Using statistical and correlative analysis, was there a relationship between the average speed of male biathletes (km/h) and their total hit rate (%) during the Biathlon World Cup season 2022-2023?

Mathematics Analysis & Approaches SL Internal Assessment

Examination session: May 2025

Table of Contents

1. Introduction 2

2. Hypothesis 3

3. Calculations and data analysis 3

Descriptive statistics: 3

Average speed of biathletes (in km/h) 3

Total hit rate of biathletes (in %) 10

Correlation and r-value: 16

4. Conclusion 18

5. Evaluation 19

6. Bibliography 20

7. Appendices 22

Appendix 1: Raw data table 22

Appendix 2: Average speed variance and standard deviation table 24

Appendix 3: Total hit rate variance and standard deviation table 26

Appendix 4: Correlation table 28

1. Introduction

I have been a fan of sports since my childhood. I enjoy watching a lot of sports, but biathlon was one of the first that took a place in my heart. It was highly popular in my home country and people were keen on watching it. My parents recall that when I was about 7 years old, I knew the results of all races, at least the top 10 sportsmen of every race and even how many seconds they lost to the winner. I was interested both in the sport process and the numbers within it. As I gained a lot of new knowledge about statistics during my IB studies, I decided to apply them while delving into the deeper investigation of my favourite sport. When I was watching biathlon with my family and friends, we always asked ourselves a question: Is it the case that when biathlete goes with a high speed on the track, he ends up shooting less accurately due to the increased heart rate and feeling of tiredness? Or are these characteristics not connected at all? This is the question I will investigate from the mathematical perspective in this exploration by finding a possible correlation between the average speed of male biathletes (in km/h) and their total hit rate (in %). Both characteristics are numerical values and, therefore, allow mathematical calculations and analysis to be made. Firstly, I will use "descriptive statistics" to analyse both characteristics separately and eventually the correlation analysis to examine how the variables relate to each other. Consequently, the aim of this exploration is to determine whether there was a correlation between the average ski speed of male biathletes and their total hit rate during the Biathlon World Cup season 2022-2023.

Therefore, my research question is as follows:

Using statistical and correlative analysis, was there a relationship between the average speed of male biathletes (km/h) and their total hit rate (%) during the Biathlon World Cup season 2022-2023?

Statistics is the science of collecting, analysing, presenting, and interpreting data (Sweeney, 2025). In this investigation, I will present the calculations related to my independent variable – the average speed of male biathletes (in km/h), and the dependent variable – total hit rate of sportsmen. In other words, at first, I will present the univariate data analysis that will provide us a clear understanding of these two separate datasets. Thereafter, a bivariate data analysis, which includes finding the correlation coefficient and plotting the data on a scatter diagram, will be applied to determine the strength of the relationships between the variables.

"Microsoft Excel" software will be used to create the scatter diagram and box and whisker plots that will be presented in the section of the same names. In addition to this it will be used in calculating the values for the tables of standard deviation and variance as well as those used in finding the correlation. Worth mentioning that all calculated values will be rounded to 2 decimal places, according to the data sets.

2. Hypothesis

Even though I have been following biathlon for many years, I approach this investigation without any prior answer, as the relationship between chosen variables cannot be observed with your own eyes and only mathematics can serve as a useful tool that enable us to figure this out. However, due to the reason I stated in the introduction, it can be hypothesized that average speed of biathletes and their total hit rate at the shooting range have a strong positive correlation. In addition to this, as the level of competition is extremely high, I hypothesize that the statistical variation within the variables is not high.

3. Calculations and data analysis

Descriptive statistics:

Average speed of biathletes (in km/h)

The set of data that will be analysed in this section is based on the statistics given on webpage https://www.realbiathlon.com/seasons.html?year=2023&gender=M&type=athletes&period=Season

&level=WorldCup&discipline=NonTeam&category=skiing&min=16&chart=true&stat=totalPercent&rank=false&movingavg=none&limit=200 (accessed 10th of December 2024) and it contains the average speeds of 46 biathletes, whose characteristics are applied in this investigation (Season Statistics).

26.67, 26.24, 26.15, 26.11, 26.06, 26.04, 26.03, 25.99, 25.98, 25.88, 25.86, 25.83, 25.79, 25.74, 25.74, 25.65, 25.58, 25.57, 25.56, 25.51, 25.50, 25.45, 25.38, 25.35, 25.32, 25.29, 25.22, 25.19, 25.15, 25.14, 25.14, 25.12, 25.11, 25.10, 24.91, 24.88, 24.88, 24.79, 24.79, 24.77, 24.49, 24.49, 24.46, 24.25, 24.21, 24.11

Table 1. The set of data of average speed of male biathletes during the World Cup season 2022-2023.

There were in total 173 sportsmen who participated in biathlon non-team competitions in the World Cup season 2022-2023 (Season Statistics). However, it was decided to select only those participants who ran at least 16 races to increase the reliability of my data by reducing the influence of outliers that could have occurred with athletes who participated in fewer races. The final sample selected is 46 biathletes and only their characteristics are analysed. Therefore, the number of values n, that is used throughout the calculations equals 46 (the raw data table can be found in Appendix 1).

1. Minimum value:

The minimum value of the data set can be defined as the smallest mathematical value in the data set. The minimum value can be seen from the set of data demonstrated above.

minimum value = 24.11 km/h

This specific biathlete had very a very low speed, especially in comparison to others. According to the hypothesis, his chances on the shooting range are better because he was probably feeling less tired. Being slower however does not allow him to succeed high scores, unless he shoots almost ideally, and others make a lot of misses at the shooting range.

