**Title:** In this project we used a dataset of cereals along with their different percentage of calories, protein, fat etc.

**Description:** This dataset has different 77 cereals with vitamin. After analyzing it helps to select the nutritional food for health requirements. The people who have different diseases like diabetes, high blood pressure, low blood pressure can choose food for their breakfast according to the elements they need and avoid those elements which contains less amount. This will help to choose all the necessary food elements that they need.

**Literature Survey:**

(With 'vitamin' and 'rating' columns): 77 x 11 here: http://www.cs.umd.edu/hcil/hce/examples/cereal/cereal-updated.txt

The meaning of each column:

1. 1st column: Name of cereal

2. Calories: calories per serving

3. Protein: grams of protein

4. Fat: grams of fat

5. Sodium: milligrams of sodium

6. Fiber: grams of dietary fiber

7. Carbo: grams of complex carbohydrates

8. Sugars: grams of sugars

9. Potass: milligrams of potassium

10. Vitamins: vitamins and minerals - 0, 25, or 100, indicating the typical percentage of FDA recommended

11. Shelf: display shelf (1, 2, or 3, counting from the floor)

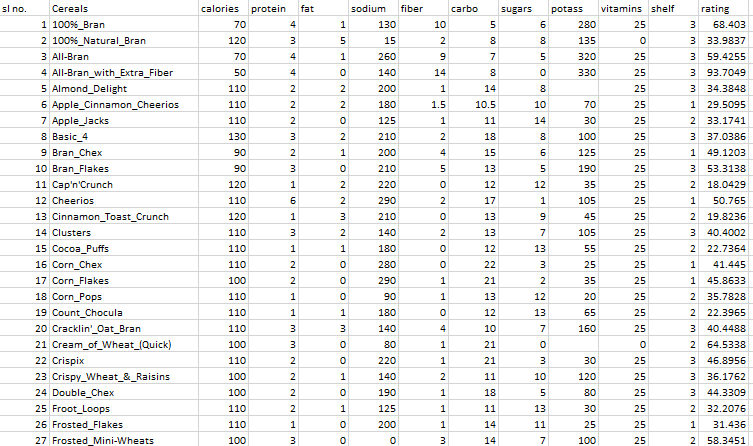
12. Rating: a rating of the cereals (calculated by Consumer Reports)

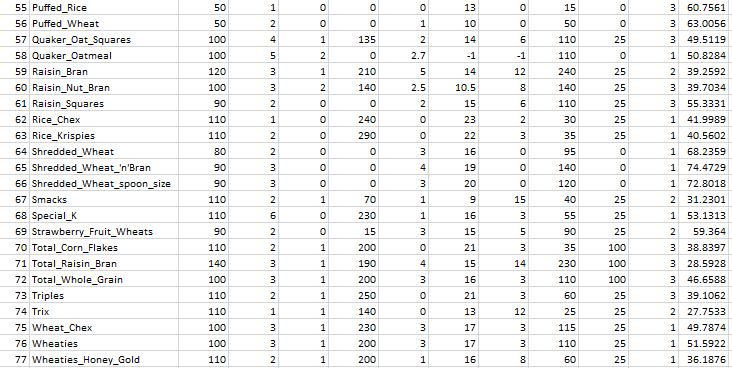
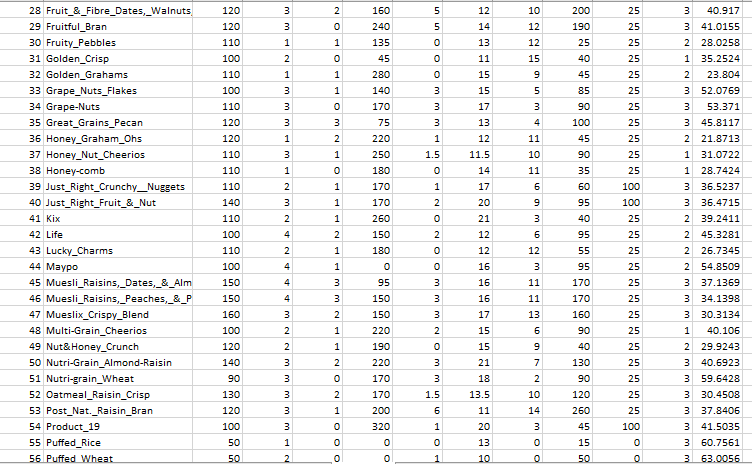
**Methodology:**

The dataset has been clustered by the hierarchical clustering technique. The cluster tree has been cut in several places. Then similarities between instances of individual clusters and dissimilarities between instances of different clusters have been analyzed.

