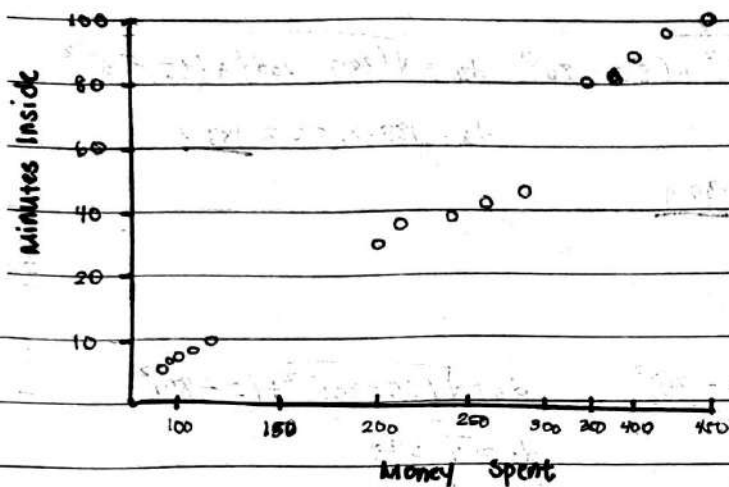


Almadrones, Kurt Justine M. CM4232

1.) Plot the data in a scatter plot



2.) Using the randomly selected initial clusters, calculate the distances of all data points using Euclidean distance

| Customer Name | Money Spent | Minutes Inside | d_1 | d_2 | d_3 | Cluster |
|------------------|-------------|----------------|-------|-------|-------|------------------------------|
| Chloe Mendoza | 380 | 85 | 310.5 | 188.2 | 30.4 | High Spender |
| Anna Reyes | 80 | 5 | 0 | 122.6 | 280.2 | Low Spender |
| Mika Tan | 280 | 45 | 209.0 | 81.3 | 78.3 | Average High Spender |
| Zach Vy | 400 | 90 | 331.1 | 384.7 | 51.0 | High Spender |
| Kevin Ramos | 220 | 35 | 143.2 | 90.6 | 137.6 | High Spender Average Spender |
| Sofia Peler Peña | 110 | 9 | 30.3 | 92.4 | 250.3 | Low Spender |
| Brian Lim | 360 | 80 | 280.6 | 158.1 | 0 | High Spender |
| Caleb Ong | 450 | 100 | 380.0 | 259.6 | 102.0 | High Spender |
| Liam Cruz | 100 | 8 | 20.2 | 102.9 | 260.2 | Low Spender |
| Ella Navarro | 240 | 38 | 163.4 | 40.8 | 117.7 | Average Spender |
| John Mercado | 120 | 10 | 40.3 | 82.5 | 240.4 | Low Spender |
| James Pines | 260 | 42 | 183.8 | 61.2 | 97.7 | Average Spender |
| Mark Santos | 95 | 1 | 15.1 | 107.5 | 265.2 | Low Spender |
| Paula Gomez | 200 | 20 | 112.6 | 0 | 158.1 | Average Spender |
| Hannah Roque | 420 | 95 | 351.7 | 229.9 | 71.6 | High Spender |

$$= D = \sqrt{(x-a)^2 + (y-b)^2}$$

1.) Chloe Mendoza

$$d_1 = \sqrt{(380-80)^2 + (85-5)^2} \quad d_2 = \sqrt{(380-200)^2 + (85-30)^2} \quad d_3 = \sqrt{(380-350)^2 + (85-80)^2}$$

$$d_1 = 20\sqrt{241} \quad d_2 = 180.2155 \approx 180.2 \quad d_3 = 5\sqrt{37}$$

$$d_1 = 310.4834 \approx 310.5 \quad d_2 = 30.4138 \approx 30.4 \quad d_3 = 30.4138 \approx 30.4$$