2. Maximum value:

Maximum value can be explained as the largest mathematical value in the data set. The maximum value of the studied characteristic can be found from the data set above.

maximum value = 26.67 km/h

This biathlete, in contrast, was the most successful on the ski track. Nevertheless, he could have made more misses on the range, as his fast speed could have contributed to the feeling of fatigue.

3. Range:

The range in statistics for a given data set is defined as the difference between the highest and lowest values. In other words, it is the difference between the maximum and the minimum values. In the characteristic of the average ski speed of biathletes, the maximum value is 26.67 and the minimum value is 24.11, hence,

$$range = 26.67 - 24.11 = 2.56\ km/h$$

The range of 2.56 km/h demonstrates that the sample has little variation between average ski speeds. However, in biathlon, even low differences in ski speed might significantly affect the results, therefore the range of 2.56 is quite large. So, the range is optimal as larger range would probably mean that the fastest and the slowest sportsmen should not participate in the competitions of the same level.

The following calculations (4-6) are the measures of central tendency. It tells us where the middle of a set of data lies (Awada, 2019).

4. Mean:

The mean is the sum of all the values in the data set, divided by the number of values in this data set (Awada, 2019). Therefore, the formula for the mean is the following:

$$mean = \frac{sum\ of\ all\ data\ values}{numberof\ data\ values} = \bar{x} = \frac{\sum x}{n}$$

The sum of the data values $\sum x$ was calculated by adding the average speeds of all 46 biathletes together: $\sum x = 1166.47$. This number is then divided by the number of values $(n = 46)$ to calculate the mean value:

$$mean = \frac{1166.47}{46} = 25.35804348\ km/h$$

$$mean \approx 25.36\ km/h$$

Therefore, it can be said that in average every biathlete had the average speed of 25.36 km/h on the ski track during the World Cup season 2022-2023.

5. Mode:

Mode is the value that occurs most frequently in a set of data (Awada, 2019) From the set of data being investigated, it can be claimed that the modes are 25.74 km/h, 25.14 km/h, 24.88 km/h, 24.79 km/h, 24.49 km/h as all these values appear in the dataset 2 times. Thus, our dataset is multimodal.

The 5 most occurred value appears in the dataset for only 2 times; hence it can be stated that the dataset is quite widely spread and contains many various values of average speeds of male biathletes. It confirms that all people are different and have different abilities in skiing. According to our hypothesis, the hit rate of biathletes with the same average speed should be very similar. However, it cannot be stated yet as the correlation is unknown.

6. Median:

Median is defined as the middle data value when the data values are arranged in order of size (Awada, 2019). As in our case, the number of values is even, the median is the mean of the two middle numbers. The position of the median can be found by the following formula:

$$The\ position\ of\ the\ median = \left(\frac{n+1}{2}\right)th$$

$$The\ position\ of\ the\ median = \left(\frac{46+1}{2}\right) = 23.5th\ value$$

It is seen that the median is found between 23rd and 24th value in the data set:

$$median = \frac{25.38 + 25.35}{2} = 25.365\ km/h$$

$$median \approx 25.37\ km/h$$

The median value of 25.37 km/h is very close to the mean value. It leads us to the conclusion that almost equal number of values are located below and above the mean value (about 50% of biathletes ski with the speed that is below the mean value and another 50 % ski with the speed that is higher than the mean). To make the strong correlation between this characteristic and their total hit rate of biathletes possible, the same pattern should be observed in the second variable.

7. Variance:

Variance is a statistical measurement of dispersion. In more details, it measures how far each number of the data set is from the mean and thus from every other in the set (Hayes, 2024). It can be calculated by finding the arithmetic mean of the squared differences between each of the values and the mean value. The correction to $n-1$ instead of n in the denominator, known as Bessel´s correction, is made to mend the bias in the estimation of the population variance (Emory. Oxford college. The Department of Mathematics and Computer Science), as in our case, the n value accounts for the number of sportsmen in the sample not a population. Hence, this formula is commonly used:

$$s^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}$$

The value of $\sum(x-\bar{x})^2$ (15.82752391) is calculated by me in a special table shown in appendices (Appendix 2)

$$\frac{\sum_{i=1}^{n}(x_i - \mu)^2}{n-1} = \frac{15.82752391}{46-1} = 0.351722754$$

Therefore, $variance \approx 0.35$

8. Standard deviation:

Standard deviation, which is also a measure of dispersion, is a square root of variance (Awada, 2019). Hence, its formula is:

$$s = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}}$$

$$\sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}} = \sqrt{\frac{15.82752391}{46-1}} = \sqrt{0.351722754} = 0.593062184$$

$$standard\ deviation \approx 0.59\ km/h$$

This value of standard deviation is relatively small, which means that data points are closely clustered around the mean. This proves the research hypothesis that the statistical variation within the independent variable is little due to the high level of competitions.