**Dataset:**

The following dataset was used for hierarchical cluster tree with cutting point:





**Hierarchical cluster tree with cutting point:**

=== Run information ===

Scheme: weka.clusterers.HierarchicalClusterer -N 2 -L SINGLE -P -A "weka.core.EuclideanDistance -R first-last"

Relation: cerealWithVitamin

Instances: 75

Attributes: 12

cerealName

cal

protein

fat

sodium

fiber

carbo

suger

potassium

vitamin

shelf

rating

Test mode: evaluate on training data

=== Clustering model (full training set) ===

Cluster 0

(((((1:0.45956,3:0.45956):0.21047,4:0.67003):0.00706,(((2:0.63675,(((((((((5:0.34946,((8:0.24314,(46:0.21537,75:0.21537):0.02777):0.07118,(73:0.09803,74:0.09803):0.2163):0.03513):0.01125,35:0.36071):0.0113,(24:0.10622,36:0.10622):0.26578):0.11218,(((15:0.04606,61:0.04606):0.10861,16:0.15467):0.09347,60:0.24814):0.23605):0.04578,((((((((6:0.20993,23:0.20993):0.01771,(((14:0.03206,18:0.03206):0.15431,(28:0.01603,72:0.01603):0.17034):0.02973,41:0.2161):0.01153):0.0032,65:0.23084):0.02228,47:0.25311):0.02191,17:0.27503):0.00006,((10:0.11214,34:0.11214):0.14118,12:0.25332):0.02177):0.07193,30:0.34702):0.17681,((((((((((((7:0.21278,48:0.21278):0.06881,50:0.28159):0.03189,((13:0.19164,58:0.19164):0.06953,19:0.26118):0.05231):0.00829,21:0.32177):0.00038,((31:0.24174,55:0.24174):0.0414,(32:0.21347,49:0.21347):0.06966):0.03902):0.00223,((20:0.27122,71:0.27122):0.002,22:0.27322):0.05116):0.01553,26:0.33991):0.02772,33:0.36763):0.03526,((43:0.17638,44:0.17638):0.15406,45:0.33044):0.07245):0.00072,(27:0.37831,51:0.37831):0.02529):0.05821,9:0.46182):0.04317,39:0.50499):0.01885):0.00614):0.01013,57:0.5401):0.00472,(((25:0.2627,67:0.2627):0.18199,42:0.44469):0.07307,59:0.51776):0.02706):0.00696,40:0.55179):0.0068,29:0.55859):0.07816):0.01932,(53:0.27195,54:0.27195):0.38412):0.0045,(62:0.29323,(63:0.10657,64:0.10657):0.18666):0.36734):0.01652):0.00506,(11:0.49375,66:0.49375):0.18839):0.08807,(((((37:0.28993,68:0.28993):0.10692,70:0.39685):0.03047,38:0.42733):0.0605,52:0.48783):0.11208,69:0.59991):0.1703)

Time taken to build model (full training data) : 0.01 seconds

=== Model and evaluation on training set ===

Clustered Instances

0 74 ( 99%)

1 1 ( 1%)

