Exploring the Relationship between Environmental Conditions and Marathon Performances

Abstract

Past research revealed significant differences in marathon performance influenced by gender, age. Meanwhile, age may influence thermoregulation as well. We therefore designated to explore the impact of environmental conditions on marathon performances, including weather and air quality. Exploratory data analysis and polynomial regression were employed to verify the influence of weather conditions and air quality on marathon performances. We also identified weather parameters that have the largest impact.

Introduction

The factors influencing marathon performance have always been a topic of interest, and existing research has confirmed the impact of weather conditions on performance. For instance, increasing WBGT has been shown to slow marathon performance (M. R. Ely et al. 2007; B. R. Ely et al. 2010), and the effects of weather are less evident in female runners. Additionally, marathon performance exhibits differences related to age and gender: men are more likely than women to experience a slowdown during a marathon, and older runners tend to have less pace variance compared to younger runners. Moreover, masters men and women are negatively impacted by higher temperatures and humidity during the New York City marathon. Endurance performance degrades with increasing environmental temperature, and this decline in performance associated with warmer temperatures is magnified in longer-distance events such as the marathon (42.2 km). Older adults also face thermoregulatory challenges that impair their ability to dissipate heat, further exacerbating the impact of elevated temperatures. Lastly, there are well-documented sex differences in endurance performance [Besson2022] and in the physiological processes related to thermoregulation. Additionally, research has indicated the impact of age on thermoregulation (Yanovich, Ketko, and Charkoudian 2020), highlighting a reduced ability to tolerate heat stress in older individuals (Kenney and Munce 2003). However, the relationship between marathon performance and environmental conditions, particularly air quality, as well as the differences based on age and gender, remains to be explored.

Therefore, this study aims to confirm the relationships between weather conditions, air quality, and marathon performance, while further identifying the weather parameters that have the greatest impact on performance. We hypothesize that the slowing of performance would be more pronounced with high WBGT or AQI in older individuals compared to younger ones, and similar trends would be observed in both men and women. To achieve our goals, we will employ exploratory data analysis and polynomial regression.

Data Collection

The current study uses the data collected from marathon race results for Boston, Chicago, New York, Twin Cities, Grandma's Marathons for 17 to 24 years, 1993-2016. To evaluate marathon performance, we extracted fastest finishing time among men and women at each year of age, compared with the course record (%CR). The data also attaches parameters that recorded local weather conditions, including Td(dry bulb temperature in Celsius), Tw(wet bulb temperature in Celsius), %rh(percent relative humidity) and so on. Wet bulb globe temperature (WGBT) is calculated by weighting these weather parameters and is regarded as a general measurement that integrate all of weather information. Additionally, another data set records local air quality by AQI of each race, which is regarded as a reflection of environmental conditions as well.

Data Processing

The raw marathon data included 11,564 cases with 8 variables that recorded weather conditions, 2 demographic variables and a variable that records marathon performances, while the air quality data mainly contains AQI(Air Quality Index). According to the city and year of each race, we merged two data sets to incorporate both weather information and air quality. Before the analysis, a summary table of each variable was constructed to obtain a rough picture of our combined data set. The missingness of weather conditions and AQI is explicitly inspected. As Table 1 suggests, most of the variables didn't show significant differences between genders, while women tend to achieve better marathon performances and have younger age than men do. It is believed that the data was collected from equal samples, and avoided the possibility of confoundness between variables. In addition, we observed large missingness in weather parameters and AQI, whose number of missing cases reaches to approximately 500 and 4000, respectively.

Table 1: Summary table of marathon data

Characteristic	**Overall** N = 11,564[1]	**Female** N = 5,452[2]	**Male** N = 6,112[2]	**p-value**[3]
race				0.8
Boston	2,088/11564 (18%)	984/5452 (18%)	1,104/6112 (18%)	
Chicago	2,553/11564 (22%)	1,210/5452 (22%)	1,343/6112 (22%)	
Duluth	2,000/11564 (17%)	934/5452 (17%)	1,066/6112 (17%)	
NYC	2,930/11564 (25%)	1,402/5452 (26%)	1,528/6112 (25%)	
TC	1,993/11564 (17%)	922/5452 (17%)	1,071/6112 (18%)	
year	[1,993, 2,016]	[1,993, 2,016]	[1,993, 2,016]	0.7
flag				>0.9
1	3,753/11073 (34%)	1,769/5218 (34%)	1,984/5855 (34%)	
2	4,706/11073 (42%)	2,222/5218 (43%)	2,484/5855 (42%)	
3	2,022/11073 (18%)	948/5218 (18%)	1,074/5855 (18%)	
4	592/11073 (5.3%)	279/5218 (5.3%)	313/5855 (5.3%)	
Unknown	491	234	257	
age	[14, 91]	[15, 88]	[14, 91]	<0.001
CR	48.84 (44.89)	48.74 (41.53)	48.93 (47.70)	<0.001
Td	13.32 (5.81)	13.31 (5.80)	13.32 (5.81)	>0.9
Unknown	491	234	257	
Tw	9.42 (5.37)	9.41 (5.37)	9.43 (5.38)	0.9
Unknown	491	234	257	
rh	42.32 (32.01)	42.47 (31.95)	42.19 (32.06)	0.7
Unknown	491	234	257	
Tg	24.95 (7.71)	24.93 (7.71)	24.96 (7.72)	0.9
Unknown	491	234	257	
SR	513.82 (187.57)	512.61 (187.33)	514.91 (187.80)	0.5
Unknown	491	234	257	
DP	5.46 (6.95)	5.45 (6.94)	5.47 (6.95)	0.9
Unknown	491	234	257	
wind	9.94 (4.07)	9.94 (4.07)	9.95 (4.07)	>0.9
Unknown	491	234	257	
WBGT	12.91 (5.59)	12.91 (5.59)	12.92 (5.59)	0.9
Unknown	491	234	257	
aqi	37.30 (14.39)	37.32 (14.39)	37.29 (14.39)	0.9
Unknown	3,993	1,856	2,137	
arithmetic_mean	3.5 (1.9, 5.8)	3.5 (1.9, 5.8)	3.5 (1.9, 5.8)	0.5
Unknown	3,993	1,856	2,137	

