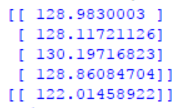
First Project Using Real Data - Blackbird Data

Dataset:

|  |  |  |  |
| --- | --- | --- | --- |
| ID | Sex | Wingspan | Weight  Upon analysis it is noticeable that the 14th female blackbird is missing a value for the wingspan. This will be the focus of the project- to generate a value for the wingspan and an inputation for the dataset.  x - sex, weight  y - wingspan  Organising data using stratified sampling on males and females and simple sampling on each:  Training - 80% male and female  Testing - rest  Application - ID 14  Training data: 1, 2, 3, 4, 5, 6, 9, 11, 12, 13, 15, 16, 17, 18, 19, 20  Testing data: 7, 8, 10, 21  Application data: 14  Males denoted by 1, females denoted by 0.  Min - Max Normalization:  x2: (x-88)/(126-88)  y: (x-122)/(135-122)  Expansion:  x = z \* (135-122) + 122 |
| 1 | M | 135 | 101 |
| 2 | M | 129 | 100 |
| 3 | M | 129 | 98 |
| 4 | M | 129 | 97 |
| 5 | M | 132 | 107 |
| 6 | F | 123 | 94 |
| 7 | F | 124 | 104 |
| 8 | F | 122 | 99  M:  6/8 datum will be chosen  5, 3, 2, 4, 8, 6  ID 2, 3, 4, 5, 9, 11 |
| 9 | M | 129 | 102 |
| 10 | M | 127 | 106 |
| 11 | M | 128 | 106 |
| 12 | F | 128 | 96 |
| 13 | F | 128 | 96 |
| 14 | F |  | 88 |
| 15 | F | 128 | 100 |
| 16 | F | 129 | 105 |
| 17 | F | 127 | 97 |
| 18 | F | 130 | 126 |
| 19 | F | 129 | 104 |
| 20 | F | 130 | 100 |
| 21 | F | 130 | 103 |

Training data results

Application data result; wingspan of 122 for case ID 14



F:

10/12 datum will be chosen

11, 9, 5, 4, 1, 12, 10, 6, 7, 8

ID 20, 18, 13, 12, 6, 21, 19, 15, 16, 17