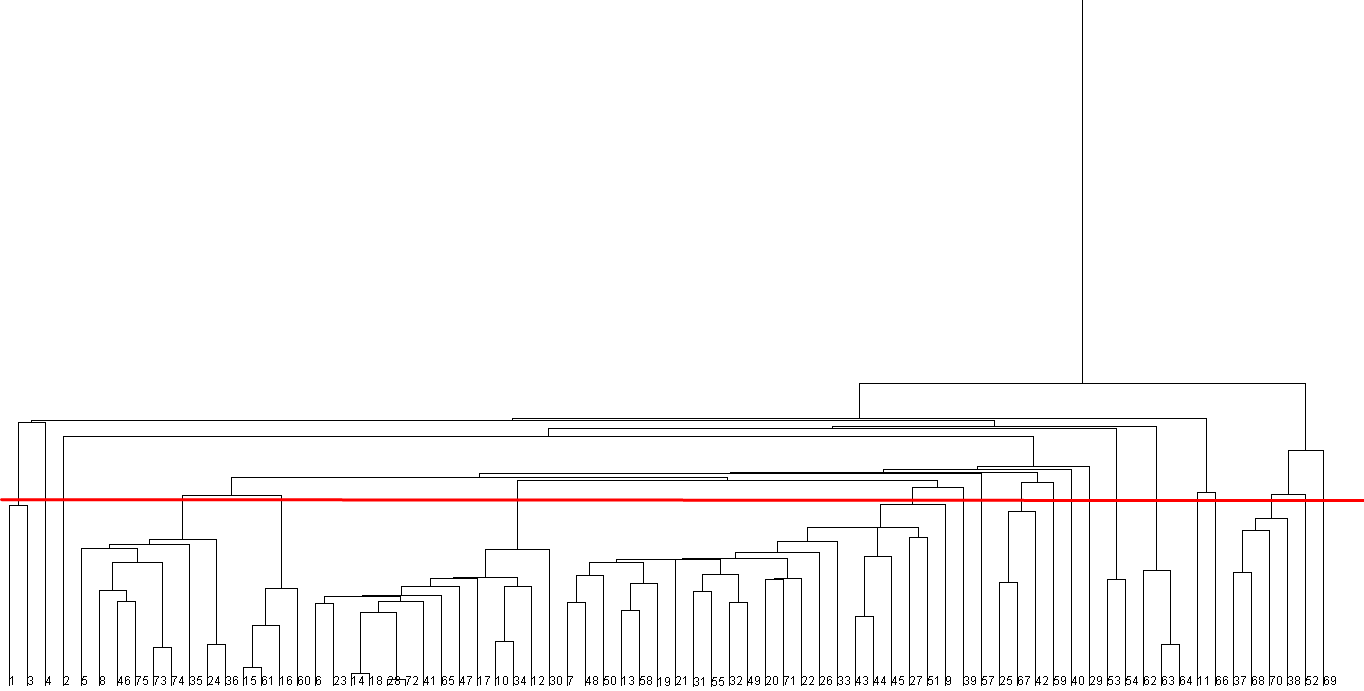


Fig 1: Hierarchical cluster tree with cutting point

**Cluster Analysis:**

**For cluster-1 (2 instances):**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Cereal  Name | calories | protein | fat | sodium | fiber | carbo | sugars | potass | vitamins | Rating |
| 100%\_Bran | 70 | 4 | 1 | 130 | 10 | 5 | 6 | 280 | 25 | 68.40297 |
| All-Bran | 70 | 4 | 1 | 260 | 9 | 7 | 5 | 320 | 25 | 59.42551 |

**Findings:** High potass, low fat, High fiber

**For cluster-2 (1 instances):**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Cereal  Name | calories | protein | fat | sodium | fiber | carbo | sugars | potass | vitamins | Rating |
| All-Bran\_with\_Extra\_Fiber | 50 | 4 | 0 | 140 | 14 | 8 | 0 | 330 | 25 | 93.70491 |

**Findings:** 0 sugars, 0 fat, high potass, high rating, highest fiber

**For cluster-3 (1 instances):**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Cereal  Name | calories | protein | fat | sodium | fiber | carbo | sugars | potass | vitamins | Rating |
| 100%\_Natural\_Bran | 120 | 3 | 5 | 15 | 2 | 8 | 8 | 135 | 0 | 33.98368 |