9. Quartile 1:

"Quartiles are values that divide the data into quarters" (Awada, 2019). The first quartile (Q_1) has 25% of the data below it. To determine the position of the first quartile in the data set, the following formula is applied:

$$Q_1 = \frac{1}{4}(n+1)th\ value$$

$$Q_1 = \frac{1}{4}(46+1) = 11.75th\ value$$

Since my data is presented in the descending order, we should assume that the first value is located at the end of the data set, and thereby find the 11,75th value from the end of the set:

$$Q_1 = x_{11} + 0.75 \times (x_{12} - x_{11})$$

$$Q_1 = 24.88 + 0.75 \times (24.91 - 24.88) = 24.9025$$

$$quartile\ 1 \approx 24.90\ km/h$$

This means that the first 25% of data lie between the values of 24.11 km/h and 24.90 km/h and this range is quite large, especially because of 3 first values that indicate that those 3 biathletes were much weaker on the race in comparison to another sportsmen.

10. Quartile 3:

The third quartile (Q_3) has 75% of the data below it. The position of the third quartile in studied data set is determined below:

$$Q_3 = \frac{3}{4}(n+1)th\ value$$

$$Q_3 = \frac{3}{4}(46+1) = 35.25th\ value$$

To calculate the third quartile, we use 35th and 36th values from the end of the list and obtain:

$$Q_3 = x_{35} + 0.25 \times (x_{36} - x_{35})$$

$$Q_3 = 25.83 + 0.25 \times (25.86 - 25.83) = 25.8375$$

$$quartile\ 3 \approx 25.84\ km/h$$

This shows that the 75% of male biathletes had the average speed between 24.11 km/h (the minimum value) and 25.84 km/h. It also let us to conclude that the last quartile (quartile 4) is the largest in this dataset, due to the maximum value (26.67 km/h), which is caused by the speed of Johannes Boe, who is by far one of the fastest biathletes in 21st century.

11. Interquartile range (IQR):

The IQR is the difference between Q_3 and Q_1 (Awada, 2019). Therefore, its formula is the following:

$$IQR = Q_3 - Q_1$$

$$IQR = 25.8375 - 24.9025 = 0.935$$

$$interquartile\ range \approx 0.94$$

This value showcases that the difference between the third and the first quartiles is 0.94 km/h. This implies that the half of data (50%) lies within this range of average speed. The interquartile range is important component of the box plot that will be demonstrated below.

12. Box and whisker plot:

Box and whisker plot is a graph summarising the set of data (Newcastle University). It includes 6 statistical components such as minimum and maximum values, first and third quartiles as well as mean and median of the data set. The box itself displays the interquartile range.

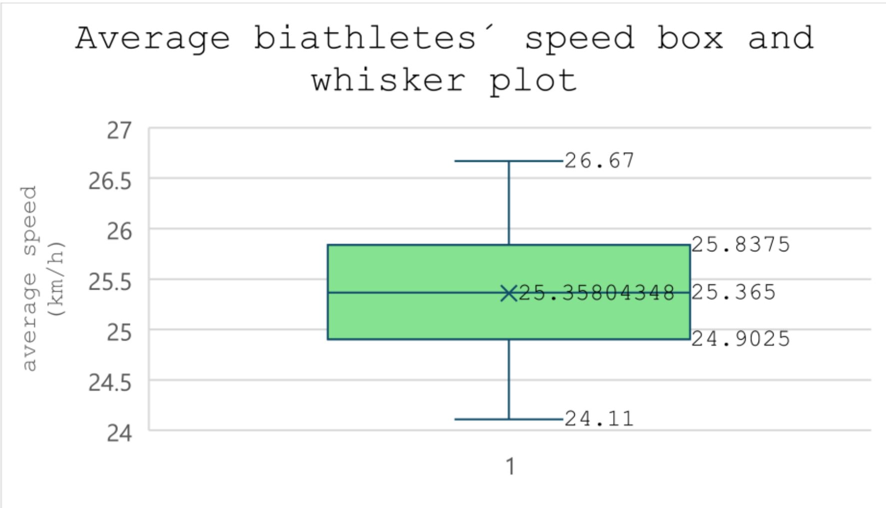

Figure 1. Box and whisker plot for average speeds (km/h) of male biathletes in the season 2022-2023.

The values calculated via “Microsoft Excel” software match the respectful values, which I calculated myself above. As it was stated before, the green box represents the interquartile range (the 50 % between the first and the third quartile). The so-called whiskers represent the values beyond these quartiles. The box is on almost equal distance from maximum and minimum values, meaning that the data is evenly spread.

<u>Total hit rate of biathletes (in %)</u>

The sat of data demonstrated below is based on the statistics given on the webpage https://www.realbiathlon.com/seasons.html?year=2023&gender=M&type=athletes&period=Season

&level=WorldCup&discipline=NonTeam&category=shooting&min=16&chart=true&stat=skiRank&rank=false&movingavg=none&limit=200 (accessed 10th of December 2024) and it consists of the total hit rates of the same sample of biathletes as previously on calculations with the average ski speed (Season Statistics).

93.51, 93.33, 92.73, 91.48, 91.25, 90.00, 90.00, 90.00, 89.66, 89.44, 89.23, 89.17, 88.61, 87.74, 87.63, 87.14, 87.08, 87.04, 86.67, 86.67, 86.43, 85.83, 85.79, 85.00, 85.00, 84.76, 84.72, 84.57, 84.40, 83.75, 83.57, 83.33, 83.33, 82.73, 81.72, 81.50, 81.21, 80.71, 80.00, 79.63, 79.29, 79.17, 78.08, 77.00, 76.15, 74.29

Table 2. The set of data of the total hit rate of male biathletes during the World Cup season 2022-2023.