2.) Anna Reyes

$$d_1 = \sqrt{(80-80)^2 + (5-5)^2} \quad d_2 = \sqrt{(80-200)^2 + (5-30)^2} \quad d_3 = \sqrt{(80-350)^2 + (5-80)^2}$$

$$d_1 = 0 \quad d_2 = 5\sqrt{401} \quad d_3 = \sqrt{349}$$

$$d_2 = 122.5765 \approx 122.6 \quad d_3 = 280.2231 \approx 280.2$$

3.) Mika Tan

$$d_1 = \sqrt{(280-80)^2 + (45-5)^2} \quad d_2 = \sqrt{(280-200)^2 + (45-30)^2} \quad d_3 = \sqrt{(280-350)^2 + (45-80)^2}$$

$$d_1 = 40\sqrt{26} \quad d_2 = 5\sqrt{265} \quad d_3 = \sqrt{5}$$

$$d_1 = 33.9407 \approx 34.0 \quad d_2 = 81.3941 \approx 81.4 \quad d_3 = 78.2623 \approx 78.3$$

4.) Zach Uy

$$d_1 = \sqrt{(400-80)^2 + (90-5)^2} \quad d_2 = \sqrt{(400-200)^2 + (90-30)^2} \quad d_3 = \sqrt{(400-350)^2 + (90-80)^2}$$

$$d_1 = 381.0966 \approx 381.1 \quad d_2 = 20\sqrt{370} \quad d_3 = 10\sqrt{26}$$

$$d_2 = 384.7076 \approx 384.7 \quad d_3 = 50.9901 \approx 51.0$$

5.) Kevin Ramos

$$d_1 = \sqrt{(220-80)^2 + (35-5)^2} \quad d_2 = \sqrt{(220-200)^2 + (35-30)^2} \quad d_3 = \sqrt{(220-350)^2 + (35-80)^2}$$

$$d_1 = 10\sqrt{205} \quad d_2 = 5\sqrt{17} \quad d_3 = 5\sqrt{757}$$

$$d_1 = 143.1788 \approx 143.2 \quad d_2 = 20.6155 \approx 20.6 \quad d_3 = 137.5681 \approx 137.6$$

6.) Sofia Dela Peña

$$d_1 = \sqrt{(110-80)^2 + (9-5)^2} \quad d_2 = \sqrt{(110-200)^2 + (9-30)^2} \quad d_3 = \sqrt{(110-350)^2 + (9-80)^2}$$

$$d_1 = 2\sqrt{229} \quad d_2 = 3\sqrt{949} \quad d_3 = \sqrt{250.2818} \approx 250.3$$

$$d_1 = 30.2659 \approx 30.3 \quad d_2 = 92.4175 \approx 92.4$$

7.) Brian Lim

$$d_1 = \sqrt{(350-80)^2 + (80-5)^2} \quad d_2 = \sqrt{(350-200)^2 + (80-30)^2} \quad d_3 = \sqrt{(350-350)^2 + (80-80)^2}$$

$$d_1 = 15\sqrt{349}$$

$$d_2 = 50\sqrt{10}$$

$$d_3 = 0$$

$$d_1 = 280.2231 \approx \underline{280.2}$$

$$d_2 = 158.1138 \approx \underline{158.1}$$

8.) Caleb Ong

$$d_1 = \sqrt{(450-80)^2 + (100-5)^2} \quad d_2 = \sqrt{(450-200)^2 + (100-30)^2} \quad d_3 = \sqrt{(450-350)^2 + (100-80)^2}$$

$$d_1 = 380.0001 \approx \underline{380.0}$$

$$d_2 = 10\sqrt{674}$$

$$d_3 = 20\sqrt{26}$$

$$d_2 = 269.6150 \approx \underline{269.6}$$

$$d_3 = 101.9803 \approx \underline{102.0}$$

$$(2.5, 101): \text{mid} =$$

9.) Liam Cruz

$$d_1 = \sqrt{(100-80)^2 + (8-5)^2} \quad d_2 = \sqrt{(100-200)^2 + (8-30)^2} \quad d_3 = \sqrt{(100-350)^2 + (8-80)^2}$$