**Findings:** 0 vitamin, low sodium, high fat

**For cluster-4 (9 instances):**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Cereal  Name | calories | protein | fat | sodium | fiber | carbo | sugars | potass | vitamins | Rating |
| Almond\_Delight | 110 | 2 | 2 | 200 | 1 | 14 | 8 | 25 | 25 | 34.38484 |
| Basic\_4 | 130 | 3 | 2 | 210 | 2 | 18 | 8 | 100 | 25 | 37.03856 |
| Double\_Chex | 100 | 2 | 0 | 190 | 1 | 18 | 5 | 80 | 25 | 44.33086 |
| Great\_Grains\_Pecan | 120 | 3 | 3 | 75 | 3 | 13 | 4 | 100 | 25 | 45.81172 |
| Honey\_Graham\_Ohs | 120 | 1 | 2 | 220 | 1 | 12 | 11 | 45 | 25 | 21.87129 |
| Muesli\_Raisins,\_Peaches,\_&\_Pecans | 150 | 4 | 3 | 150 | 3 | 16 | 11 | 170 | 25 | 34.13977 |
| Triples | 110 | 2 | 1 | 250 | 0 | 21 | 3 | 60 | 25 | 39.10617 |
| Trix | 110 | 1 | 1 | 140 | 0 | 13 | 12 | 25 | 25 | 27.7533 |
| Wheat\_Chex | 100 | 3 | 1 | 230 | 3 | 17 | 3 | 115 | 25 | 49.78745 |

**Findings:** low fat, high sodium

**For cluster-5 (4 instances):**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Cereal  Name | calories | protein | fat | sodium | fiber | carbo | sugars | potass | vitamins | Rating |
| Cocoa\_Puffs | 110 | 1 | 1 | 180 | 0 | 12 | 13 | 55 | 25 | 22.73645 |
| Corn\_Chex | 110 | 2 | 0 | 280 | 0 | 22 | 3 | 25 | 25 | 41.44502 |
| Raisin\_Squares | 90 | 2 | 0 | 0 | 2 | 15 | 6 | 110 | 25 | 55.33314 |
| Raisin\_Nut\_Bran | 100 | 3 | 2 | 140 | 2.5 | 10.5 | 8 | 140 | 25 | 39.7034 |

**Findings:** Average rating, high calories

**For cluster-6 (14 instances):**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Cereal  Name | calories | protein | fat | sodium | fiber | carbo | sugars | potass | vitamins | Rating |
| Apple\_Cinnamon\_Cheerios | 110 | 2 | 2 | 180 | 1.5 | 10.5 | 10 | 70 | 25 | 29.50954 |
| Crispy\_Wheat\_&\_Raisins | 100 | 2 | 1 | 140 | 2 | 11 | 10 | 120 | 25 | 36.1762 |
| Clusters | 110 | 3 | 2 | 140 | 2 | 13 | 7 | 105 | 25 | 40.40021 |
| Corn\_Pops | 110 | 1 | 0 | 90 | 1 | 13 | 12 | 20 | 25 | 35.78279 |
| Fruit\_&\_Fibre\_Dates,\_Walnuts,\_and\_Oats | 120 | 3 | 2 | 160 | 5 | 12 | 10 | 200 | 25 | 40.91705 |
| Total\_Whole\_Grain | 100 | 3 | 1 | 200 | 3 | 16 | 3 | 110 | 100 | 46.65884 |
| Kix | 110 | 2 | 1 | 260 | 0 | 21 | 3 | 40 | 25 | 39.24111 |
| Shredded\_Wheat\_'n'Bran | 90 | 3 | 0 | 0 | 4 | 19 | 0 | 140 | 0 | 74.47295 |
| Mueslix\_Crispy\_Blend | 160 | 3 | 2 | 150 | 3 | 17 | 13 | 160 | 25 | 30.31335 |
| Corn\_Flakes | 100 | 2 | 0 | 290 | 1 | 21 | 2 | 35 | 25 | 45.86332 |
| Bran\_Flakes | 90 | 3 | 0 | 210 | 5 | 13 | 5 | 190 | 25 | 53.31381 |
| Grape-Nuts | 110 | 3 | 0 | 170 | 3 | 17 | 3 | 90 | 25 | 53.37101 |
| Cheerios | 110 | 6 | 2 | 290 | 2 | 17 | 1 | 105 | 25 | 50.765 |
| Fruity\_Pebbles | 110 | 1 | 1 | 135 | 0 | 13 | 12 | 25 | 25 | 28.02577 |