[1]n/N (%); [Min, Max]; Mean (SD); Median (Q1, Q3)
[2]Mean (SD), Min - Max are reported for continuous variables. Count (%) for categorical variables.
[3]Pearson's Chi-squared test; Wilcoxon rank sum test

To further investigate the missing patterns of weather parameters and AQI, we counted the number of missing cases across races. We first focus on the missingness of weather parameters. It has been suggested that during 2011-2012, races held in Chicago, New York, Twin Cities, and Duluth experienced significant data missing issues (Table 2). Notably, both Twin Cities and Duluth lacked air quality data. Given the geographical proximity of these two cities (both located in Minnesota), it is plausible that the regional climate center encountered major disruptions, possibly due to disasters or other issues, which impacted data recording and reporting.

Table 2: Summary table of cases that missing WBGT

Year	Race	Number of cases that missing weather conditions	AQI	AQI_ext
2011	Chicago	126	65.46154	11.292162
2011	NYC	131	36.37079	3.366759
2011	TC	118	NA	NA
2012	Duluth	116	NA	NA

Furthermore, we paid attention to the missing air quality data. As shown by Figure 1, Duluth and Twin Cities ceased reporting air quality data after 2000, while other cities reported no such issues. Combined with previous evidence, it can be inferred that Minnesota may have discontinued reporting these statistics in favor of other air quality measurements, or the data may have experienced systematic missingness during the collection process. In conclusion, the missingness of weather parameters and AQI is likely Missing Not at Random (MNAR) and should be carefully considered in the analysis.

Results

Marathon Performance across Age

Our first aim is to examine the effects of increasing age on marathon performance in men and women. As shown in the Figure 2, marathon performance generally exhibits an asymmetric concave quadratic relationship with age. From around 14 to 25 years old, individuals' marathon times continuously decrease, reaching a minimum at the lowest point, after which the times

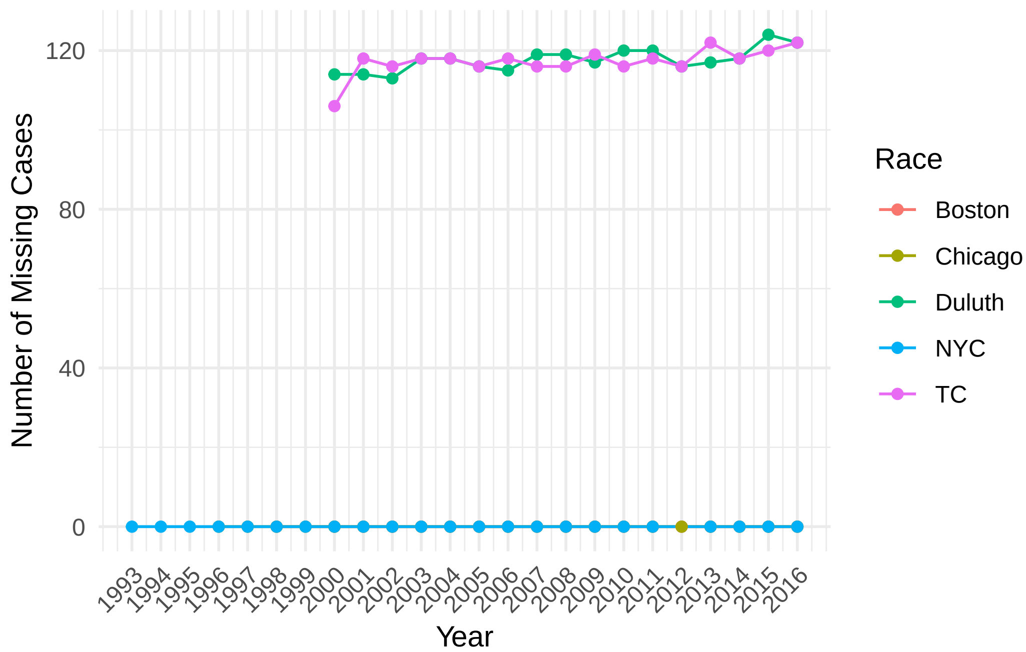

Figure 1: Number of cases that missing AQI

steadily increase, eventually reaching up to three times the best record. At the same time, men and women show different declining trends: overall, men perform better in marathons than women, and the age distribution of marathon participants is broader for men, but their performance also declines more significantly with increasing age.