In this section, the total hit rates of 46 biathletes will be analysed in the same way as their average speeds were analysed in the previous section. The calculations will be explained in less details to prevent to appearance of repetitive information in the investigation. Worth mentioning that the values are presented in the descending order and the position of the value therefore does not necessarily relate to the value of the average speed of the same sportsman. The average speed and the corresponding total hit rate will be presented in the next section dedicated to finding the correlation.

1. Minimum value:

minimum value = 74.29 %

This athlete demonstrated the worst shooting and the result of 74.29 % indicates that, on average, he hits less than 4 targets on each shooting range (as one shooting session in biathlon consists of 5 targets). According to our hypothesis, he should have run at a high pace. However, it can be seen in the correlation table, which will be discussed in the following section, that this biathlete had the 19th average speed out of 46 athletes, which implies that he was a middle skier and bad shooter. Therefore, it questions our hypothesis.

2. Maximum value:

maximum value = 93.51 %

This sportsman showed himself as the best shooter, and the value of 93.51 % says that he passed many shooting sessions without a single miss, allowing him to exit the shooting range before many of rivals, who went to the penalty loop more often. According to the hypothesis, he should have been the worst skier and that is the reason he shot well on ranges. Nonetheless, the correlation table demonstrated he was the 7^{th} best skier, which contradicts our hypothesis.

3. Range:

As the maximum value of total hit rate is 93.51 % and the minimum value is 74.29 %, the range is:

$$range = 93.51 - 74.29 = 19.22\ \%$$

The range of 19.22 % implies that the data is quite spread within this set, and the best and the worst sportsmen in shooting component differ in about 1 hit target. Hence, on average, the worst shooter closes 1 less target on each range than the best shooter. However, standard deviation will better describe the variation within the set.

4. Mean:

The sum of the data values $\sum x$ was calculated:

$$\sum x = 3920.34$$

Therefore, the mean is:

$$mean = \frac{3920.34}{46} = 85.22478261\ \%$$

$$mean \approx 85.22\ \%$$

Thus, it can be claimed that on average, every player hits 85.22 % targets when they shoot, meaning that it is more likely for them to close 4 targets out of 5.

5. Mode:

From the set of data being investigated, it can be claimed that the mode is 90.00 % as this value appears in the dataset 3 times.

6. Median:

$$The\ position\ of\ the\ median = \left(\frac{n+1}{2}\right)th \rightarrow The\ position\ of\ the\ median = \left(\frac{46+1}{2}\right)$$

$$= 23.5th\ value$$

$$median = \frac{85.79 + 85.00}{2} = 85.395\ \%$$

$$median \approx 85.40\ \%$$

The median is close to the mean value, again, as it was observed when analysing the average speed variable. It says that the data contains almost equal number of points below and above the mean. It makes it possible for the correlation to be strong as the statistical features are identical in regards to mean and median, however it does not guarantee this, as it can be just a coincidence.

7. Variance:

$$s^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}$$

The value of $\sum(x - \bar{x})^2$ (977.5999478) is calculated via the special table shown in appendices (Appendix 3)

$$\frac{\sum_{i=1}^{n}(x_i - \mu)^2}{n} = \frac{977.5999478}{46-1} = 21.72444328$$

Therefore, $variance \approx 21.72$

8. Standard deviation:

$$s = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}}$$

$$\sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}} = \sqrt{\frac{977.5999478}{46-1}} = \sqrt{21.72444328} = 4.660948754$$

$$standard\ deviation \approx 4.66\ \%$$

This value of standard deviation is not high, which demonstrates little statistical variation within this variable, for the same reason as with the average speed, discussed above.

9. Quartile 1:

$$Q_1 = \frac{1}{4}(n+1)th\ value \rightarrow Q_1 = \frac{1}{4}(46+1) = 11.75th\ value$$

Again, since the data is in the descending order, we assume that the first value is located at the end of the data set, and thereby find the 11,75th value from the end of the set:

$$Q_1 = x_{11} + 0.75 \times (x_{12} - x_{11})$$

$$Q_1 = 81.50 + 0.75 \times (81.72 - 81.50) = 81.665\ \%$$

$$quartile\ 1 \approx 81.67\ \%$$

Even though the variation in the whole data set is not so significant, it is quite high in the first 25% of the data. The minimum value is 74.29 % and the first quartile´s 81.67 %, meaning that more than one third of the range is in the first 25% of values.

10. Quartile 3:

$$Q_3 = \frac{3}{4}(n+1)th\ value \rightarrow Q_3 = \frac{3}{4}(46+1) = 35.25th\ value$$

$$Q_3 = x_{35} + 0.25 \times (x_{36} - x_{35})$$

$$Q_3 = 89.17 + 0.25 \times (89.23 - 89.17) = 89.185\ \%$$

$$quartile\ 3 \approx 89.19\ \%$$

This value reveals that first 75% of data values in this set are in interval from 74.29 % (as it is the minimum) and 89.19 %. It also implies that the last quartile contains values from 89.19 % to 93.51 %, and the sportsmen with these percentage of correct shots were quite successful on the range.

11. Interquartile range (IQR):

$$IQR = Q_3 - Q_1 = 89.185 - 81.665 = 7.52\ \%$$

$$interquartile\ range = 7.52\ \%$$

The value of 7.52 % for the interquartile range means that 50 % of biathletes´ total hit rates lie within this interval.