**Findings:** Low fat, high calories & high sodium

**For cluster-7 (22 instances):**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Cereal  Name | calories | protein | fat | sodium | fiber | carbo | sugars | potass | vitamins | Rating |
| Apple\_Jacks | 110 | 2 | 0 | 125 | 1 | 11 | 14 | 30 | 25 | 33.17409 |
| Multi-Grain\_Cheerios | 100 | 2 | 1 | 220 | 2 | 15 | 6 | 90 | 25 | 40.10597 |
| Nutri-Grain\_Almond-Raisin | 140 | 3 | 2 | 220 | 3 | 21 | 7 | 130 | 25 | 40.69232 |
| Cinnamon\_Toast\_Crunch | 120 | 1 | 3 | 210 | 0 | 13 | 9 | 45 | 25 | 19.82357 |
| Quaker\_Oatmeal | 100 | 5 | 2 | 0 | 2.7 | -1 | -1 | 110 | 25 | 50.82839 |
| Count\_Chocula | 110 | 1 | 1 | 180 | 0 | 12 | 13 | 65 | 25 | 22.39651 |
| Cream\_of\_Wheat\_(Quick) | 100 | 3 | 0 | 80 | 1 | 21 | 0 | 0 | 25 | 64.53382 |
| Golden\_Crisp | 100 | 2 | 0 | 45 | 0 | 11 | 15 | 40 | 25 | 35.25244 |
| Puffed\_Rice | 50 | 1 | 0 | 0 | 0 | 13 | 0 | 15 | 25 | 60.75611 |
| Golden\_Grahams | 110 | 1 | 1 | 280 | 0 | 15 | 9 | 45 | 25 | 23.80404 |
| Nut&Honey\_Crunch | 120 | 2 | 1 | 190 | 0 | 15 | 9 | 40 | 25 | 29.92429 |
| Cracklin'\_Oat\_Bran | 110 | 3 | 3 | 140 | 4 | 10 | 7 | 160 | 25 | 40.44877 |
| Total\_Raisin\_Bran | 140 | 3 | 1 | 190 | 4 | 15 | 14 | 230 | 100 | 28.59279 |
| Crispix | 110 | 2 | 0 | 220 | 1 | 21 | 3 | 30 | 25 | 46.89564 |
| Frosted\_Flakes | 110 | 1 | 0 | 200 | 1 | 14 | 11 | 25 | 25 | 31.43597 |
| Grape\_Nuts\_Flakes | 100 | 3 | 1 | 140 | 3 | 15 | 5 | 85 | 25 | 52.0769 |
| Lucky\_Charms | 110 | 2 | 1 | 180 | 0 | 12 | 12 | 55 | 25 | 26.73452 |
| Maypo | 100 | 4 | 1 | 0 | 0 | 16 | 3 | 95 | 25 | 54.85092 |
| Muesli\_Raisins,\_Dates,\_&\_Almonds | 150 | 4 | 3 | 95 | 3 | 16 | 11 | 170 | 25 | 37.13686 |
| Frosted\_Mini-Wheats | 100 | 3 | 0 | 0 | 3 | 14 | 7 | 100 | 25 | 58.34514 |
| Nutri-grain\_Wheat | 90 | 3 | 0 | 170 | 3 | 18 | 2 | 90 | 25 | 59.64284 |
| Bran\_Chex | 90 | 2 | 1 | 200 | 4 | 15 | 6 | 125 | 25 | 49.12025 |

**Findings:** In this cluster, all the food ingredients are available on the cereals.

**For cluster-8 (1 instances):**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Cereal  Name | calories | protein | fat | sodium | fiber | carbo | sugars | potass | vitamins | Rating |
| Just\_Right\_Crunchy\_\_Nuggets | 110 | 2 | 1 | 170 | 1 | 17 | 6 | 60 | 100 | 36.52368 |

