Info about the data here

1. A graph with blue and black bars

   Description automatically generatedIndicate 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

Figure 1: Histogram of lap times for Max Verstappen

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

57 observations

1. Given the 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. Find if the lap time 93.5 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 IQR Rule, no, 93.5 would not be considered an outlier

* 1. Lap time 106.123?

By the IQR Rule, yes, 106.123 would be considered an outlier

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. A graph with blue bars

   Description automatically generatedDescribe 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