Marathon Performances across Weather Conditions

Next, we aim to examine the effects of weather conditions on marathon performance and whether these effects differ between genders. We first transformed the continuous WBGT variable into a flag variable with five levels to reflect the risk of heat illness. WBGT values of <10°C, 10-18°C, 18-23°C, 23-28°C, and >28°C correspond to the flags white, green, yellow, red, and black, respectively. As shown in Figure 3, marathon performance does not appear to differ across weather conditions, and the marathon times for men and women under the same weather conditions are also very similar. However, it is worth noting that there are many outliers in all groups, which may need to be removed in future analyses to avoid the influence of extreme values.

Moreover, it is also worth investigating whether the effect of weather conditions on marathon performance varies with age. To this end, we plotted the relationship between marathon times and age under different weather conditions. As shown in Figure 3, as the risk of heat illness increases, marathon times tend to rise overall, and when the flag is red, the age range of

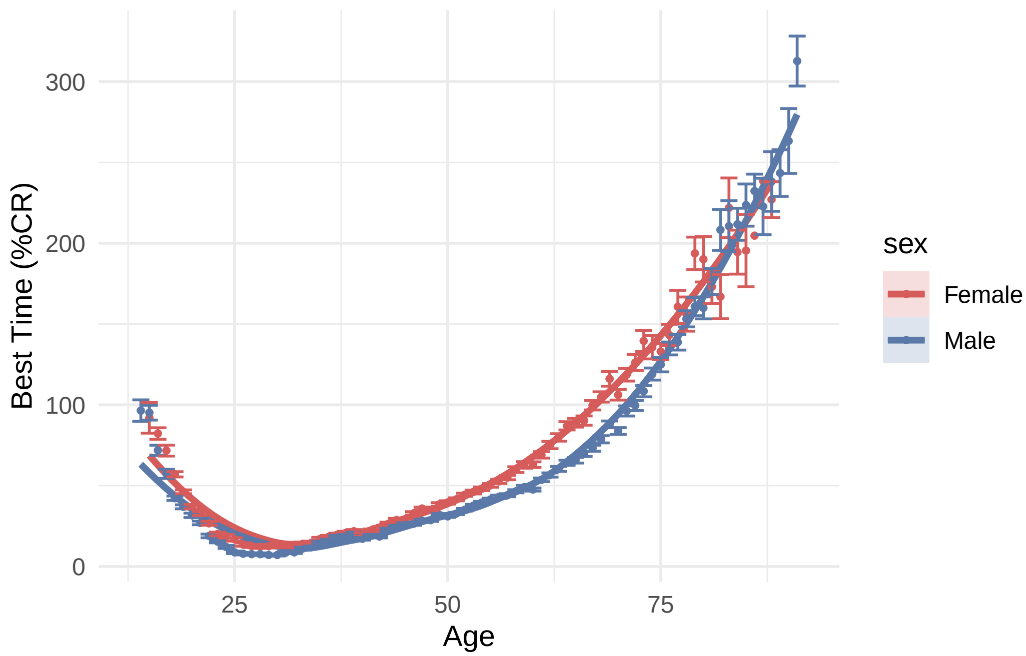

Figure 2: Marathon performance between men and women across age

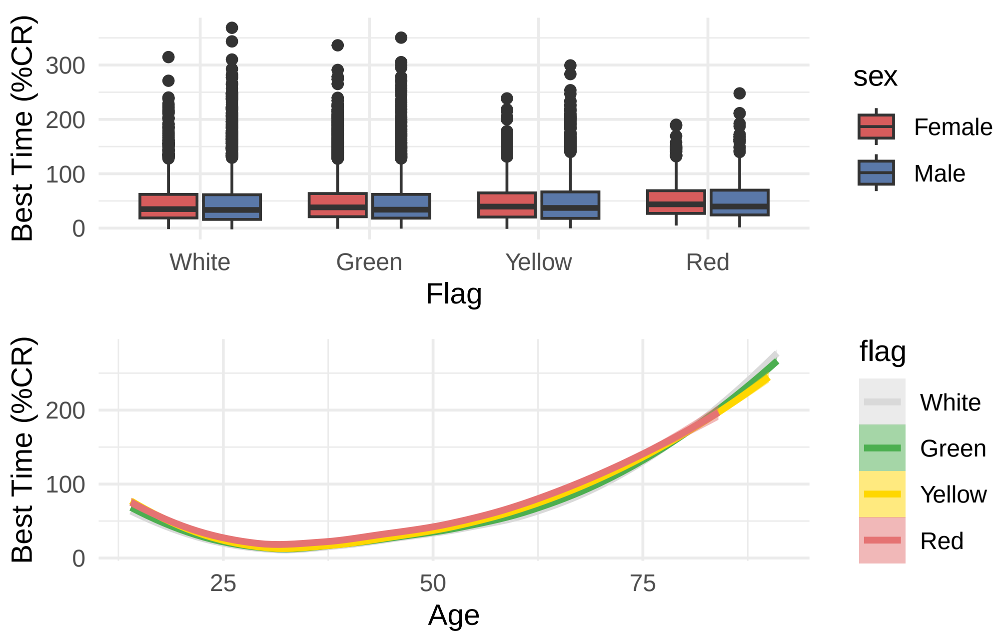

Figure 3: Marathon performance under different weather conditions

individuals completing the marathon significantly narrows. Thus, it can be inferred that in extreme conditions, the increased demand for thermoregulation affects the energy expenditure in marathon running, requiring more time to complete the race. Additionally, Figure 3 suggests that older individuals, around 60 years old, may be the most affected by weather conditions, with the differences in marathon performance across different flags being more pronounced in this age group compared to other age groups.

Marathon Performances across Air Quality

