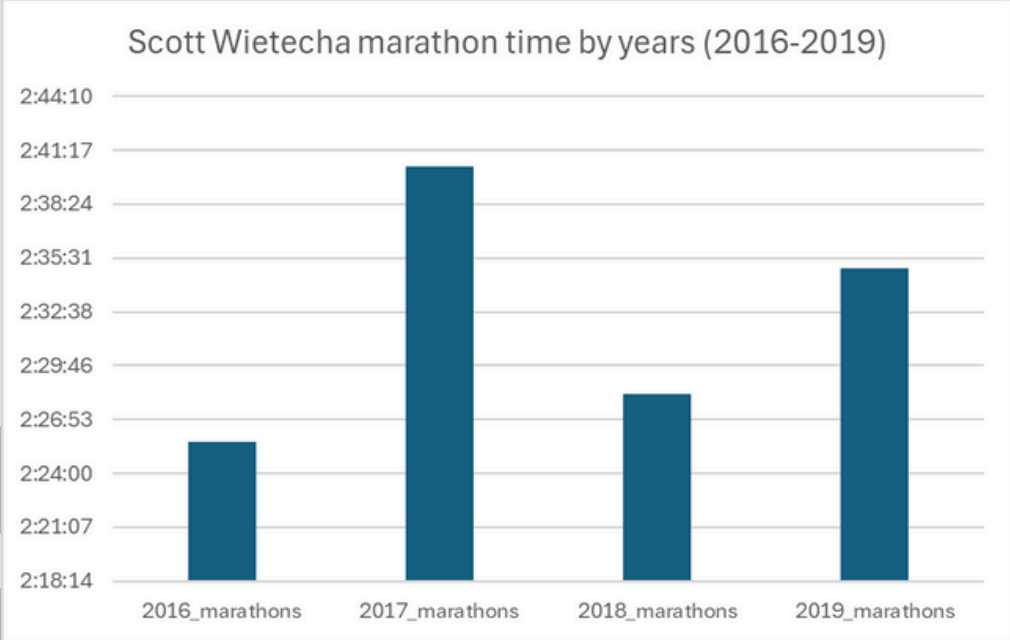2. **Runner distribution:**
   a. **The remaining 8 participants appear on the list only once (I also checked that over the years they not only did not make it into the top 3, but also did not participate at all).**
3. **Runners' time changes:**
   a. **Scott Wietecha: His time has changed every year:**
      i. **2016: 2:25:42**
      ii. **2017: 2:40:25 (+14:43 compared to 2016)**
      iii. **2018: 2:28:16 (-12:09 compared to 2017)**
      iv. **2019: 2:34:59 (+6:43 compared to 2018)**
      v. **There is a trend towards increasing times, which may be due to age changes or the conditions of the marathon.**
4. **Hypotheses:**
   a. **Scott Wietecha shows a consistent advantage over other runners, but his results are gradually becoming slower. This may be a result of a natural decline in physical performance or increased competition.**

Scott Wietecha marathon time by years (2016-2019)

Thank You

For your attention