$$d_1 = \sqrt{409}$$

$$d_2 = \underline{102.4}$$

$$d_3 = \underline{260.2}$$

$$d_1 = \underline{20.2}$$

$$(8.08, 0.08): \text{mid} =$$

10.) Ella Navarro

$$d_1 = \sqrt{(240-80)^2 + (38-5)^2} \quad d_2 = \sqrt{(240-200)^2 + (38-30)^2} \quad d_3 = \sqrt{(240-350)^2 + (38-80)^2}$$

$$d_1 = \underline{163.4}$$

$$d_2 = \underline{40.8}$$

$$d_3 = \underline{117.7}$$

11.) John Mercado

$$d_1 = \sqrt{(120-80)^2 + (10-5)^2} \quad d_2 = \sqrt{(120-200)^2 + (10-30)^2} \quad d_3 = \sqrt{(120-350)^2 + (10-80)^2}$$

$$d_1 = \underline{40.3}$$

$$d_2 = \underline{82.5}$$

$$d_3 = \underline{240.4}$$

12.) Jared Pines

$$d_1 = \sqrt{(260-80)^2 + (42-5)^2} \quad d_2 = \sqrt{(260-200)^2 + (42-30)^2} \quad d_3 = \sqrt{(260-350)^2 + (42-80)^2}$$

$$d_1 = \underline{183.8}$$

$$d_2 = \underline{61.2}$$

$$d_3 = \underline{97.7}$$

13.) Mark Santos

$$d_1 = \sqrt{(95-80)^2 + (7-5)^2} \quad d_2 = \sqrt{(95-200)^2 + (7-30)^2} \quad d_3 = \sqrt{(95-350)^2 + (7-80)^2}$$

$$d_1 = \underline{15.1}$$

$$d_2 = \underline{107.5}$$

$$d_3 = \underline{265.2}$$

14.) Paula Gomes

$$d_1 = \sqrt{(200-80)^2 + (30-5)^2} \quad d_2 = \sqrt{(200-200)^2 + (30-30)^2} \quad d_3 = \sqrt{(200-300)^2 + (30-80)^2}$$
$$d_1 = \underline{122.6} \quad d_2 = \underline{0} \quad d_3 = \underline{158.1}$$

15.) Hannah Roque

$$d_1 = \sqrt{(420-80)^2 + (95-5)^2} \quad d_2 = \sqrt{(420-200)^2 + (95-30)^2} \quad d_3 = \sqrt{(420-300)^2 + (95-80)^2}$$
$$d_1 = \underline{359.7} \quad d_2 = \underline{229.4} \quad d_3 = \underline{71.6}$$

4.) Calculate the mean of each cluster

1.) Low Spender

= Low: (101, 7.8)

$$a_1 = \frac{(80 + 110 + 100 + 120 + 95)}{5} = \underline{101}$$

$$b_1 = \frac{(5 + 9 + 8 + 10 + 7)}{5} = \underline{7.8}$$

2.) Average Spender

= Average: (230, 36.3)

$$a_2 = \frac{(220 + 240 + 260 + 200)}{4} = \underline{230}$$

$$b_2 = \frac{(35 + 38 + 42 + 30)}{4} = \underline{36.3}$$

3.) High Spender

= High: (380, 82.5)

$$a_3 = \frac{(380 + 280 + 400 + 350 + 450 + 420)}{6} = \underline{380}$$

$$b_3 = \frac{(85 + 45 + 90 + 80 + 100 + 95)}{6} = \underline{82.5}$$

5.) Recalculate the distance of all data points from the mean using Euclidean distance