Similarly, we also explored the impact of air quality on marathon performance, as well as the differences in this effect across various age and gender groups. The upper part of Figure 4 illustrates the relationship between AQI and performance, as well as the differences between men and women. We fitted the data using a cubic curve, with the shaded areas representing confidence intervals. It can be seen that the relationship between air quality and performance is quite complex: when the AQI is around 25 or greater than 70, marathon times tend to be longer, while the best performance occurs around an AQI of 45. This overall trend is evident in both male and female groups. Additionally, men appear to be more affected by air quality than women, as reflected by the greater variability in their marathon performance.

To investigate differences across age groups, we fitted curves showing the relationship between age and marathon times under different AQI levels (lower part of Figure 4). While there is no clear clustering of the curves, marathon performance seems to worsen under high AQI conditions, with a greater degree of variation as age increases.

Finding The Most Effective Weather Parameter

Identifying The Weather Parameters

To identify the weather parameters (such as WBGT, flag conditions, temperature, etc.) that have the largest impact on marathon performance, a polynomial regression was applied. To avoid issues related to differing units when making comparisons, we first standardized each weather parameter. Additionally, considering the curvilinear relationships observed in previous exploratory data analysis (EDA), we decided to include cubic regression terms to fit the data. We then filtered out regression terms with significant predictive effects based on their p-values. Finally, we selected the weather parameter that had the largest impact on marathon performance based on their respective effect sizes. As shown in Table 3, the predictive effects of Td^2, Tw^3, rh^2, Tg^2, SR, and SR^2 are significant (with $p \leq .01$). Among these, the first three terms have the largest effect sizes. Therefore, we can conclude that dry bulb temperature (Td), wet bulb temperature (Tw), and relative humidity (rh) have the greatest impacts on marathon performance.

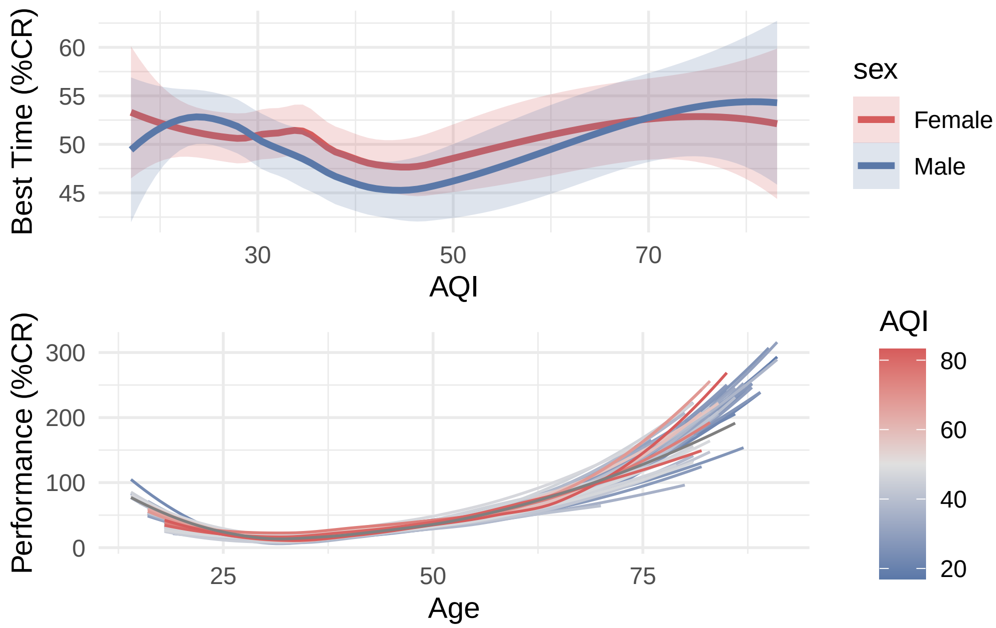

Figure 4: Marathon performance under different environmental conditions

Table 3: Coefficients of each weather variables

term	estimate	p.value
Td^2	-1019.88	0.01
Tw^3	-484.35	0.01
Tg^2	303.55	0.00
rh^2	394.46	0.01
SR	-271.45	0.01
SR^2	-246.88	0.00

