Formula 1 is the highest class of single seater racing category with a season featuring a series races, known as Grand Prix, which are held in a variety of countries. A single race has drivers each complete a number of laps at a specific circuit. While laps are generally around the same time, in every Grand Prix, a car will complete anywhere between 1 and 5 pit stops in a race to change tires, replace parts, or check damage on the car. Races can also introduce red flags and yellow flags (that indicate race stoppage or caution due to accidents) that could slow a car’s lap time. This dataset has times from the 2023 F1 Miami Grand Prix, where each driver completed 57 laps (minus the lapped cars in last and second-to-last who only completed 56) as there were no red or yellow flags brought out. This worksheet focuses on only lap times from a single driver, Max Verstappen, the winner of the 2023 F1 Miami Grand Prix.

1. Indicate the observations (cases) of the data set for the 2023 F1 Miami Grand Prix

All lap times for each driver that participated in the Grand Prix

1. Use the histogram to describe shape, center, and spread of Max Verstappen’s Lap Times during the race

Shape: Unimodal, right-skewed

Center: Around 92.5 seconds

Spread: Most of the data located around 90 to 93 seconds, with some data around 95, and more outliers around 102 and 106 seconds

1. How many observations are there in the histogram to the right?

57 observations (i.e., the number of laps in the race)

1. Given the following summary statistics of Verstappen’s lap times…

Min.: 1st Qu: Median: Mean: 3rd Qu: Max:

89.71 91.18 91.83 92.25 92.43 106.12

* 1. Determine which lap(s) would be considered an outlier?

**1.5 \* IQR Rule:**
IQR = 92.43 – 91.18 = 1.25

1.5(IQR) = 1.5(1.25) = 1.875

Q1 – 1.875 = 91.18 – 1.875= 89.305
Q3 + 1.875 = 92.43 + 1.875= 94.305

By the 1.5 IQR Rule, only the three longest laps would be outliers as they are all above the upper bound of 94.305. (There is a lap at 94.26 sec. While not an outlier according to this rule, it is visually close enough to the upper bound that students might label it as such if they don’t also have access to the data set.)

1. Given the general description of an F1 race above, what is one reason that there are some laps that are so far away from the rest of the data?

Answers may vary. One possible answer: In a race with no crashes like the Miami GP, pit stop lengths are likely the reason for outliers

1. Would removing these lap times provide a more accurate analysis of a driver’s or contractor’s race?

Answers may vary. One possible answer: Yes, it could be helpful to remove outliers that way we can see the majority spread of the laps in a grand prix.

Helpful data to have to include outliers would be length of pit stops on those laps in order to calculate the “base” lap time.

1. Describe the histogram with outliers removed from the data

Shape: Unimodal, approximately symmetric

Center: Around 92 seconds

Spread: Most of the data located around 90 to 93 seconds, no outliers

1. Would you expect the mean and median lap time to increase or decrease after the outliers are removed? Why?

Mean and median would decrease, because all of the outliers in the data were above Q3 + 1.5(IQR), meaning that the data will shift left.

New median: New mean:

91.80 91.75