**Mean Calculation:**

The formula for the mean (average) of a set of values is:

*Mean=Sum of all values/Number of values*

**Population Standard Deviation Calculation:**

The formula for the population standard deviation of a set of values is:

*Population Standard Deviation = sqrt(Sum of squared differences / (n))*

Sepal length: [6.3, 5.1, 5.7, 5.0, 4.8, 6.6, 6.3, 6.3, 6.0, 4.7, 5.0, 6.1, 6.8, 4.5, 7.7]

Sepal width: [2.9, 3.4, 2.5, 3.5, 3.4, 2.9, 3.3, 3.4, 3.4, 3.2, 3.3, 2.9, 3.2, 2.3, 3.0]

Sepal length mean:

Sum of values = 6.3 + 5.1 + 5.7 + ... + 4.5 + 7.7 = 86.899

**Mean(length) = Sum of values / Number of values = 86.0 / 15 = 5.793**

Sepal width mean:

Sum of values = 2.9 + 3.4 + 2.5 + ... + 2.3 + 3.0 = 46.59

**Mean(width) = Sum of values / Number of values = 48.6 / 15 = 3.106**

Now, let's calculate the standard deviation:

Sepal length standard deviation:

Mean = 5.793

Sum of squared differences = (6.3 - 5.793)^2 + (5.1 - 5.793)^2 + ... + (7.7 - 5.793)^2 = 11.809

**Standard Deviation(length) = sqrt(Sum of squared differences / (n)) = sqrt(11.809 / 15) ≈ 0.88729**

Sepal width standard deviation:

Mean = 3.24

Sum of squared differences = (2.9 - 3.24)^2 + (3.4 - 3.24)^2 + ... + (3.0 - 3.24)^2 = 1.749

**Standard Deviation(width) = sqrt(Sum of squared differences / (n ) = sqrt(1.749 / 15) ≈ 0.34149**

After Z-score normalization

New\_value(length) = old\_value(length) – mean(length)/std(length)

New\_value(length) = old\_value(length) – mean(length)/std(length)

|  | **Sepal length** | **Sepal width** | **Class** |
| --- | --- | --- | --- |
| **0** | 0.571025 | -0.605173 | virginica |
| **1** | -0.781403 | 0.858956 | setosa |
| **2** | -0.105189 | -1.776477 | virginica |
| **3** | -0.894105 | 1.151781 | setosa |
| **4** | -1.119510 | 0.858956 | setosa |
| **5** | 0.909132 | -0.605173 | versicolor |
| **6** | 0.571025 | 0.566130 | versicolor |
| **7** | 0.571025 | 0.858956 | versicolor |
| **8** | 0.232918 | 0.858956 | versicolor |
| **9** | -1.232212 | 0.273304 | ??? |
| **10** | -0.894105 | 0.566130 | ??? |
| **11** | 0.345621 | -0.605173 | ??? |
| **12** | 1.134537 | 0.273304 | ??? |
| **13** | -1.457617 | -2.362128 | ??? |
| **14** | 2.148858 | -0.312348 | ??? |

Pairwise distances

0 1 2 3 4 5 6 7 8

0 2.006 0.739 2.339 0.941 0.596 2.315 1.827 1.896 1.578

1 1.876 0.314 2.472 0.586 0.370 2.150 1.465 1.494 1.164

2 0.225 1.848 1.255 2.150 2.071 0.564 1.193 1.481 1.468

3 1.044 2.003 2.396 2.211 2.329 0.907 0.635 0.813 1.075

4 2.684 3.291 1.474 3.559 3.239 2.948 3.562 3.807 3.638

5 1.605 3.156 2.688 3.377 3.472 1.274 1.806 1.965 2.246

*Top 3 KNN distances for*

| **Sepal length** | **Sepal width** | **Class** |
| --- | --- | --- |
| **9** | -1.232212 | 0.273304 | ??? |
| **10** | -0.894105 | 0.566130 | ??? |
| **11** | 0.345621 | -0.605173 | ??? |
| **12** | 1.134537 | 0.273304 | ??? |
| **13** | -1.457617 | -2.362128 | ??? |
| **14** | 2.148858 | -0.312348 | ??? |

are

**0 1 2 CLASS**

**0 (0.596, setosa) (0.739, setosa) (0.941, setosa) setosa**

**1 (0.314, setosa) (0.37, setosa) (0.586, setosa) setosa**

**2 (0.225, virginica) (0.564, versicolor) (1.193, versicolor) versicolor**

**3 (0.635, versicolor) (0.813, versicolor) (0.907, versicolor) versicolor**

**4 (1.474, virginica) (2.684, virginica) (2.948, versicolor) virginica**

**5 (1.274, versicolor) (1.605, virginica) (1.806, versicolor) versicolor**

After taking majority votes,

The final classes are indicated above in column CLASS.

['setosa', 'setosa', 'versicolor', 'versicolor', 'virginica', 'versicolor']