Visualizing The Impact

To further visualize the impact of these three weather parameters, we plotted the curve relationships between each parameter and marathon performance. As shown in Figure 5, the relationship between Td and marathon performance remains complex and curved. Individuals have the shortest average completion times when Td is around 10, after which the times increase with either higher or lower Td, and the variation is greater for women than for men. Additionally, Td affects the relationship between age and marathon performance to varying degrees. Individuals perform better at moderate Td levels, while their performance declines at higher Td levels. Older adults have a weaker tolerance for such extreme conditions, resulting in a more pronounced decrease in performance.

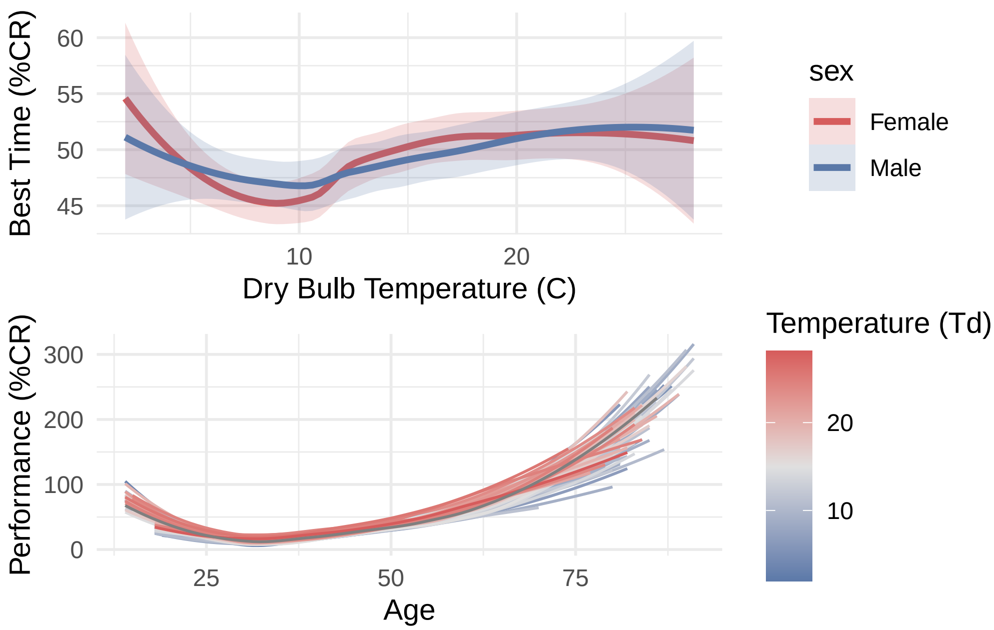

Figure 5

Same things happen with Tw (wet bulb temperature).Individuals perform the best when Tw is around $5 - 7.5$, and female is less adaptive than men across Tw (Figure 6). Similarly, individuals tend to achieve better performance at moderate Tw levels, but their performance deteriorates as Td levels increase. Older adults exhibit a reduced ability to tolerate these extreme conditions, leading to a more significant decline in their performance.

As for relative humidity, its relationship with marathon performances presents a beautiful curve, whose minimum at around 25% while maximum at around 0% or 60% (Figure 7). Men and women show no significant differences on their adaptability. Compared to younger or middle-aged individuals, older adults still exhibit significant differences in marathon performance across different relative humidity environments.

Conclusion

Slowing would be more pronounced with high WBGT in older individuals compared to younger individuals, and similar trends would be observed in both men and women. Individuals in different air quality environments would also exhibit similar performance. Additionally, dry

