Assignment 5: Hierarchical Clustering

Jared Bartee

2023-11-29

# Step 1: Data Preprocessing

# Load the data

cereals\_data <- read.csv("Cereals.csv")  
head(cereals\_data)

## name mfr type calories protein fat sodium fiber carbo  
## 1 100%\_Bran N C 70 4 1 130 10.0 5.0  
## 2 100%\_Natural\_Bran Q C 120 3 5 15 2.0 8.0  
## 3 All-Bran K C 70 4 1 260 9.0 7.0  
## 4 All-Bran\_with\_Extra\_Fiber K C 50 4 0 140 14.0 8.0  
## 5 Almond\_Delight R C 110 2 2 200 1.0 14.0  
## 6 Apple\_Cinnamon\_Cheerios G C 110 2 2 180 1.5 10.5  
## sugars potass vitamins shelf weight cups rating  
## 1 6 280 25 3 1 0.33 68.40297  
## 2 8 135 0 3 1 1.00 33.98368  
## 3 5 320 25 3 1 0.33 59.42551  
## 4 0 330 25 3 1 0.50 93.70491  
## 5 8 NA 25 3 1 0.75 34.38484  
## 6 10 70 25 1 1 0.75 29.50954

# Drop mfr & type columns

cereals\_data <- cereals\_data[-c(2,3)]  
cereals\_data

## name calories protein fat sodium fiber  
## 1 100%\_Bran 70 4 1 130 10.0  
## 2 100%\_Natural\_Bran 120 3 5 15 2.0  
## 3 All-Bran 70 4 1 260 9.0  
## 4 All-Bran\_with\_Extra\_Fiber 50 4 0 140 14.0  
## 5 Almond\_Delight 110 2 2 200 1.0  
## 6 Apple\_Cinnamon\_Cheerios 110 2 2 180 1.5  
## 7 Apple\_Jacks 110 2 0 125 1.0  
## 8 Basic\_4 130 3 2 210 2.0  
## 9 Bran\_Chex 90 2 1 200 4.0  
## 10 Bran\_Flakes 90 3 0 210 5.0  
## 11 Cap'n'Crunch 120 1 2 220 0.0  
## 12 Cheerios 110 6 2 290 2.0  
## 13 Cinnamon\_Toast\_Crunch 120 1 3 210 0.0  
## 14 Clusters 110 3 2 140 2.0  
## 15 Cocoa\_Puffs 110 1 1 180 0.0  
## 16 Corn\_Chex 110 2 0 280 0.0  
## 17 Corn\_Flakes 100 2 0 290 1.0  
## 18 Corn\_Pops 110 1 0 90 1.0  
## 19 Count\_Chocula 110 1 1 180 0.0  
## 20 Cracklin'\_Oat\_Bran 110 3 3 140 4.0  
## 21 Cream\_of\_Wheat\_(Quick) 100 3 0 80 1.0  
## 22 Crispix 110 2 0 220 1.0  
## 23 Crispy\_Wheat\_&\_Raisins 100 2 1 140 2.0  
## 24 Double\_Chex 100 2 0 190 1.0  
## 25 Froot\_Loops 110 2 1 125 1.0  
## 26 Frosted\_Flakes 110 1 0 200 1.0  
## 27 Frosted\_Mini-Wheats 100 3 0 0 3.0  
## 28 Fruit\_&\_Fibre\_Dates,\_Walnuts,\_and\_Oats 120 3 2 160 5.0  
## 29 Fruitful\_Bran 120 3 0 240 5.0  
## 30 Fruity\_Pebbles 110 1 1 135 0.0  
## 31 Golden\_Crisp 100 2 0 45 0.0  
## 32 Golden\_Grahams 110 1 1 280 0.0  
## 33 Grape\_Nuts\_Flakes 100 3 1 140 3.0  
## 34 Grape-Nuts 110 3 0 170 3.0  
## 35 Great\_Grains\_Pecan 120 3 3 75 3.0  
## 36 Honey\_Graham\_Ohs 120 1 2 220 1.0  
## 37 Honey\_Nut\_Cheerios 110 3 1 250 1.5  
## 38 Honey-comb 110 1 0 180 0.0  
## 39 Just\_Right\_Crunchy\_\_Nuggets 110 2 1 170 1.0  
## 40 Just\_Right\_Fruit\_&\_Nut 140 3 1 170 2.0  
## 41 Kix 110 2 1 260 0.0  
## 42 Life 100 4 2 150 2.0  
## 43 Lucky\_Charms 110 2 1 180 0.0  
## 44 Maypo 100 4 1 0 0.0  
## 45 Muesli\_Raisins,\_Dates,\_&\_Almonds 150 4 3 95 3.0  
## 46 Muesli\_Raisins,\_Peaches,\_&\_Pecans 150 4 3 150 3.0  
## 47 Mueslix\_Crispy\_Blend 160 3 2 150 3.0  
## 48 Multi-Grain\_Cheerios 100 2 1 220 2.0  
## 49 Nut&Honey\_Crunch 120 2 1 190 0.0  
## 50 Nutri-Grain\_Almond-Raisin 140 3 2 220 3.0  
## 51 Nutri-grain\_Wheat 90 3 0 170 3.0  
## 52 Oatmeal\_Raisin\_Crisp 130 3 2 170 1.5  
## 53 Post\_Nat.\_Raisin\_Bran 120 3 1 200 6.0  
## 54 Product\_19 100 3 0 320 1.0  
## 55 Puffed\_Rice 50 1 0 0 0.0  
## 56 Puffed\_Wheat 50 2 0 0 1.0  
## 57 Quaker\_Oat\_Squares 100 4 1 135 2.0  
## 58 Quaker\_Oatmeal 100 5 2 0 2.7  
## 59 Raisin\_Bran 120 3 1 210 5.0  
## 60 Raisin\_Nut\_Bran 100 3 2 140 2.5  
## 61 Raisin\_Squares 90 2 0 0 2.0  
## 62 Rice\_Chex 110 1 0 240 0.0  
## 63 Rice\_Krispies 110 2 0 290 0.0  
## 64 Shredded\_Wheat 80 2 0 0 3.0  
## 65 Shredded\_Wheat\_'n'Bran 90 3 0 0 4.0  
## 66 Shredded\_Wheat\_spoon\_size 90 3 0 0 3.0  
## 67 Smacks 110 2 1 70 1.0  
## 68 Special\_K 110 6 0 230 1.0  
