**Scenario A: 10k Split vs Half Marathon Time**

1.a.

The x variable is 10K split time, and the y variable is half marathon time.

2.a. There is strong, positive, linear association between 10K split times and half marathon times. There are no outliers or influential points.

3.a. The explanatory variable is the 10K split time because it is being used to predict half marathon times, which is the response variable. The slope is 2. For every one minute increase in the 10K split time, there is a predicted 2 minute increase in half marathon time. The y-intercept is 6.3569. When the 10K split time is 0 minutes, the half marathon time is 6.3569 minutes. This value is not realistically feasible because a runner can’t complete a 10K in 0 minutes.

The LSRL equation is ŷ = 6.3569 + 2x or (predicted half marathon time) = 6.3569 + 2(10K split time)

4.a. The correlation coefficient is r = .9825. There is a strong, positive, linear association between 10K split times and half marathon times. The correlation coefficient matches the description of the scatterplot. The coefficient of determination is r² = .96525. Approximately 96.53% of the variation in the predicted half marathon times can be explained by the LSRL of 10K split times.

5.a.

### Based on the residual plot, the LSRL is a good model for the data given, as there is no leftover pattern. The equation for calculating residuals is: y – ŷ or actual half marathon time minus predicted half marathon time. For example, the runner that took 36.53 minutes to run the 10K took 77.78 minutes to run the half marathon. Plugging the point 36.53 into the LSRL equation gives us 79.42 minutes. Actual y-value minus predicted y-value gives us the residual -1.64, meaning the runner completed the half marathon 1.64 minutes faster than expected based on the LSRL line. The residual is graphed in relation to the x-value, and a residual for every y value is calculated and plotted.

**Scenario A: Half Marathon Time vs Final Marathon Time**

1.b.

The x variable is half marathon time, and the y variable is full marathon time.

2.b. There is a strong, positive, linear association between half marathon time and final marathon time. There is a possible outlier at (126.92, 347.58), and there are no influential points.

3.b. The explanatory variable is half marathon time, and the response variable is final marathon time. The slope is 1.9991. For every one minute increase in half marathon time, there is a predicted 1.9991 minute increase in final marathon time. The y-intercept is 13.484. Where half marathon time is 0, final marathon time is 13.484. This is not realistically feasible because a half marathon time can’t equal 0.

The LSRL equation is ŷ = 13.484 + 1.9991x or (predicated final marathon time) = 13.484 + 1.9991(half marathon time)

4.b. The correlation coefficient is r =.9475. There is a strong, positive, linear association between half marathon times and final marathon times. The correlation coefficient matches the description of the scatterplot. The coefficient of determination is r² = .897787. Approximately 89.78% of the variation in the predicted final marathon times can be explained by the LSRL of half marathon times.

5.b.

Based on the residual plot, the LSRL is a good model for the data given, as there is no leftover pattern. The equation for calculating residuals is: y – ŷ or actual half marathon time minus predicted half marathon time. For example, the runner that took 87.08 minutes to run the half marathon took 180.58 minutes to run the full marathon. Plugging the point 87.08 into the LSRL equation gives us 187.57 minutes. Actual y-value minus predicted y-value gives us the residual -6.99, meaning the runner completed the full marathon 6.99 minutes faster than expected based on the LSRL line. The residual is graphed in relation to the x-value, and a residual for every y value is calculated and plotted.

6. Predicted Half Marathon / Full Marathon

Runner 1: 103.96 min / 216.43 min

Runner 2: 118.12 min / 238.50 min

Runner 3: 129.50min/ 267.83 min

Runner 4: 217.26 min / 493.41 min

Runner 5: 79.12 min/ 165.16 min

Discuss things relating to extrapolation and its effects. Answers may vary