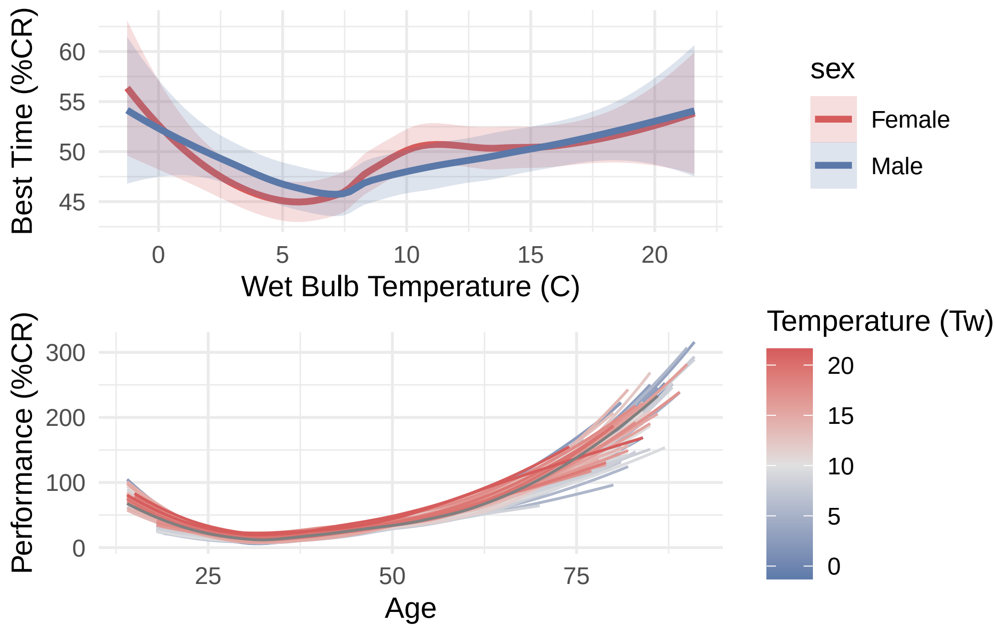

Figure 6

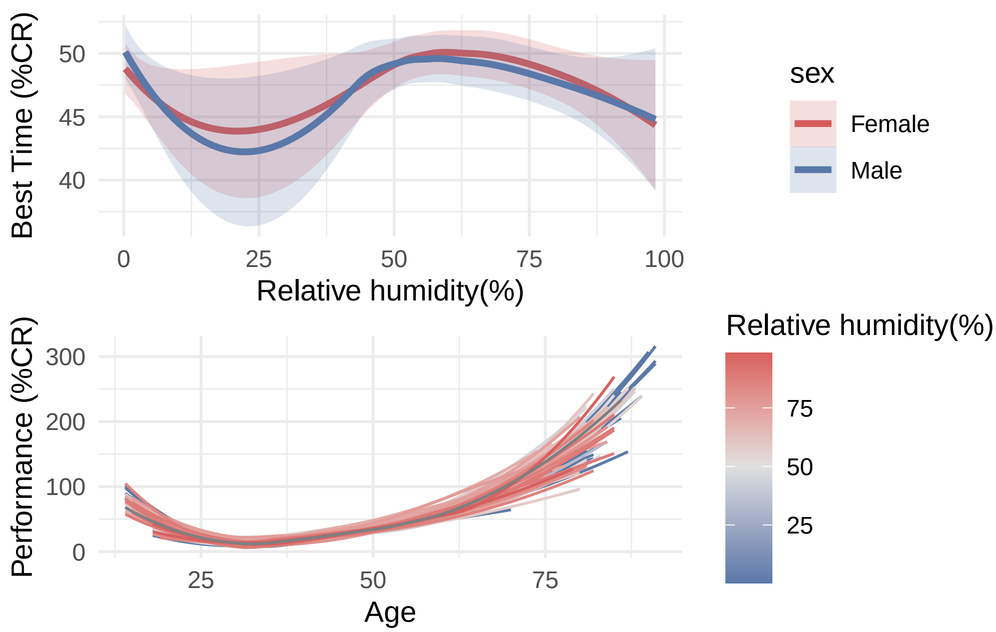

Figure 7

bulb temperature, wet bulb temperature, and relative humidity have relatively significant impacts on marathon performance.

References

Ely, B. R., S. N. Cheuvront, R. W. Kenefick, and M. N. Sawka. 2010. "Aerobic Performance Is Degraded, Despite Modest Hyperthermia, in Hot Environments." *Med Sci Sports Exerc* 42 (1): 135–41.

Ely, M. R., S. N. Cheuvront, W. O. Roberts, and S. J. Montain. 2007. "Impact of Weather on Marathon-Running Performance." *Medicine and Science in Sports and Exercise* 39 (3): 487–93.

Kenney, W. L., and T. A. Munce. 2003. "Invited Review: Aging and Human Temperature Regulation." *Journal of Applied Physiology* 95 (6): 2598–2603.

Yanovich, R., I. Ketko, and N. Charkoudian. 2020. "Sex Differences in Human Thermoregulation: Relevance for 2020 and Beyond." *Physiology* 35 (3): 177–84.