## 69 Strawberry\_Fruit\_Wheats 90 2 0 15 3.0  
## 70 Total\_Corn\_Flakes 110 2 1 200 0.0  
## 71 Total\_Raisin\_Bran 140 3 1 190 4.0  
## 72 Total\_Whole\_Grain 100 3 1 200 3.0  
## 73 Triples 110 2 1 250 0.0  
## 74 Trix 110 1 1 140 0.0  
## 75 Wheat\_Chex 100 3 1 230 3.0  
## 76 Wheaties 100 3 1 200 3.0  
## 77 Wheaties\_Honey\_Gold 110 2 1 200 1.0  
## carbo sugars potass vitamins shelf weight cups rating  
## 1 5.0 6 280 25 3 1.00 0.33 68.40297  
## 2 8.0 8 135 0 3 1.00 1.00 33.98368  
## 3 7.0 5 320 25 3 1.00 0.33 59.42551  
## 4 8.0 0 330 25 3 1.00 0.50 93.70491  
## 5 14.0 8 NA 25 3 1.00 0.75 34.38484  
## 6 10.5 10 70 25 1 1.00 0.75 29.50954  
## 7 11.0 14 30 25 2 1.00 1.00 33.17409  
## 8 18.0 8 100 25 3 1.33 0.75 37.03856  
## 9 15.0 6 125 25 1 1.00 0.67 49.12025  
## 10 13.0 5 190 25 3 1.00 0.67 53.31381  
## 11 12.0 12 35 25 2 1.00 0.75 18.04285  
## 12 17.0 1 105 25 1 1.00 1.25 50.76500  
## 13 13.0 9 45 25 2 1.00 0.75 19.82357  
## 14 13.0 7 105 25 3 1.00 0.50 40.40021  
## 15 12.0 13 55 25 2 1.00 1.00 22.73645  
## 16 22.0 3 25 25 1 1.00 1.00 41.44502  
## 17 21.0 2 35 25 1 1.00 1.00 45.86332  
## 18 13.0 12 20 25 2 1.00 1.00 35.78279  
## 19 12.0 13 65 25 2 1.00 1.00 22.39651  
## 20 10.0 7 160 25 3 1.00 0.50 40.44877  
## 21 21.0 0 NA 0 2 1.00 1.00 64.53382  
## 22 21.0 3 30 25 3 1.00 1.00 46.89564  
## 23 11.0 10 120 25 3 1.00 0.75 36.17620  
## 24 18.0 5 80 25 3 1.00 0.75 44.33086  
## 25 11.0 13 30 25 2 1.00 1.00 32.20758  
## 26 14.0 11 25 25 1 1.00 0.75 31.43597  
## 27 14.0 7 100 25 2 1.00 0.80 58.34514  
## 28 12.0 10 200 25 3 1.25 0.67 40.91705  
## 29 14.0 12 190 25 3 1.33 0.67 41.01549  
## 30 13.0 12 25 25 2 1.00 0.75 28.02576  
## 31 11.0 15 40 25 1 1.00 0.88 35.25244  
## 32 15.0 9 45 25 2 1.00 0.75 23.80404  
## 33 15.0 5 85 25 3 1.00 0.88 52.07690  
## 34 17.0 3 90 25 3 1.00 0.25 53.37101  
## 35 13.0 4 100 25 3 1.00 0.33 45.81172  
## 36 12.0 11 45 25 2 1.00 1.00 21.87129  
## 37 11.5 10 90 25 1 1.00 0.75 31.07222  
## 38 14.0 11 35 25 1 1.00 1.33 28.74241  
## 39 17.0 6 60 100 3 1.00 1.00 36.52368  
## 40 20.0 9 95 100 3 1.30 0.75 36.47151  
## 41 21.0 3 40 25 2 1.00 1.50 39.24111  
## 42 12.0 6 95 25 2 1.00 0.67 45.32807  
## 43 12.0 12 55 25 2 1.00 1.00 26.73451  
## 44 16.0 3 95 25 2 1.00 1.00 54.85092  
## 45 16.0 11 170 25 3 1.00 1.00 37.13686  
## 46 16.0 11 170 25 3 1.00 1.00 34.13976  
## 47 17.0 13 160 25 3 1.50 0.67 30.31335  
## 48 15.0 6 90 25 1 1.00 1.00 40.10596  
## 49 15.0 9 40 25 2 1.00 0.67 29.92429  
## 50 21.0 7 130 25 3 1.33 0.67 40.69232  
## 51 18.0 2 90 25 3 1.00 1.00 59.64284  
## 52 13.5 10 120 25 3 1.25 0.50 30.45084  
## 53 11.0 14 260 25 3 1.33 0.67 37.84059  
## 54 20.0 3 45 100 3 1.00 1.00 41.50354  
## 55 13.0 0 15 0 3 0.50 1.00 60.75611  
## 56 10.0 0 50 0 3 0.50 1.00 63.00565  
## 57 14.0 6 110 25 3 1.00 0.50 49.51187  
## 58 NA NA 110 0 1 1.00 0.67 50.82839  
## 59 14.0 12 240 25 2 1.33 0.75 39.25920  
## 60 10.5 8 140 25 3 1.00 0.50 39.70340  
## 61 15.0 6 110 25 3 1.00 0.50 55.33314  
## 62 23.0 2 30 25 1 1.00 1.13 41.99893  
## 63 22.0 3 35 25 1 1.00 1.00 40.56016  
## 64 16.0 0 95 0 1 0.83 1.00 68.23588  
## 65 19.0 0 140 0 1 1.00 0.67 74.47295  
## 66 20.0 0 120 0 1 1.00 0.67 72.80179  
## 67 9.0 15 40 25 2 1.00 0.75 31.23005  
## 68 16.0 3 55 25 1 1.00 1.00 53.13132  
## 69 15.0 5 90 25 2 1.00 1.00 59.36399  
## 70 21.0 3 35 100 3 1.00 1.00 38.83975  
## 71 15.0 14 230 100 3 1.50 1.00 28.59278  
## 72 16.0 3 110 100 3 1.00 1.00 46.65884  
## 73 21.0 3 60 25 3 1.00 0.75 39.10617  
## 74 13.0 12 25 25 2 1.00 1.00 27.75330  
## 75 17.0 3 115 25 1 1.00 0.67 49.78744  
## 76 17.0 3 110 25 1 1.00 1.00 51.59219  
## 77 16.0 8 60 25 1 1.00 0.75 36.18756

