1. Which of these histograms of *Victories* for each *Weapon* and *Gender* combination looks normally distributed?

A graph of different levels of a graph

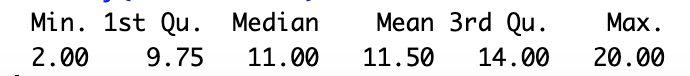
Description automatically generated with medium confidence

The distribution of `Victories` for womens' foil looks to have a normal distribution, due to its symmetrical appearance.

1. The mean number of *Victories* for women’s' foil is 11.5 and the standard deviation is 3.97. Can it be concluded that the distribution of *Victories* for women’s' foil is normal?

No, normal distribution has to have a mean of 0 and a standard deviation of 1, just because a variable has a symmetrical distribution does not mean it is normally distributed.

1. The 5 number summary of *Victories* for women’s' foil is provided below. Calculate the IQR of *Victories* for women’s foil.



IQR = 14-9.75= 4.25

1. This table shows the 5 fencers with the least *Victories* in women's foil. Identify any outliers using the IQR method.

A table with numbers and symbols

Description automatically generated

1.5 \* 4.25= 6.375; 9.75 - 6.375 = 3.375; 2 < 3.375

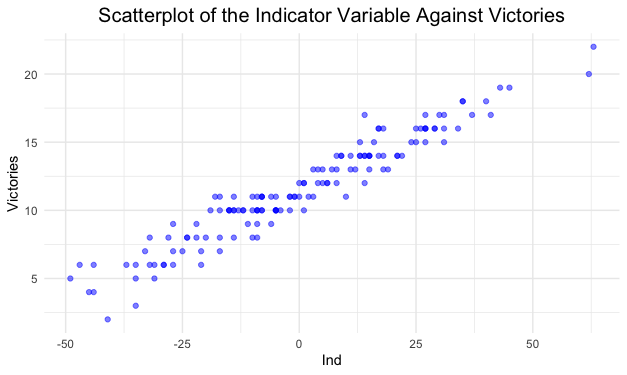
Alina Lee is an outlier since her number of Victories is less than the lower quantile minus 1.5(IQR).

1. Draw a boxplot of *Victories* for women’s' foil.

A diagram of a graph

Description automatically generated

1. Describe the trend of this scatterplot of *Victories* and *Ind*.



Victories increase with a higher indicator variable.

1. The least squares line of *Victories* predicted by *Ind* is . Interpret the slope of the line.

For every unit increase in the indicator variable, a fencer's victories increase by 0.16.

1. A fencer has an indicator score of 16, calculate their predicted *Victories*.

11.5 + (0.16\*16) = 14.06

A fencer with an indicator score of 16 should have 14.06 or roughly 14 victories.

1. The actual value of *Victories* for a fencer with an indicator score of 16 is 15, calculate the residual.

15 – 14 = 1

1. The model has an R2 value of 92%. Explain the significance of this value.

R-squared refers to the proportion of total variability in Y explained by X. In terms of this model, this means that 92% of variability in victories is explained by the indicator variable. This means that the model does a good job predicting the number of victories.