Code Appendix

```r
library(tidyverse)
library(ggplot2)
library(knitr)
library(kableExtra)
library(lubridate)
library(patchwork)
library(gt)
library(gtsummary)
library(broom)
aqi <- read.csv("aqi_values_ext.csv") %>%
  mutate(year=date_local %>% ymd %>% year) %>%
  group_by(marathon, year) %>%
  summarise(aqi=mean(aqi, na.rm=T),
            arithmetic_mean=mean(arithmetic_mean))
df <- read.csv("project1.csv", header=T,
               col.names=c("race", "year", "sex", "flag", "age",
                           "CR", "Td", "Tw", "rh", "Tg", "SR",
                           "DP", "wind", "WBGT")) %>%
  mutate(race=case_when(race==0 ~ "Boston",
                        race==1 ~ "Chicago",
                        race==2 ~ "NYC",
```

```
                        race==3 ~ "TC",
                        race==4 ~ "Duluth",
                        TRUE ~ NA),
         sex=case_when(sex==0 ~ "Female",
                       sex==1 ~ "Male",
                       TRUE ~ NA),
         flag=factor(flag, levels=c("White", "Green", "Yellow", "Red", "Black"))) %>%
  left_join(aqi, by=c("race"="marathon", "year"))
df %>%
  mutate(flag=ifelse(flag=="", NA, flag)) %>%
  tbl_summary(
    by = "sex",  # Group by sex
    statistic = list(
      year ~ "[{min}, {max}]",  # Only range for 'year'
      age ~ "[{min}, {max}]",   # Only range for 'age'
      c(CR, Td, Tw, rh, Tg, SR, DP, wind, WBGT, aqi) ~ "{mean} ({sd})",  # For other continuo
      all_categorical() ~ "{n}/{N} ({p}%)"  # For categorical variables
    ),
    missing = "ifany",  # Add missing data information
    digits = list(c(CR, Td, Tw, rh, Tg, SR, DP, wind, WBGT, aqi) ~ c(2, 2))
  ) %>%
  # Remove the 'Mean (SD), Min - Max' text from the table body
  modify_header(stat_0 = "**Statistics**") %>%
  modify_footnote(
    all_stat_cols() ~ "Mean (SD), Min - Max are reported for continuous variables. Count (%)
  ) %>%
  add_overall() %>%
  add_p()
df %>%
  filter(is.na(WBGT)) %>%
  group_by(year, race) %>%
  summarise(num = n(), .groups = "drop") %>% # Summarise with group by race and year
  left_join(aqi, by=c("race"="marathon", "year")) %>%
  kable(col.names = c("Year", "Race", "Number of cases that missing weather conditions", "AQI
  kable_styling(full_width=F, position="center") %>%
  column_spec(2, width = "3em") %>%
  column_spec(3, width = "10em")
df %>%
  group_by(year, race) %>%
  summarise(num=sum(is.na(aqi))) %>%
  ggplot(aes(x = year, y = num)) +  # Set year as x and num as y
  geom_line(aes(color = factor(race))) +
```

```
  geom_point(aes(color = factor(race))) +  # Plot points
  scale_x_continuous(breaks = seq(min(df$year), max(df$year), by = 1)) +  # Ensure x-axis tic
  labs(
    x = "Year",  # Change x-axis label
    y = "Number of Missing Cases",  # Change y-axis label
    color = "Race",  # Change legend label
  ) +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, vjust = 1, hjust = 1))
df_summary <- df %>%
  group_by(age, sex) %>%
  summarise(mean=mean(CR),
            se=sd(CR)/sqrt(n()),
            .groups="drop")

ggplot(df, aes(x=age, color=sex, fill=sex)) +
  geom_smooth(aes(y = CR), method = "loess", se = TRUE, size=1.2, alpha=0.2) +  # Fit a curve
  geom_point(data=df_summary, aes(y=mean), size=0.8) +
  geom_errorbar(data=df_summary, aes(y=mean, ymin=mean-se, ymax=mean+se), width=2, size=0.5)
  scale_color_manual(values = c("Male" = "#5A78A8", "Female" = "#D65C5C")) +
  scale_fill_manual(values = c("Male" = "#5A78A8", "Female" = "#D65C5C")) +
  labs(x = "Age", y = "Best Time (%CR)") +
  theme_minimal()
plt1 <- df %>%
  filter(flag!="") %>%
  ggplot(aes(x = flag, fill = sex)) +
  geom_boxplot(aes(y = CR)) +  # Set alpha for transparent CIs
  scale_fill_manual(values = c("Male" = "#5A78A8", "Female" = "#D65C5C")) +   # Same color fo
  labs(x = "Flag", y = "Best Time (%CR)") +
  theme_minimal()
plt2 <- df %>%
  filter(flag!="") %>%
  ggplot(aes(x = age, color = flag, fill = flag)) +
  geom_smooth(aes(y = CR), method = "loess", se = TRUE, size = 1.2, alpha = 0.5) +  # Set alp
  scale_color_manual(values = c("White" = "#D9D9D9", "Green" = "#4CAF50",
                                "Yellow"="#FFD700", "Red"="#E57373", "Black"="black")) +  # C
  scale_fill_manual(values = c("White" = "#D9D9D9", "Green" = "#4CAF50",