12. Box and whisker plot:

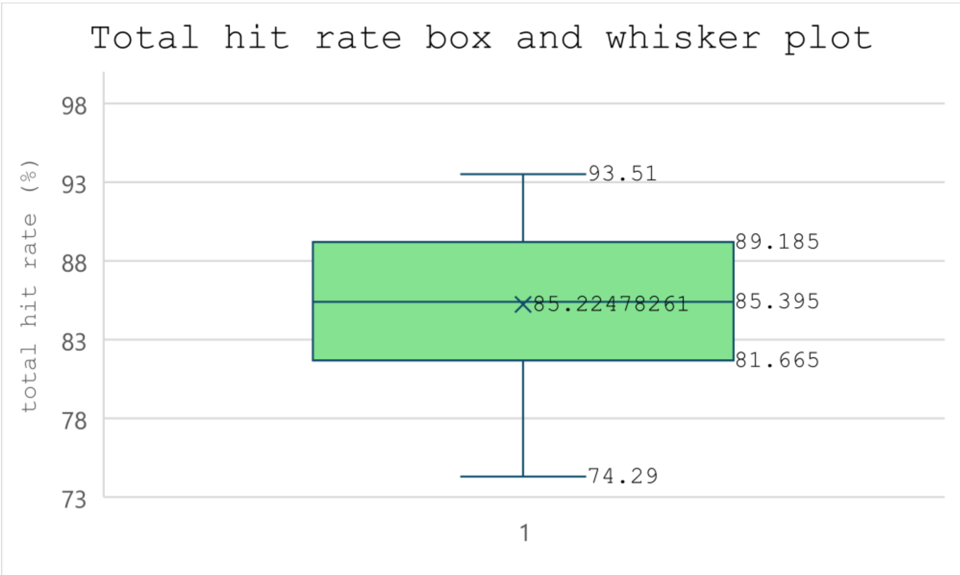

Figure 2. Box and whisker plot for the total hit rate (%) of male biathletes in the season 2022-2023.

The box and whisker plot provides us with the great demonstration of how the values are distributed in the given set of data. It can be seen that the values in the lower quartile are more spread than the values in the fourth quartile as the whisker below the box is longer than the whisker above it. The box is closer to the maximum value and that implies the people with the total hit rate closer to minimum value were quite rare during the season 2022-2023 in male competitions. The trend is not similar to the trend in the independent variable, where the data is more equally spread.

Correlation and r-value:

In this section, the r-value, also known as the Pearson product-moment correlation coefficient, of the correlation between two studied variables, average speed of biathletes and their total hit rate on the shooting range, is calculated according to the formula (Awada, 2019) provided:

$$r = \frac{S_{xy}}{S_x S_y}$$

The Pearson´s correlation coefficient helps us understand both the strength and the direction of possible relationship between the variables (Statistics Solutions, 2025). The range of this coefficient is from -1 value that indicates perfect negative correlation to +1, showing the correlation to be perfectly positive. Hence, if the value is 0, there is no correlation between the variables. The other values of the correlation coefficient should be interpreted in the way shown below (Awada, 2019):

$$0 < |r| \leq 0.25 \rightarrow very\ weak$$

$$0.25 < |r| \leq 0.5 \rightarrow weak$$

$$0.5 < |r| \leq 0.75 \rightarrow moderate$$

$$0.75 < |r| \leq 1 \rightarrow strong$$

Values used in the calculations below are calculated via Microsoft Excel and are shown in the correlation table seen in the appendices (Appendix 4). In the formula of r-value, the numerator equals:

$$S_{xy} = \sum xy - \frac{(\sum x)(\sum y)}{n} = 99412.881 - \frac{1166.47 \times 3920.34}{46} = 0.728830435$$

Calculating the denominators:

$$S_x = \sqrt{\sum x^2 - \frac{(\sum x)^2}{n}} = \sqrt{29595.2245 - \frac{1166.47^2}{46}} = 3.978382072$$

$$S_y = \sqrt{\sum y^2 - \frac{\left(\sum y\right)^2}{n}} = \sqrt{335087.7242 - \frac{3920.34^2}{46}} = 31.26659476$$

Using these values, the r-value can be found:

$$r = \frac{S_{xy}}{S_x S_y} = \frac{0.728830435}{3.978382072 \times 31.26659476}$$

$$r = 0.005859215$$

$r = 0.005859215 \approx 0.00586 \approx 5.86 \times 10^{-3} \rightarrow the\ correlation\ is\ very\ weak\ (0 < |r| \leq 0.25)$

The r-value demonstrates that the correlation between the average speed of male biathletes (km/h) and their total hit rate (%) during the Biathlon World Cup season 2022-2023 was very weak. The value is really close to zero and it can be stated that there was almost no correlation, and thus relationship between them.

Scatter diagram provides as with the visual representation of the relationship between variables. In this scatter diagram, average speed of biathletes is on the horizontal x-axis as it is the independent variable and total hit rate on vertical y-axis, because it is the dependent variable.

The scatter diagram shown below is created via the Microsoft Excel software.