1.) Chloe Mendoza

$$d_1 = \sqrt{(380-101)^2 + (85-7.8)^2}$$

$$d_1 = 289.4837 \approx 289.5$$

$$d_2 = \sqrt{(380-230)^2 + (85-36.3)^2}$$

$$d_2 = 157.7076 \approx 157.7$$

$$d_3 = \sqrt{(380-380)^2 + (85-82.5)^2}$$

$$d_3 = 2.5$$

2.) Anna Reyes

$$d_1 = \sqrt{(80-101)^2 + (5-7.8)^2}$$

$$d_1 = 21.1858 \approx 21.2$$

$$d_2 = \sqrt{(80-230)^2 + (5-36.3)^2}$$

$$d_2 = 153.2308 \approx 153.2$$

$$d_3 = \sqrt{(80-380)^2 + (5-82.5)^2}$$

$$d_3 = 309.8487 \approx 309.8$$

3.) Mika Tan

$$d_1 = \sqrt{(280-101)^2 + (45-7.8)^2}$$

$$d_1 = 182.8246 \approx 182.8$$

$$d_2 = \sqrt{(280-230)^2 + (45-36.3)^2}$$

$$d_2 = 50.7512 \approx 50.8$$

$$d_3 = \sqrt{(280-380)^2 + (45-82.5)^2}$$

$$d_3 = 106.8000 \approx 106.1$$

4.) Zach Uy

$$d_1 = \sqrt{(400-101)^2 + (90-7.8)^2}$$

$$d_1 = 310.0932 \approx 310.1$$

$$d_2 = \sqrt{(400-230)^2 + (90-36.3)^2}$$

$$d_2 = 178.2798 \approx 178.3$$

$$d_3 = \sqrt{(400-380)^2 + (90-82.5)^2}$$

$$d_3 = 21.3600 \approx 21.4$$

5.) Kevin Ramos

$$d_1 = \sqrt{(220-101)^2 + (35-7.8)^2}$$

$$d_1 = 122.0689 \approx 122.1$$

$$d_2 = \sqrt{(220-230)^2 + (35-36.3)^2}$$

$$d_2 = 10.0841 \approx 10.1$$

$$d_3 = \sqrt{(220-380)^2 + (35-82.5)^2}$$

$$d_3 = 166.9019 \approx 166.9$$

6.) Sofia Dela Peña

$$d_1 = \sqrt{(110-101)^2 + (9-7.8)^2}$$

$$d_1 = 9.0796 \approx 9.1$$

$$d_2 = \sqrt{(110-230)^2 + (9-36.3)^2}$$

$$d_2 = 123.0662 \approx 123.1$$

$$d_3 = \sqrt{(110-380)^2 + (9-82.5)^2}$$

$$d_3 = 279.8253 \approx 279.8$$

7.) Brian Lim

$$d_1 = \sqrt{(350-101)^2 + (80-7.8)^2}$$

$$d_1 = 259.2569 \approx 259.3$$

$$d_2 = \sqrt{(350-230)^2 + (80-36.3)^2}$$

$$d_2 = 127.7093 \approx 127.7$$

$$d_3 = \sqrt{(350-380)^2 + (80-82.5)^2}$$

$$d_3 = 30.1039 \approx 30.1$$

8.) Liam Cruz

$$d_1 = \sqrt{(100-101)^2 + (8-7.8)^2}$$

$$d_1 = 1.0$$

$$d_2 = \sqrt{(100-230)^2 + (8-36.3)^2}$$

$$d_2 = 133.0$$

$$d_3 = \sqrt{(100-380)^2 + (8-82.5)^2}$$

$$d_3 = 289.7$$

9.) Ella Navarro

$$d_1 = \sqrt{(240-101)^2 + (38-7.8)^2}$$

$$d_1 = 142.2$$

$$d_2 = \sqrt{(240-230)^2 + (38-36.3)^2}$$

$$d_2 = 10.1$$

$$d_3 = \sqrt{(240-380)^2 + (38-82.5)^2}$$

$$d_3 = 146.9$$

11.) Jared Flores

$$d_1 = \sqrt{(240-101)^2 + (42-7.8)^2}$$

$$d_1 = 162.4$$

$$d_2 = \sqrt{(240-230)^2 + (42-36.3)^2}$$

$$d_2 = 30.5$$

$$d_3 = \sqrt{(240-380)^2 + (42-82.5)^2}$$