**Findings:** High calories, Average rating, high vitamin, low fat, low protein

**For cluster-9 (1 instances):**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Cereal  Name | calories | protein | fat | sodium | fiber | carbo | sugars | potass | vitamins | Rating |
| Quaker\_Oat\_Squares | 100 | 4 | 1 | 135 | 2 | 14 | 6 | 110 | 25 | 49.51187 |

**Findings:** High calories, Average rating, low fat, low fiber, high protein

**For cluster-10 (3 instances):**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Cereal  Name | calories | protein | fat | sodium | fiber | carbo | sugars | potass | vitamins | Rating |
| Smacks | 110 | 2 | 1 | 70 | 1 | 9 | 15 | 40 | 25 | 31.23005 |
| Life | 100 | 4 | 2 | 150 | 2 | 12 | 6 | 95 | 25 | 45.32807 |
| Froot\_Loops | 110 | 2 | 1 | 125 | 1 | 11 | 13 | 30 | 25 | 32.20758 |

**Findings:** Average rating, high calories, low fiber

**For cluster-11 (1 instances):**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Cereal  Name | calories | protein | fat | sodium | fiber | carbo | sugars | potass | vitamins | Rating |
| Raisin\_Bran | 120 | 3 | 1 | 210 | 5 | 14 | 12 | 240 | 25 | 39.2592 |

**Findings:** Average rating, High sodium, high calories, high potass

**For cluster-12 (1 instances):**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Cereal  Name | calories | protein | fat | sodium | fiber | carbo | sugars | potass | vitamins | Rating |
| Just\_Right\_Fruit\_&\_Nut | 140 | 3 | 1 | 170 | 2 | 20 | 9 | 95 | 100 | 36.47151 |

**Findings:** Average rating, high calories, low fat, high potass

**For cluster-13 (1 instances):**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Cereal  Name | calories | protein | fat | sodium | fiber | carbo | sugars | potass | vitamins | Rating |
| Fruitful\_Bran | 120 | 3 | 0 | 240 | 5 | 14 | 12 | 190 | 25 | 41.01549 |

**Findings:** high potass,0 fat, high calories, high sugar

**For cluster-14 (2 instances):**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Cereal  Name | calories | protein | fat | sodium | fiber | carbo | sugars | potass | vitamins | Rating |
| Post\_Nat.\_Raisin\_Bran | 120 | 3 | 1 | 200 | 6 | 11 | 14 | 260 | 25 | 37.84059 |
| Product\_19 | 100 | 3 | 0 | 320 | 1 | 20 | 3 | 45 | 100 | 41.50354 |

**Findings:** High calories, low fat, high sodium, High carbon

**For cluster-15 (3 instances):**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Cereal  Name | calories | protein | fat | sodium | fiber | carbo | sugars | potass | vitamins | Rating |
| Rice\_Chex | 110 | 1 | 0 | 240 | 0 | 23 | 2 | 30 | 25 | 41.99893 |
| Rice\_Krispies | 110 | 2 | 0 | 290 | 0 | 22 | 3 | 35 | 25 | 40.56016 |
| Shredded\_Wheat | 80 | 2 | 0 | 0 | 3 | 16 | 0 | 95 | 0 | 68.23589 |

**Findings:** 0 fat, High calories, high carbon, low sugar

**For cluster-16 (1 instances):**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Cereal  Name | calories | protein | fat | sodium | fiber | carbo | sugars | potass | vitamins | Rating |
| Cap'n'Crunch | 120 | 1 | 2 | 220 | 0 | 12 | 12 | 35 | 25 | 18.04285 |

**Findings:** low fat, 0 fiber, high sodium, high sugar, low protein, high calories

**For cluster-17 (1 instances):**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Cereal  Name | calories | protein | fat | sodium | fiber | carbo | sugars | potass | vitamins | Rating |
| Shredded\_Wheat\_spoon\_size | 90 | 3 | 0 | 0 | 3 | 20 | 0 | 120 | 0 | 72.80179 |