# Convert the names of breakfast cereals to row names

rownames(cereals\_data) <- cereals\_data$name

#Drop the name column

cereals\_data <- cereals\_data[-c(1)]  
cereals\_data

## calories protein fat sodium fiber carbo  
## 100%\_Bran 70 4 1 130 10.0 5.0  
## 100%\_Natural\_Bran 120 3 5 15 2.0 8.0  
## All-Bran 70 4 1 260 9.0 7.0  
## All-Bran\_with\_Extra\_Fiber 50 4 0 140 14.0 8.0  
## Almond\_Delight 110 2 2 200 1.0 14.0  
## Apple\_Cinnamon\_Cheerios 110 2 2 180 1.5 10.5  
## Apple\_Jacks 110 2 0 125 1.0 11.0  
## Basic\_4 130 3 2 210 2.0 18.0  
## Bran\_Chex 90 2 1 200 4.0 15.0  
## Bran\_Flakes 90 3 0 210 5.0 13.0  
## Cap'n'Crunch 120 1 2 220 0.0 12.0  
## Cheerios 110 6 2 290 2.0 17.0  
## Cinnamon\_Toast\_Crunch 120 1 3 210 0.0 13.0  
## Clusters 110 3 2 140 2.0 13.0  
## Cocoa\_Puffs 110 1 1 180 0.0 12.0  
## Corn\_Chex 110 2 0 280 0.0 22.0  
## Corn\_Flakes 100 2 0 290 1.0 21.0  
## Corn\_Pops 110 1 0 90 1.0 13.0  
## Count\_Chocula 110 1 1 180 0.0 12.0  
## Cracklin'\_Oat\_Bran 110 3 3 140 4.0 10.0  
## Cream\_of\_Wheat\_(Quick) 100 3 0 80 1.0 21.0  
## Crispix 110 2 0 220 1.0 21.0  
## Crispy\_Wheat\_&\_Raisins 100 2 1 140 2.0 11.0  
## Double\_Chex 100 2 0 190 1.0 18.0  
## Froot\_Loops 110 2 1 125 1.0 11.0  
## Frosted\_Flakes 110 1 0 200 1.0 14.0  
## Frosted\_Mini-Wheats 100 3 0 0 3.0 14.0  
## Fruit\_&\_Fibre\_Dates,\_Walnuts,\_and\_Oats 120 3 2 160 5.0 12.0  
## Fruitful\_Bran 120 3 0 240 5.0 14.0  
## Fruity\_Pebbles 110 1 1 135 0.0 13.0  
## Golden\_Crisp 100 2 0 45 0.0 11.0  
## Golden\_Grahams 110 1 1 280 0.0 15.0  
## Grape\_Nuts\_Flakes 100 3 1 140 3.0 15.0  
## Grape-Nuts 110 3 0 170 3.0 17.0  
## Great\_Grains\_Pecan 120 3 3 75 3.0 13.0  
## Honey\_Graham\_Ohs 120 1 2 220 1.0 12.0  
## Honey\_Nut\_Cheerios 110 3 1 250 1.5 11.5  
## Honey-comb 110 1 0 180 0.0 14.0  
## Just\_Right\_Crunchy\_\_Nuggets 110 2 1 170 1.0 17.0  
## Just\_Right\_Fruit\_&\_Nut 140 3 1 170 2.0 20.0  
## Kix 110 2 1 260 0.0 21.0  
## Life 100 4 2 150 2.0 12.0  
## Lucky\_Charms 110 2 1 180 0.0 12.0  
## Maypo 100 4 1 0 0.0 16.0  
## Muesli\_Raisins,\_Dates,\_&\_Almonds 150 4 3 95 3.0 16.0  
## Muesli\_Raisins,\_Peaches,\_&\_Pecans 150 4 3 150 3.0 16.0  
## Mueslix\_Crispy\_Blend 160 3 2 150 3.0 17.0  
## Multi-Grain\_Cheerios 100 2 1 220 2.0 15.0  
## Nut&Honey\_Crunch 120 2 1 190 0.0 15.0  
## Nutri-Grain\_Almond-Raisin 140 3 2 220 3.0 21.0  
## Nutri-grain\_Wheat 90 3 0 170 3.0 18.0  
## Oatmeal\_Raisin\_Crisp 130 3 2 170 1.5 13.5  
## Post\_Nat.\_Raisin\_Bran 120 3 1 200 6.0 11.0  
## Product\_19 100 3 0 320 1.0 20.0  
## Puffed\_Rice 50 1 0 0 0.0 13.0  
## Puffed\_Wheat 50 2 0 0 1.0 10.0  
## Quaker\_Oat\_Squares 100 4 1 135 2.0 14.0  
## Quaker\_Oatmeal 100 5 2 0 2.7 NA  
## Raisin\_Bran 120 3 1 210 5.0 14.0  
## Raisin\_Nut\_Bran 100 3 2 140 2.5 10.5  
## Raisin\_Squares 90 2 0 0 2.0 15.0  
## Rice\_Chex 110 1 0 240 0.0 23.0  
## Rice\_Krispies 110 2 0 290 0.0 22.0  
## Shredded\_Wheat 80 2 0 0 3.0 16.0  
## Shredded\_Wheat\_'n'Bran 90 3 0 0 4.0 19.0  
## Shredded\_Wheat\_spoon\_size 90 3 0 0 3.0 20.0  
## Smacks 110 2 1 70 1.0 