$$d_3 = 126.7$$

3.) Caleb Ong

$$d_1 = \sqrt{(450-101)^2 + (100-7.8)^2}$$

$$d_1 = 360.9734 \approx 361.0$$

$$d_2 = \sqrt{(450-230)^2 + (100-36.3)^2}$$

$$d_2 = 229.6364 \approx 229.0$$

$$d_3 = \sqrt{(450-380)^2 + (100-82.5)^2}$$

$$d_3 = 72.1543 \approx 72.2$$

10.) John Mercado

$$d_1 = \sqrt{(120-101)^2 + (10-7.8)^2}$$

$$d_1 = 19.1$$

$$d_2 = \sqrt{(120-230)^2 + (10-36.3)^2}$$

$$d_2 = 113.1$$

$$d_3 = \sqrt{(120-380)^2 + (10-82.5)^2}$$

$$d_3 = 269.9$$

12.) Mark Santos

$$d_1 = \sqrt{(95-101)^2 + (7-7.8)^2}$$

$$d_1 = 6.1$$

$$d_2 = \sqrt{(95-230)^2 + (7-36.3)^2}$$

$$d_2 = 138.1$$

$$d_3 = \sqrt{(95-380)^2 + (7-82.5)^2}$$

$$d_3 = 294.8$$

14.) Paula Gomez

$$d_1 = \sqrt{(200-101)^2 + (30-7.8)^2}$$

$$d_1 = 101.5$$

$$d_2 = \sqrt{(200-230)^2 + (30-36.3)^2}$$

$$d_2 = 30.7$$

$$d_3 = \sqrt{(200-380)^2 + (30-82.5)^2}$$

$$d_3 = 187.5$$

15.) Hannah Roque

$$d_1 = \sqrt{(820-101)^2 + (95-7.8)^2}$$

$$d_1 = 820.7$$

$$d_2 = \sqrt{(820-230)^2 + (95-36.3)^2}$$

$$d_2 = 801.9$$

$$d_3 = \sqrt{(820-380)^2 + (95-82.5)^2}$$

$$d_3 = 41.9$$

6.) Assign a cluster for all data points

| Customer Name | Money Spent | Minutes Inside | d1 | d2 | d3 | Cluster |
|-----------------|-------------|----------------|-------|-------|-------|-----------------|
| Chloe Mendoza | 360 | 85 | 289.5 | 157.7 | 2.5 | High Spender |
| Anna Reyes | 80 | 5 | 21.2 | 153.2 | 309.8 | Low Spender |
| Mika Tan | 280 | 45 | 182.8 | 50.8 | 104.1 | Average Spender |
| Zach Uy | 400 | 90 | 310.1 | 178.3 | 21.4 | High Spender |
| Kevin Ramos | 220 | 35 | 122.1 | 10.1 | 164.9 | Average Spender |
| Sofia Dela Peña | 110 | 9 | 9.1 | 123.1 | 279.8 | Low Spender |
| Brian Lim | 350 | 80 | 259.3 | 127.7 | 30.1 | High Spender |
| Caleb Ong | 450 | 100 | 341.0 | 229.0 | 72.2 | High Spender |
| Liam Cruz | 100 | 8 | 1.0 | 113.0 | 284.7 | Low Spender |
| Ella Navarro | 240 | 38 | 142.2 | 10.1 | 146.9 | Average Spender |
| John Mercado | 120 | 10 | 19.1 | 113.1 | 269.9 | Low Spender |
| Jared Flores | 260 | 42 | 142.6 | 30.1 | 126.7 | Average Spender |
| Mark Santos | 95 | 7 | 6.1 | 138.1 | 294.8 | Low Spender |
| Paula Gomez | 200 | 30 | 101.5 | 30.7 | 187.5 | Average Spender |
| Hannah Roque | 420 | 95 | 330.7 | 198.9 | 41.9 | High Spender |

7.) Plot the final cluster in a scatter plot