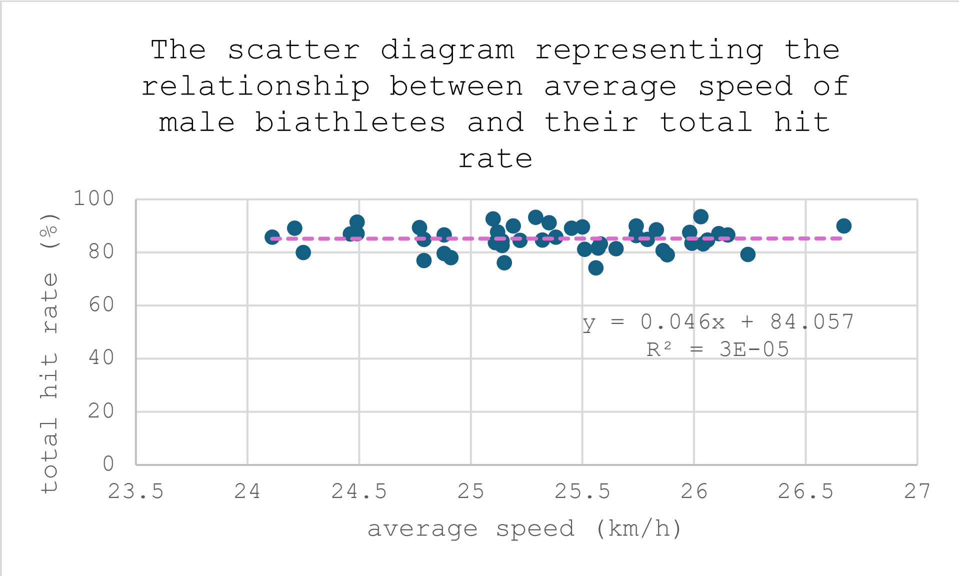

Figure 3. The scatter diagram demonstrating the correlation between average speed and total hit rate of male biathletes during World Cup season 2022-2023.

There is not upward or downward trend of the line of best fit, and it is almost flat. That represents the correlation to be insignificant.

4. Conclusion

During this investigation, it was found out via the statistical and correlation analysis, that there is no relationship between the average speed of male biathletes and their total hit rate on shooting ranges during the World Cup season 2022-2023. Therefore, the main research hypothesis can be rejected. Even though there were some similar trends in statistical analysis, some patterns varied and contradicted our hypothesis about the relationship between variables. It was supported by the correlation found to be about 5.86×10^{-3}. In contrast, the additional hypothesis about the little statistical variation within both data sets was confirmed.

Implying that the correlation is neither negative nor positive, it leads us to conclude that these components of biathlon athletes´ performance are independent of each other and their quality depends only on separate skill in this component of each biathlete.

5. Evaluation

The strength of this investigation was that the data was taken from the reliable website, where I often search for some statistical characteristics of biathletes. As I have been into learning biathlon and watching almost all the races, I can determine when the website gives unreliable information, which is not the case with this website. In addition to this, the advantage of this investigation is my personal interest in biathlon, because I was motivated to explore chosen topic and calculate the correlation, implying I did all the steps thoroughly and consulted some mathematical websites as well as IB Mathematics book to prevent using incorrect formulas. Since I know now that the variables are not related to each other, I will share this information with my family and friends, who enjoy watching biathlon. Another advantage of this work is using a sufficient sample size, 46 people. This allows us to state that the findings can be generalized, at least to all professional sportsmen.

One serious limitation of my investigation is that average speed values was used that reflect the speed during the whole distance. It might be that some of the athletes slow down before the shooting range that allows them to shoot better and their highly increase their speed right after the shooting session is over, which makes their average speed throughout the whole distance still high enough. It can be improved by taking into account only the speed during the last, for example, minute before the range. It would probably provide us with more meaningful insights and either confirm my findings or contradict them. This is what could be done in future research. In addition, shooting accuracy is influenced by the time spent on shooting. In our context, if a sportsman was tired during the loop, he could have just simply spent more time on shooting range to prevent inaccurate shooting due to increased heart rate. Therefore, there are some confounding variables that need to be considered along with the studied variables. It can be stated that even though a conclusion with the clear evidence was made, this topic can be further studied to get a more complete picture of it.

6. Bibliography

Awada, e. a. (2019). In e. a. Awada, *Mathematics: Analysis and Approaches. Standard level* (pp. 276-351). Oxford: Oxford University Press.

Emory. Oxford college. The Department of Mathematics and Computer Science. (n.d.). *Bessel's correction*. Retrieved from https://mathcenter.oxford.emory.edu/site/math117/besselCorrection/

Hayes, A. (2024, September 20). *What is variance in statistics? Definition, formula, and example*. Retrieved from Investopedia: https://www.investopedia.com/terms/v/variance.asp

Newcastle University. (n.d.). *Box and Whisker Plots*. Retrieved from Numeracy, Maths and Statistics - Academic Skills kit: https://www.ncl.ac.uk/webtemplate/ask-assets/external/maths-resources/statistics/data-presentation/box-and-whisker-plots.html

Season Statistics. (n.d.). Retrieved from real biathlon: https://www.realbiathlon.com/seasons.html?year=2023&gender=M&type=athletes&period=Season&level=WorldCup&discipline=NonTeam&category=skiing&min=16&chart=true&stat=skiRank&rank=false&movingavg=none&limit=200

Season Statistics . (n.d.). Retrieved from real biathlon: https://www.realbiathlon.com/seasons.html?year=2023&gender=M&type=athletes&period=Season&level=WorldCup&discipline=NonTeam&category=shooting&min=16&chart=true&stat=skiRank&rank=false&movingavg=none&limit=200

Statistics Solutions. (2025, January 17). *Understanding Pearson's Correlation Coefficient: A Guide to Statistical Association*. Retrieved from Statistics Solutions: https://www.statisticssolutions.com/free-resources/directory-of-statistical-analyses/pearsons-correlation-coefficient/