9.0  
## Special\_K 110 6 0 230 1.0 16.0  
## Strawberry\_Fruit\_Wheats 90 2 0 15 3.0 15.0  
## Total\_Corn\_Flakes 110 2 1 200 0.0 21.0  
## Total\_Raisin\_Bran 140 3 1 190 4.0 15.0  
## Total\_Whole\_Grain 100 3 1 200 3.0 16.0  
## Triples 110 2 1 250 0.0 21.0  
## Trix 110 1 1 140 0.0 13.0  
## Wheat\_Chex 100 3 1 230 3.0 17.0  
## Wheaties 100 3 1 200 3.0 17.0  
## Wheaties\_Honey\_Gold 110 2 1 200 1.0 16.0  
## sugars potass vitamins shelf weight cups  
## 100%\_Bran 6 280 25 3 1.00 0.33  
## 100%\_Natural\_Bran 8 135 0 3 1.00 1.00  
## All-Bran 5 320 25 3 1.00 0.33  
## All-Bran\_with\_Extra\_Fiber 0 330 25 3 1.00 0.50  
## Almond\_Delight 8 NA 25 3 1.00 0.75  
## Apple\_Cinnamon\_Cheerios 10 70 25 1 1.00 0.75  
## Apple\_Jacks 14 30 25 2 1.00 1.00  
## Basic\_4 8 100 25 3 1.33 0.75  
## Bran\_Chex 6 125 25 1 1.00 0.67  
## Bran\_Flakes 5 190 25 3 1.00 0.67  
## Cap'n'Crunch 12 35 25 2 1.00 0.75  
## Cheerios 1 105 25 1 1.00 1.25  
## Cinnamon\_Toast\_Crunch 9 45 25 2 1.00 0.75  
## Clusters 7 105 25 3 1.00 0.50  
## Cocoa\_Puffs 13 55 25 2 1.00 1.00  
## Corn\_Chex 3 25 25 1 1.00 1.00  
## Corn\_Flakes 2 35 25 1 1.00 1.00  
## Corn\_Pops 12 20 25 2 1.00 1.00  
## Count\_Chocula 13 65 25 2 1.00 1.00  
## Cracklin'\_Oat\_Bran 7 160 25 3 1.00 0.50  
## Cream\_of\_Wheat\_(Quick) 0 NA 0 2 1.00 1.00  
## Crispix 3 30 25 3 1.00 1.00  
## Crispy\_Wheat\_&\_Raisins 10 120 25 3 1.00 0.75  
## Double\_Chex 5 80 25 3 1.00 0.75  
## Froot\_Loops 13 30 25 2 1.00 1.00  
## Frosted\_Flakes 11 25 25 1 1.00 0.75  
## Frosted\_Mini-Wheats 7 100 25 2 1.00 0.80  
## Fruit\_&\_Fibre\_Dates,\_Walnuts,\_and\_Oats 10 200 25 3 1.25 0.67  
## Fruitful\_Bran 12 190 25 3 1.33 0.67  
## Fruity\_Pebbles 12 25 25 2 1.00 0.75  
## Golden\_Crisp 15 40 25 1 1.00 0.88  
## Golden\_Grahams 9 45 25 2 1.00 0.75  
## Grape\_Nuts\_Flakes 5 85 25 3 1.00 0.88  
## Grape-Nuts 3 90 25 3 1.00 0.25  
## Great\_Grains\_Pecan 4 100 25 3 1.00 0.33  
## Honey\_Graham\_Ohs 11 45 25 2 1.00 1.00  
## Honey\_Nut\_Cheerios 10 90 25 1 1.00 0.75  
## Honey-comb 11 35 25 1 1.00 1.33  
## Just\_Right\_Crunchy\_\_Nuggets 6 60 100 3 1.00 1.00  
## Just\_Right\_Fruit\_&\_Nut 9 95 100 3 1.30 0.75  
## Kix 3 40 25 2 1.00 1.50  
## Life 6 95 25 2 1.00 0.67  
## Lucky\_Charms 12 55 25 2 1.00 1.00  
## Maypo 3 95 25 2 1.00 1.00  
## Muesli\_Raisins,\_Dates,\_&\_Almonds 11 170 25 3 1.00 1.00  
## Muesli\_Raisins,\_Peaches,\_&\_Pecans 11 170 25 3 1.00 1.00  
## Mueslix\_Crispy\_Blend 13 160 25 3 1.50 0.67  
## Multi-Grain\_Cheerios 6 90 25 1 1.00 1.00  
## Nut&Honey\_Crunch 9 40 25 2 1.00 0.67  
## Nutri-Grain\_Almond-Raisin 7 130 25 3 1.33 0.67  
## Nutri-grain\_Wheat 2 90 25 3 1.00 1.00  
## Oatmeal\_Raisin\_Crisp 10 120 25 3 1.25 0.50  
## Post\_Nat.