                                "Yellow"="#FFD700", "Red"="#E57373", "Black"="black")) +   #
  labs(x = "Age", y = "Best Time (%CR)") +
  theme_minimal()
plt1 / plt2
plt1 <- df %>%
```

```
  ggplot(aes(x = aqi, color = sex, fill = sex)) +
  geom_smooth(aes(y = CR), method = "loess", se = TRUE, size = 1.2, alpha = 0.2) +  # Set alp
  scale_color_manual(values = c("Male" = "#5A78A8", "Female" = "#D65C5C")) +  # Custom color
  scale_fill_manual(values = c("Male" = "#5A78A8", "Female" = "#D65C5C")) +   # Same color fo
  labs(x = "AQI", y = "Best Time (%CR)") +
  theme_minimal()
plt2 <- ggplot(df, aes(x = age, y = CR, group = aqi, color = aqi)) +
  geom_smooth(method = "loess", se = FALSE, size = 0.5, alpha=0.2) +  # Smooth curve with no
  scale_color_gradient2(low = "#5A78A8", mid="#E0E0E0", high = "#D65C5C", midpoint=50) +  # C
  labs(x = "Age", y = "Performance (%CR)", color = "AQI") +
  theme_minimal()
plt1 / plt2
var_weather <- c("Td", "Tw", "Tg", "rh", "SR", "DP", "wind")
data <- df %>%
  drop_na() %>%
  mutate(across(var_weather, scale))
form <- paste("CR ~", paste(paste0("poly(", var_weather, ", 3)"),
                            collapse = " + ")
              )
plymdl <- lm(form, data = data)
tidy(plymdl) %>%
  # Filter for significant effects (p-value < 0.05)
  filter(p.value < 0.05 & term != "(Intercept)") %>%
  # Select relevant columns and rename them
  select(term, estimate, p.value) %>%
  mutate(across(c(estimate, p.value), ~ round(., 2)),
         term = gsub("poly\\((\\w+), 3\\)(2)", "\\1^2", term),
         term = gsub("poly\\((\\w+), 3\\)(3)", "\\1^3", term),
         term = gsub("poly\\((\\w+), 3\\)(1)", "\\1", term),
         term = gsub("poly\\((\\w+), 2\\)(2)", "\\1^2", term),
         term = gsub("poly\\((\\w+), 2\\)(1)", "\\1", term)) %>%
  kable() %>%
  kable_styling(full_width=F, position="center")
plt1 <- df %>%
  ggplot(aes(x = Td, color = sex, fill = sex)) +
  geom_smooth(aes(y = CR), method = "loess", se = TRUE, size = 1.2, alpha = 0.2) +  # Set alp
  scale_color_manual(values = c("Male" = "#5A78A8", "Female" = "#D65C5C")) +  # Custom color
  scale_fill_manual(values = c("Male" = "#5A78A8", "Female" = "#D65C5C")) +   # Same color fo
  labs(x = "Dry Bulb Temperature (C)", y = "Best Time (%CR)") +
  theme_minimal()
plt2 <- ggplot(df, aes(x = age, y = CR, group = Td, color = Td)) +
  geom_smooth(method = "loess", se = FALSE, size = 0.5, alpha=0.2) +  # Smooth curve with no
```

```
  scale_color_gradient2(low = "#5A78A8", mid="#E0E0E0", high = "#D65C5C", midpoint=15) +  # C
  labs(x = "Age", y = "Performance (%CR)", color = "Temperature (Td)") +
  theme_minimal()
plt1 / plt2
plt1 <- df %>%
  ggplot(aes(x = Tw, color = sex, fill = sex)) +
  geom_smooth(aes(y = CR), method = "loess", se = TRUE, size = 1.2, alpha = 0.2) +  # Set alp
  scale_color_manual(values = c("Male" = "#5A78A8", "Female" = "#D65C5C")) +  # Custom color
  scale_fill_manual(values = c("Male" = "#5A78A8", "Female" = "#D65C5C")) +   # Same color fo
  labs(x = "Wet Bulb Temperature (C)", y = "Best Time (%CR)") +
  theme_minimal()
plt2 <- ggplot(df, aes(x = age, y = CR, group = Tw, color = Tw)) +
  geom_smooth(method = "loess", se = FALSE, size = 0.5, alpha=0.2) +  # Smooth curve with no
  scale_color_gradient2(low = "#5A78A8", mid="#E0E0E0", high = "#D65C5C", midpoint=10) +  # C
  labs(x = "Age", y = "Performance (%CR)", color = "Temperature (Tw)") +
  theme_minimal()
plt1 / plt2
plt1 <- df %>%
  ggplot(aes(x = rh, color = sex, fill = sex)) +
  geom_smooth(aes(y = CR), method = "loess", se = TRUE, size = 1.2, alpha = 0.2) +  # Set alp
  scale_color_manual(values = c("Male" = "#5A78A8", "Female" = "#D65C5C")) +  # Custom color
  scale_fill_manual(values = c("Male" = "#5A78A8", "Female" = "#D65C5C")) +   # Same color fo
  labs(x = "Relative humidity(%)", y = "Best Time (%CR)") +
  theme_minimal()
plt2 <- ggplot(df, aes(x = age, y = CR, group = rh, color = rh)) +
  geom_smooth(method = "loess", se = FALSE, size = 0.5, alpha=0.2) +  # Smooth curve with no
  scale_color_gradient2(low = "#5A78A8", mid="#E0E0E0", high = "#D65C5C", midpoint=50) +  # C
  labs(x = "Age", y = "Performance (%CR)", color = "Relative humidity(%)") +
  theme_minimal()
plt1 / plt2
```