**Findings:** 0 fat, 0 sodium, high carbon, 0 sugar

**For cluster-18 (4 instances):**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Cereal  Name | calories | protein | fat | sodium | fiber | carbo | sugars | potass | vitamins | Rating |
| Honey\_Nut\_Cheerios | 110 | 3 | 1 | 250 | 1.5 | 11.5 | 10 | 90 | 25 | 31.07222 |
| Honey-comb | 110 | 1 | 0 | 180 | 0 | 14 | 11 | 35 | 25 | 28.74241 |
| Special\_K | 110 | 6 | 0 | 230 | 1 | 16 | 3 | 55 | 25 | 53.13132 |
| Total\_Corn\_Flakes | 110 | 2 | 1 | 200 | 0 | 21 | 3 | 35 | 100 | 38.83975 |

**Findings:** High sodium, low fat, high calories

**For cluster-19 (1 instances):**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Cereal  Name | calories | protein | fat | sodium | fiber | carbo | sugars | potass | vitamins | Rating |
| Oatmeal\_Raisin\_Crisp | 130 | 3 | 2 | 170 | 1.5 | 13.5 | 10 | 120 | 25 | 30.45084 |

**Findings:** high sodium, high calories, low fat, low fiber, high potass

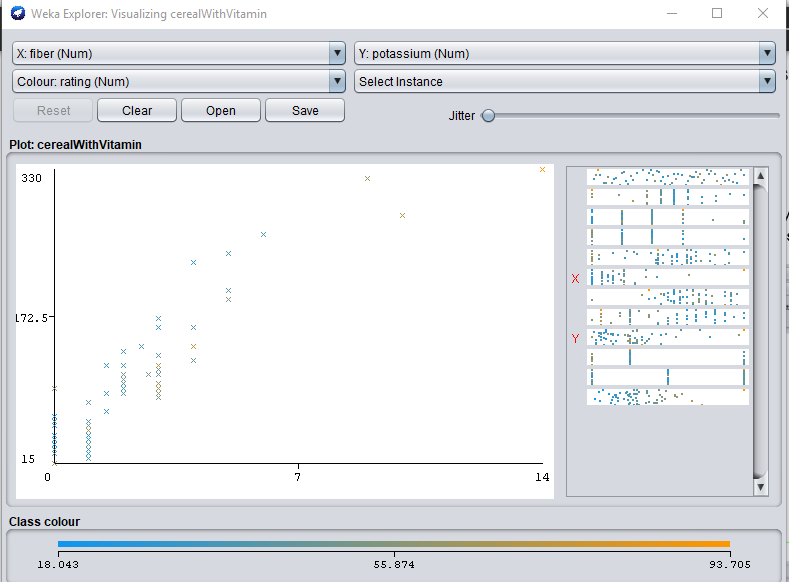
**For cluster-20 (1 instances):**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Cereal  Name | calories | protein | fat | sodium | fiber | carbo | sugars | potass | vitamins | Rating |
| Strawberry\_Fruit\_Wheats | 90 | 2 | 0 | 15 | 3 | 15 | 5 | 90 | 25 | 59.36399 |

**Findings:** 0 fat, low sugar, low sodium, high carbon

**Questions & Answers**

1. Yes, there is correlation between dietary fiber and potassium and they are linearly proportional. If the value of dietary fiber rises then the value of potassium also rises.



1. Yes, they are in groups so one can choose according to their own preferences.

* The diabetes patients can choose cereal from cluster -1
* Anyone looking for all food values he/she can choose a cereal from cluster two.
* Those who want high sodium, high potassium, high fiber can pick a cereal from cluster-11.
* Those who have both blood pressure and diabetics can chose from cluster-20
* Anyone looking for high sodium, high calories, low fat, low fiber, high potass can chose cluster-19
* Anyone looking for high sodium, low fat, high calories can choose cluster-18
* Anyone looking for high calories, high carbon, low sugar can choose cluster-15

1. There is correlation between Calories and Carbohydrates. If the value of Calories goes higher then there is a possibility of higher carbohydrate value. This confirms low carbohydrate value for low calories.