\_Raisin\_Bran 14 260 25 3 1.33 0.67  
## Product\_19 3 45 100 3 1.00 1.00  
## Puffed\_Rice 0 15 0 3 0.50 1.00  
## Puffed\_Wheat 0 50 0 3 0.50 1.00  
## Quaker\_Oat\_Squares 6 110 25 3 1.00 0.50  
## Quaker\_Oatmeal NA 110 0 1 1.00 0.67  
## Raisin\_Bran 12 240 25 2 1.33 0.75  
## Raisin\_Nut\_Bran 8 140 25 3 1.00 0.50  
## Raisin\_Squares 6 110 25 3 1.00 0.50  
## Rice\_Chex 2 30 25 1 1.00 1.13  
## Rice\_Krispies 3 35 25 1 1.00 1.00  
## Shredded\_Wheat 0 95 0 1 0.83 1.00  
## Shredded\_Wheat\_'n'Bran 0 140 0 1 1.00 0.67  
## Shredded\_Wheat\_spoon\_size 0 120 0 1 1.00 0.67  
## Smacks 15 40 25 2 1.00 0.75  
## Special\_K 3 55 25 1 1.00 1.00  
## Strawberry\_Fruit\_Wheats 5 90 25 2 1.00 1.00  
## Total\_Corn\_Flakes 3 35 100 3 1.00 1.00  
## Total\_Raisin\_Bran 14 230 100 3 1.50 1.00  
## Total\_Whole\_Grain 3 110 100 3 1.00 1.00  
## Triples 3 60 25 3 1.00 0.75  
## Trix 12 25 25 2 1.00 1.00  
## Wheat\_Chex 3 115 25 1 1.00 0.67  
## Wheaties 3 110 25 1 1.00 1.00  
## Wheaties\_Honey\_Gold 8 60 25 1 1.00 0.75  
## rating  
## 100%\_Bran 68.40297  
## 100%\_Natural\_Bran 33.98368  
## All-Bran 59.42551  
## All-Bran\_with\_Extra\_Fiber 93.70491  
## Almond\_Delight 34.38484  
## Apple\_Cinnamon\_Cheerios 29.50954  
## Apple\_Jacks 33.17409  
## Basic\_4 37.03856  
## Bran\_Chex 49.12025  
## Bran\_Flakes 53.31381  
## Cap'n'Crunch 18.04285  
## Cheerios 50.76500  
## Cinnamon\_Toast\_Crunch 19.82357  
## Clusters 40.40021  
## Cocoa\_Puffs 22.73645  
## Corn\_Chex 41.44502  
## Corn\_Flakes 45.86332  
## Corn\_Pops 35.78279  
## Count\_Chocula 22.39651  
## Cracklin'\_Oat\_Bran 40.44877  
## Cream\_of\_Wheat\_(Quick) 64.53382  
## Crispix 46.89564  
## Crispy\_Wheat\_&\_Raisins 36.17620  
## Double\_Chex 44.33086  
## Froot\_Loops 32.20758  
## Frosted\_Flakes 31.43597  
## Frosted\_Mini-Wheats 58.34514  
## Fruit\_&\_Fibre\_Dates,\_Walnuts,\_and\_Oats 40.91705  
## Fruitful\_Bran 41.01549  
## Fruity\_Pebbles 28.02576  
## Golden\_Crisp 35.25244  
## Golden\_Grahams 23.80404  
## Grape\_Nuts\_Flakes 52.07690  
## Grape-Nuts 53.37101  
## Great\_Grains\_Pecan 45.81172  
## Honey\_Graham\_Ohs 21.87129  
## Honey\_Nut\_Cheerios 31.07222  
## Honey-comb 28.74241  
## Just\_Right\_Crunchy\_\_Nuggets 36.52368  
## Just\_Right\_Fruit\_&\_Nut 36.47151  
## Kix 39.24111  
## Life 45.32807  
## Lucky\_Charms 26.73451  
## Maypo 54.85092  
## Muesli\_Raisins,\_Dates,\_&\_Almonds 37.13686  
## Muesli\_Raisins,\_Peaches,\_&\_Pecans 34.13976  
## Mueslix\_Crispy\_Blend 30.31335  
## Multi-Grain\_Cheerios 40.10596  
## Nut&Honey\_Crunch 29.92429  
## Nutri-Grain\_Almond-Raisin 40.69232  
## Nutri-grain\_Wheat 59.64284  
## Oatmeal\_Raisin\_Crisp 30.45084  
## Post\_Nat.\_Raisin\_Bran 37.84059  
## Product\_19 41.50354  
## Puffed\_Rice 60.75611  
## Puffed\_Wheat 63.00565  
## Quaker\_Oat\_Squares 49.51187  
## Quaker\_Oatmeal 50.82839  
## Raisin\_Bran 39.25920  
## Raisin\_Nut\_Bran 39.70340  
## Raisin\_Squares 55.33314  
## Rice\_Chex 41.99893  
## Rice\_Krispies 40.56016  
## Shredded\_Wheat 68.23588  
## Shredded\_Wheat\_'n'Bran 74.47295  
## Shredded\_Wheat\_spoon\_size 72.80179  
## Smacks 31.23005  
## Special\_K 53.13132  
## Strawberry\_Fruit\_Wheats 59.36399  
## Total\_Corn\_Flakes 38.83975  
## Total\_Raisin\_Bran 28.59278  
## Total\_Whole\_Grain 46.65884  
## Triples 39.10617  
## Trix 27.75330  
## Wheat\_Chex 49.78744  
## Wheaties 51.59219  
## Wheaties\_Honey\_Gold 36.18756