Sweeney, D. J. (2025, January 17). *Statistics | Definition, Types, & Importance*. Retrieved from Encyclopedia Britannica: https://www.britannica.com/science/statistics

7. Appendices

Appendix 1: Raw data table

Family name	**Average speed (in km/h)**	**Total hit rate (in %)**
Boe J.	26.67	90.00
Ponsiluoma	26.24	79.29
Samuelsson	26.15	86.67
Christiansen	26.11	87.14
Doll	26.06	84.76
Dale	26.04	83.33
Laegreid	26.03	93.51
Claude Fa.	25.99	83.57
Boe T.	25.98	87.63
Kuehn	25.88	79.17
Giacomel	25.86	80.71
Filon Maillet	25.83	88.61
Nelin	25.79	85.00
Rees	25.74	90.00
Krchmar	25.74	86.43
Guigonnat	25.65	81.50

Nawrath	25.58	83.33
Seppälä	25.57	81.72
Stvrtecky	25.56	74.29
Rastorgujevs	25.51	81.21
Fak	25.50	89.66
Claude Fl.	25.45	89.17
Zobel	25.38	85.79
Stalder	25.35	91.25
Strolia	25.32	84.72
Hartweg	25.29	93.33
Hiidensalo	25.22	84.57
Strelow	25.19	90.00
Iliev	25.15	76.15
Doherty	25.14	82.73
Lapshin	25.14	84.40
Dudchenko	25.12	87.74
Bionaz	25.11	83.75
Eder	25.10	92.73
Langer	24.91	78.08

Magazeev	24.88	79.63
Mikyska	24.88	86.67
Vidmar	24.79	85.00
Femling	24.79	77.00
Komatz	24.77	89.44
Cisar	24.49	91.48
Dombrovski	24.49	87.08
Tsymbal	24.46	87.04
Lemmerer	24.25	80.00
Shamaev	24.21	89.23
Sima	24.11	85.83

Table 3. Raw data table containing the values of average speed (km/h) and total hit rate (%) of 46 biathletes in 2022-23 season.

Appendix 2: Average speed variance and standard deviation table

x value	frequency	x-mean	$(x\text{-mean})^2$	mean
26.67	1	1.31195652	1.72122991	25.35804348
26.24	1	0.88195652	0.777847303	
26.15	1	0.79195652	0.62719513	
26.11	1	0.75195652	0.565438608	
26.06	1	0.70195652	0.492742956	
26.04	1	0.68195652	0.465064695	
26.03	1	0.67195652	0.451525565	

25.99	1	0.63195652	0.399369043	
25.98	1	0.62195652	0.386829913	
25.88	1	0.52195652	0.272438609	
25.86	1	0.50195652	0.251960348	
25.83	1	0.47195652	0.222742957	
25.79	1	0.43195652	0.186586435	
25.74	2	0.38195652	0.145890783	
25.65	1	0.29195652	0.08523861	
25.58	1	0.22195652	0.049264697	
25.57	1	0.21195652	0.044925566	
25.56	1	0.20195652	0.040786436	
25.51	1	0.15195652	0.023090784	
25.50	1	0.14195652	0.020151654	
25.45	1	0.09195652	0.008456002	
25.38	1	0.02195652	0.000482089	
25.35	1	-0.00804348	0.0000646976	
25.32	1	-0.03804348	0.001447306	
25.29	1	-0.06804348	0.004629915	
25.22	1	-0.13804348	0.019056002	
25.19	1	-0.16804348	0.028238611	
25.15	1	-0.20804348	0.04328209	
25.14	2	-0.21804348	0.047542959	
25.12	1	-0.23804348	0.056664698	
25.11	1	-0.24804348	0.061525568	

25.10	1	-0.25804348	0.066586438	
24.91	1	-0.44804348	0.20074296	
24.88	2	-0.47804348	0.228525569	
24.79	2	-0.56804348	0.322673395	
24.77	1	-0.58804348	0.345795134	
24.49	2	-0.86804348	0.753499483	
24.46	1	-0.89804348	0.806482092	
24.25	1	-1.10804348	1.227760354	
24.21	1	-1.14804348	1.318003832	
24.11	1	-1.24804348	1.557612528	
			$\sum = 15.82752391$	

Table 4. Average speed variance and standard deviation table.

Appendix 3: Total hit rate variance and standard deviation table

y value	frequency	(y-mean)	(y-mean)^2	mean
93.51	1	8.28521739	68.6448272	85.22478261
93.33	1	8.10521739	65.69454894	
92.73	1	7.50521739	56.32828807	
91.48	1	6.25521739	39.1277446	
91.25	1	6.02521739	36.3032446	
90.00	3	4.77521739	22.80270112	
89.66	1	4.43521739	19.6711533	
89.44	1	4.21521739	17.76805764	
89.23	1	4.00521739	16.04176634	
89.17	1	3.94521739	15.56474025	

88.61	1	3.38521739	11.45969678	
87.74	1	2.51521739	6.326318519	
87.63	1	2.40521739	5.785070693	
87.14	1	1.91521739	3.668057651	
87.08	1	1.85521739	3.441831564	
87.04	1	1.81521739	3.295014173	
86.67	2	1.44521739	2.088653304	
86.43	1	1.20521739	1.452548957	
85.83	1	0.60521739	0.366288089	
85.79	1	0.56521739	0.319470698	
85.00	2	-0.22478261	0.050527222	
84.76	1	-0.46478261	0.216022875	
84.72	1	-0.50478261	0.254805483	
84.57	1	-0.65478261	0.428740266	
84.40	1	-0.82478261	0.680266354	
83.75	1	-1.47478261	2.174983747	
83.57	1	-1.65478261	2.738305486	
83.33	2	-1.89478261	3.590201139	
82.73	1	-2.49478261	6.223940271	
81.72	1	-3.50478261	12.28350114	
81.50	1	-3.72478261	13.87400549	
81.21	1	-4.01478261	16.11847941	
80.71	1	-4.51478261	20.38326202	
80.00	1	-5.22478261	27.29835332	