# Find the amount of missing values and then omit them

sum(is.na(cereals\_data))

## [1] 4

cereals\_data <- na.omit(cereals\_data)  
sum(is.na(cereals\_data))

## [1] 0

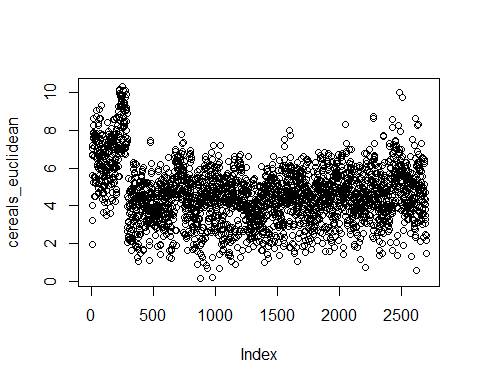
# Step 2: Hierarchical Clustering

# Normalize the data

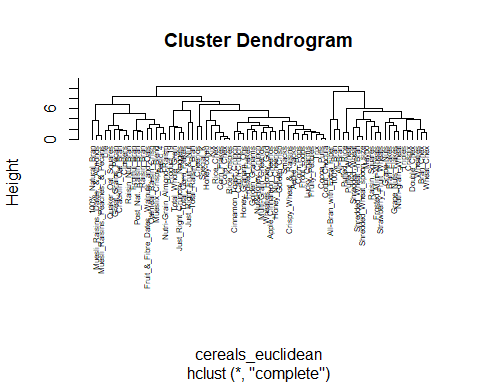
normalized\_cereals <- scale(cereals\_data[, -1])

# Use the Euclidean distance measure

cereals\_euclidean <- dist(normalized\_cereals, method = "euclidean")  
plot(cereals\_euclidean)



CC1 <- hclust(cereals\_euclidean)  
plot(CC1, cex = 0.5, hang = -1)

 # Use Agnes

library(cluster)  
library(stats)  
Cereal\_Single <- agnes(normalized\_cereals, method = "single")  
Cereal\_Complete <- agnes(normalized\_cereals, method = "complete")  
Cereal\_average <- agnes(normalized\_cereals, method = "average")  
Cereal\_Ward <- agnes(normalized\_cereals, method = "ward")

print(Cereal\_Single$ac)

## [1] 0.5892187

print(Cereal\_Complete$ac)

## [1] 0.8347316

print(Cereal\_average$ac)

## [1] 0.7677452

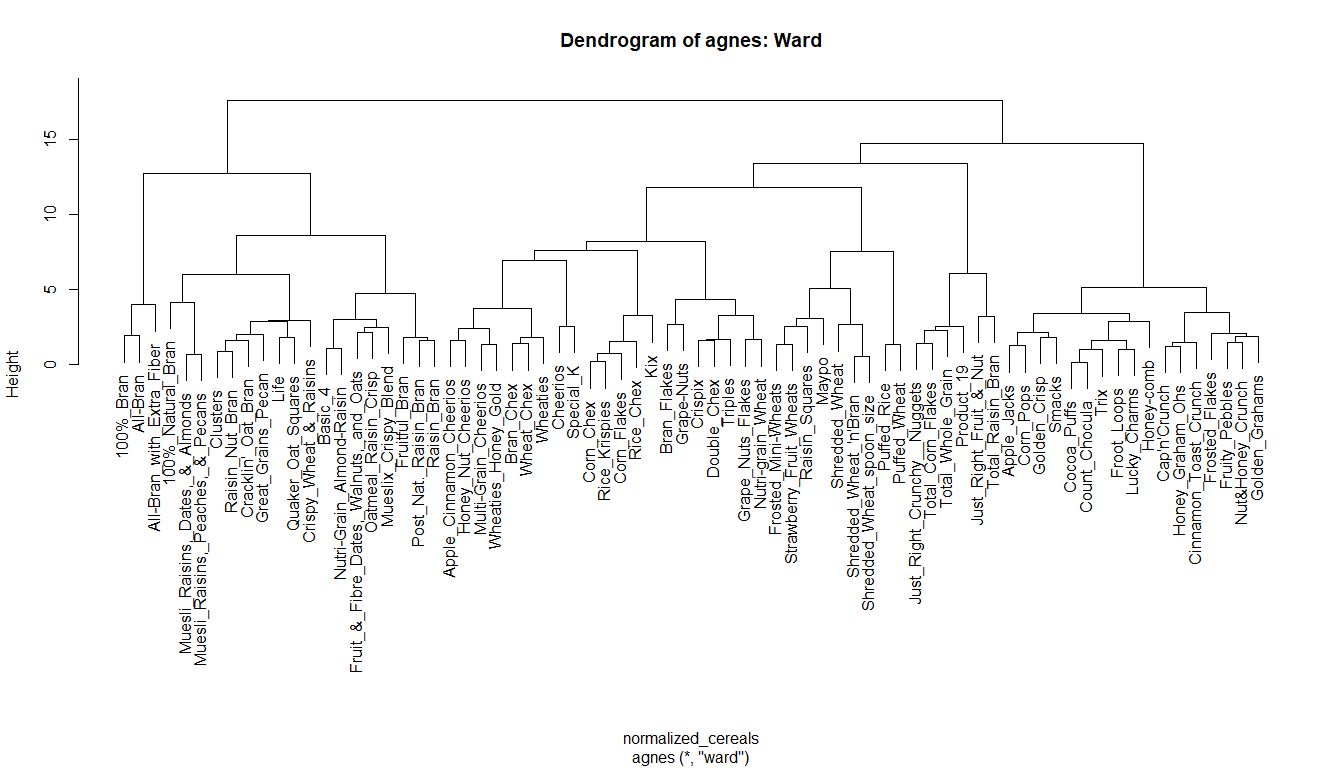
# Ward appears to be the best method

print(Cereal\_Ward$ac)