79.63	1	-5.59478261	31.30159245	
79.29	1	-5.93478261	35.22164463	
79.17	1	-6.05478261	36.66039245	
78.08	1	-7.14478261	51.04791854	
77.00	1	-8.22478261	67.64704898	
76.15	1	-9.07478261	82.35167942	
74.29	1	-10.93478261	119.5694707	
			$\sum = 977.5999478$	

Table 5. Biathletes´ total hit rate variance and standard deviation table.

Appendix 4: Correlation table

Sportsman	**x**	**y**
Boe J.	26.67	90.00
Ponsiluoma	26.24	79.29
Samuelsson	26.15	86.67
Christiansen	26.11	87.14
Doll	26.06	84.76
Dale	26.04	83.33
Laegreid	26.03	93.51
Claude Fa.	25.99	83.57
Boe T.	25.98	87.63
Kuehn	25.88	79.17
Giacomel	25.86	80.71
Fillon Maillet	25.83	88.61
Nelin	25.79	85.00

Rees	25.74	90.00
Krchmar	25.74	86.43
Guigonnat	25.65	81.50
Nawrath	25.58	83.33
Seppälä	25.57	81.72
Stvrtecky	25.56	74.29
Rastorgujevs	25.51	81.21
Fak	25.50	89.66
Claude Fl.	25.45	89.17
Zobel	25.38	85.79
Stalder	25.35	91.25
Strolia	25.32	84.72
Hartweg	25.29	93.33
Hiidensalo	25.22	84.57
Strelow	25.19	90.00
Iliev	25.15	76.15
Doherty	25.14	82.73
Lapshin	25.14	84.40
Dudchenko	25.12	87.74
Bionaz	25.11	83.75
Eder	25.10	92.73
Langer	24.91	78.08
Magazeev	24.88	79.63
Mikyska	24.88	86.67

Vidmar	24.79	85.00
Femling	24.79	77.00
Komatz	24.77	89.44
Cisar	24.49	91.48
Dombrovski	24.49	87.08
Tsymbal	24.46	87.04
Lemmerer	24.25	80.00
Shamaev	24.21	89.23
Sima	24.11	85.83
	$\sum = 1166.47$	$\sum = 3920.34$

Sportsman	xy	x^2	y^2
Boe J.	2400.3	711.2889	8100
Ponsiluoma	2080.5696	688.5376	6286.9041
Samuelsson	2266.4205	683.8225	7511.6889
Christiansen	2275.2254	681.7321	7593.3796
Doll	2208.8456	679.1236	7184.2576
Dale	2169.9132	678.0816	6943.8889
Laegreid	2434.0653	677.5609	8744.1201
Claude Fa.	2171.9843	675.4801	6983.9449
Boe T.	2276.6274	674.9604	7679.0169
Kuehn	2048.9196	669.7744	6267.8889
Giacomel	2087.1606	668.7396	6514.1041
Fillon Maillet	2288.7963	667.1889	7851.7321
Nelin	2192.15	665.1241	7225
Rees	2316.6	662.5476	8100
Krchmar	2224.7082	662.5476	7470.1449
Guigonnat	2090.475	657.9225	6642.25
Nawrath	2131.5814	654.3364	6943.8889
Seppälä	2089.5804	653.8249	6678.1584
Stvrtecky	1898.8524	653.3136	5519.0041
Rastorgujevs	2071.6671	650.7601	6595.0641
Fak	2286.33	650.25	8038.9156
Claude Fl.	2269.3765	647.7025	7951.2889
Zobel	2177.3502	644.1444	7359.9241

Stalder	2313.1875	642.6225	8326.5625
Strolia	2145.1104	641.1024	7177.4784
Hartweg	2360.3157	639.5841	8710.4889
Hiidensalo	2132.8554	636.0484	7152.0849
Strelow	2267.1	634.5361	8100
Iliev	1915.1725	632.5225	5798.8225
Doherty	2079.8322	632.0196	6844.2529
Lapshin	2121.816	632.0196	7123.36
Dudchenko	2204.0288	631.0144	7698.3076
Bionaz	2102.9625	630.5121	7014.0625
Eder	2327.523	630.01	8598.8529
Langer	1944.9728	620.5081	6096.4864
Magazeev	1981.1944	619.0144	6340.9369
Mikyska	2156.3496	619.0144	7511.6889
Vidmar	2107.15	614.5441	7225
Femling	1908.83	614.5441	5929
Komatz	2215.4288	613.5529	7999.5136
Cisar	2240.3452	599.7601	8368.5904
Dombrovski	2132.5892	599.7601	7582.9264
Tsymbal	2128.9984	598.2916	7575.9616
Lemmerer	1940	588.0625	6400
Shamaev	2160.2583	586.1241	7961.9929
Sima	2069.3613	581.2921	7366.7889
	$\sum = 99412.881$	$\sum = 29595.2245$	$\sum = 335087.7242$

Table 6. Correlation table.