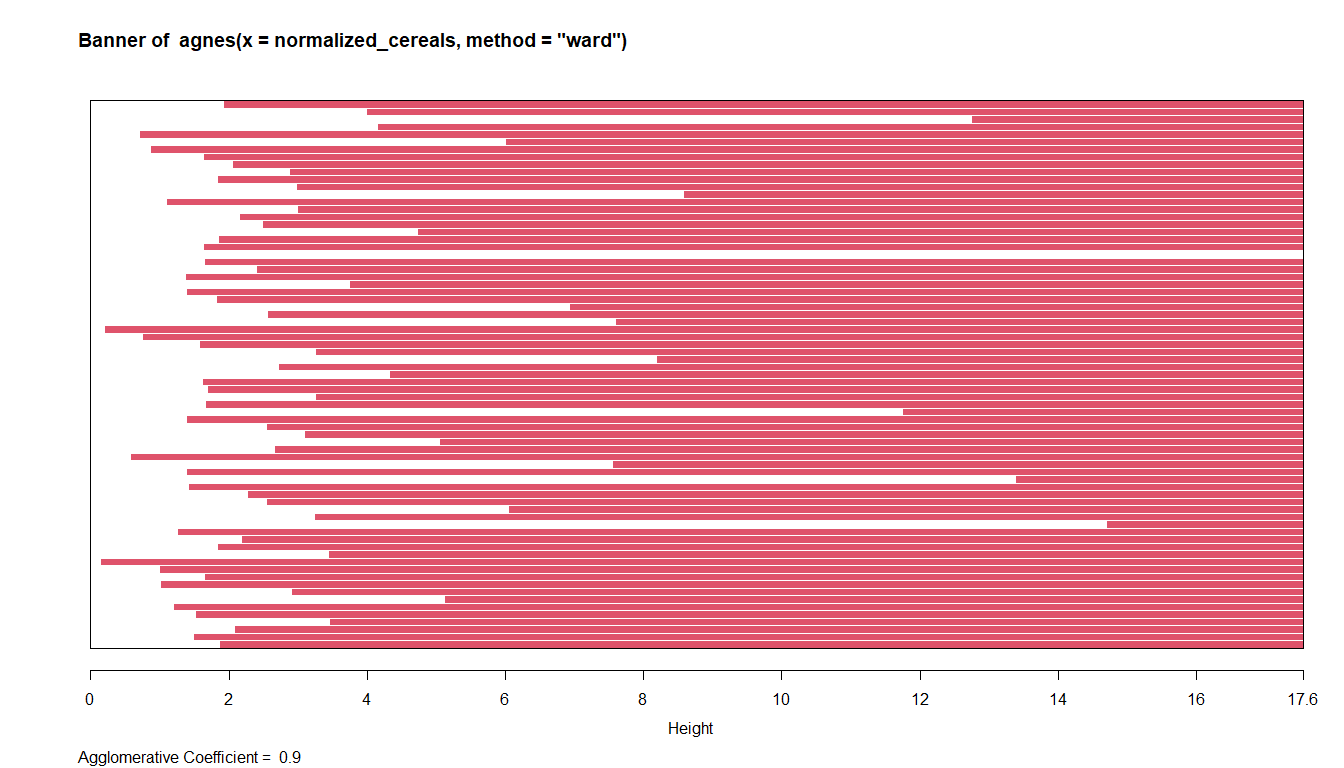
## [1] 0.9021306

# Apply hierarchical clustering with Ward’s method & Visualize for Ward’s method

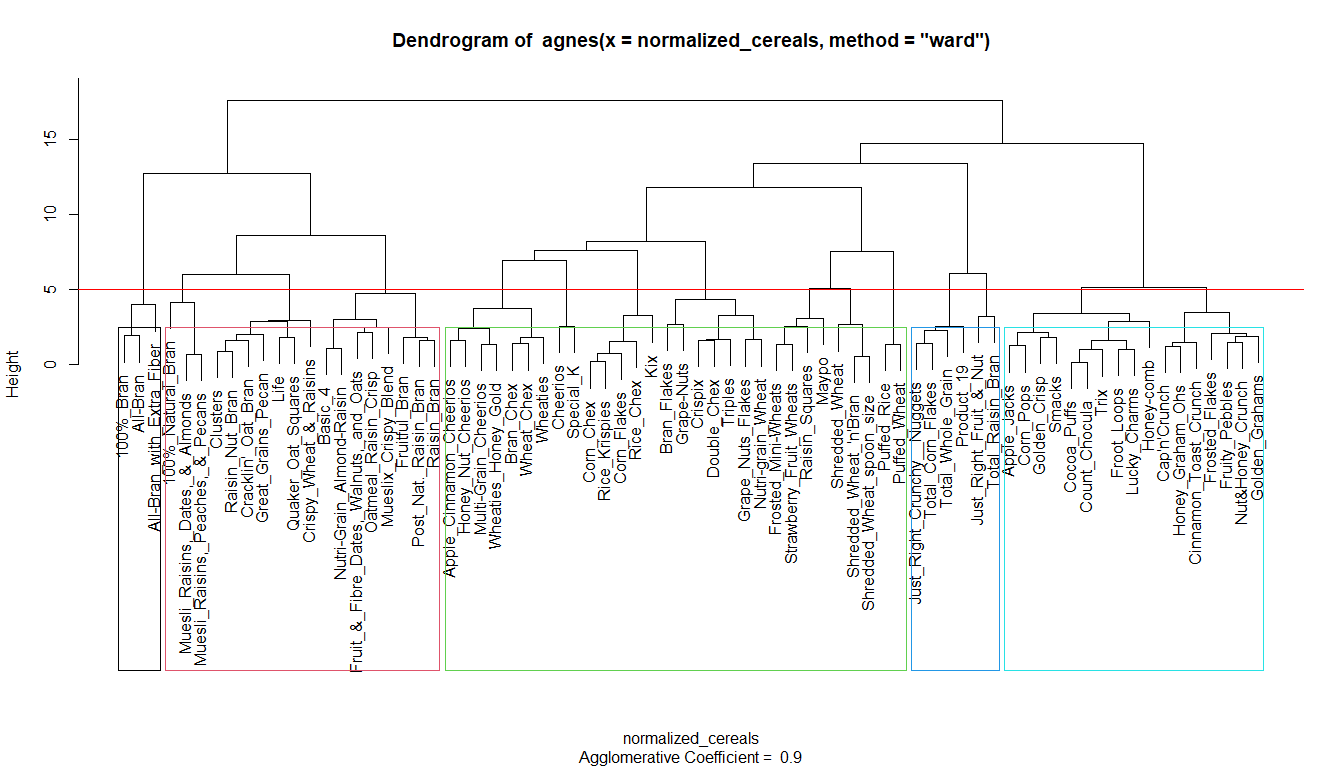
pltree(Cereal\_Ward,main = "Dendrogram of agnes: Ward")



plot(Cereal\_Ward)



rect.hclust(Cereal\_Ward, k = 5, border = 1:5) # I would choose 5 clusters.  
abline(h = 5, col = 'red')



# Step 3: Cluster Stability and Healthy Cereals

# Create cluster partitions A and B

set.seed(123)  
partition\_A <- sample(1:2, nrow(normalized\_cereals), replace = TRUE)  
partition\_B <- 3 - partition\_A

# Fit cluster on partition A

cluster\_A <- cutree(Cereal\_Ward, k = 5) # Assuming 4 clusters, adjust as needed

# Calculate centroids of clusters in partition A

library(dplyr)

## Warning: package 'dplyr' was built under R version 4.3.2

##   
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':  
##   
## filter, lag

## The following objects are masked from 'package:base':  
##   
## intersect, setdiff, setequal, union

# Combine partition\_A, cluster\_A, and normalized\_cereals for easier processing

combined\_data <- cbind(partition = partition\_A, cluster = cluster\_A)  
combined\_data\_df <- as.data.frame(combined\_data)

# Calculate centroids of clusters in partition A

cluster\_A\_centroids <- combined\_data\_df %>%  
 group\_by(cluster) %>%  
 summarise\_all(mean) %>%  
 select(-partition)

# Calculate distances between partition B records and centroids from partition A

distances\_to\_centroids <- apply(normalized\_cereals[partition\_B, ], 1, function(record) {  
 apply(cluster\_A\_centroids, 1, function(centroid) {  
 sqrt(sum((record - centroid)^2))  
 })  
})

# Assign clusters to partition B based on minimum distances

cluster\_B <- apply(distances\_to\_centroids, 2, which.min)

# Assess cluster consistency

consistency <- sum(cluster\_A == cluster\_B) / length(cluster\_B)

# Identify healthy cereals cluster

# Let’s assume ‘healthy’ cereals have low sugar and high fiber

healthy\_cereals\_cluster <- cluster\_A[which(cereals\_data$sugars < 5 & cereals\_data$fiber > 5)]

# Print the results

cat("Cluster Consistency:", consistency, "\n")

## Cluster Consistency: 0.04054054

cat("Healthy Cereals Cluster:", healthy\_cereals\_cluster, "\n")

## Healthy Cereals Cluster: 1