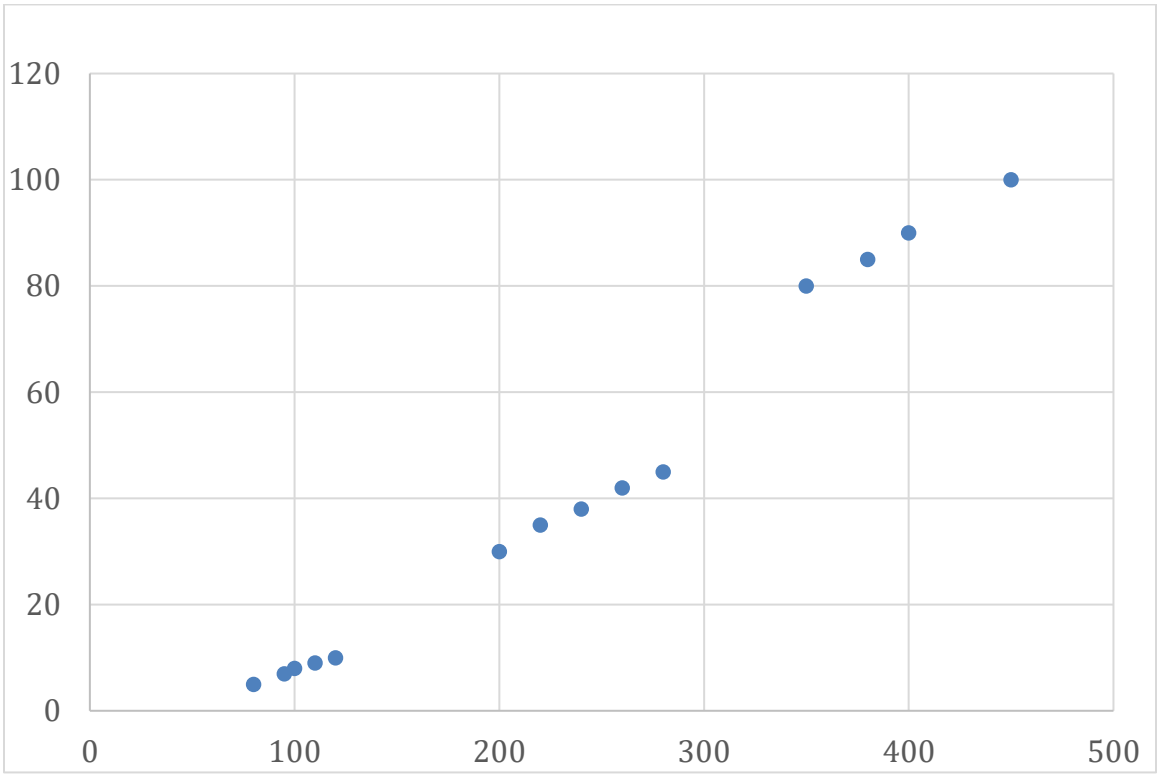
EXERCISE 1 Answer Key

1. Plot the data in a scatter plot (10 pts)



2. Using the randomly selected initial clusters, calculate the distances of all data points using Euclidean distance (15 pts)

3. Assign a cluster for all data points (15 pts)

Let D1, D2 and D3 be the distances of each data point from the initial clusters A,B, C respectively.

Customer	Money Spent	Minutes Inside	D1	D2	D3	Cluster
Chloe Mendoza	380	85	310.48	188.22	30.41	C
Anna Reyes	80	5	0.00	122.58	280.22	A
Mika Tan	280	45	203.96	81.39	78.26	C
Zach Uy	400	90	331.10	208.81	50.99	C
Kevin Ramos	220	35	143.18	20.62	137.57	B
Sofia Dela Pena	110	9	30.27	92.42	250.28	A
Brian Lim	350	80	280.22	158.11	0.00	C
Caleb Ong	450	100	382.00	259.62	101.98	C
Liam Cruz	100	8	20.22	102.39	260.16	A
Ella Navarro	240	38	163.37	40.79	117.75	B
John Mercado	120	10	40.31	82.46	240.42	B
Jared Flores	260	42	183.76	61.19	97.69	B
Mark Santos	95	7	15.13	107.49	265.24	A
Paula Gomez	200	30	122.58	0.00	158.11	B
Hannah Roque	420	95	351.71	229.40	71.59	C

4. Calculate the mean of each cluster (15 pts)

Mean of cluster A: (101, 7.8)

Mean of cluster B: (230, 36.25)

Mean of cluster C: (380, 82.5)

5. Recalculate the distance of all data points from the mean using Euclidean distance (15 pts)

6. Assign a cluster for all data points (15 pts)

Let D1, D2 and D3 be the distances of each data point from Means of cluster A,B, C respectively.

Customer	Money Spent	Minutes Inside	D1	D2	D3	Cluster
Chloe Mendoza	380	85	289.48	157.72	2.50	C
Anna Reyes	80	5	21.19	153.22	309.85	A
Mika Tan	280	45	182.82	50.76	106.80	B
Zach Uy	400	90	310.09	178.29	21.36	C
Kevin Ramos	220	35	122.07	10.08	166.90	B
Sofia Dela Pena	110	9	9.08	123.06	279.83	A
Brian Lim	350	80	259.26	127.73	30.10	C
Caleb Ong	450	100	360.97	229.05	72.15	C
Liam Cruz	100	8	1.02	133.03	289.74	A
Ella Navarro	240	38	142.24	10.15	146.90	B
John Mercado	120	10	19.13	113.09	269.92	A
Jared Flores	260	42	162.64	30.55	126.65	B
Mark Santos	95	7	6.05	138.13	294.83	A
Paula Gomez	200	30	101.46	30.64	187.50	B
Hannah Roque	420	95	330.70	198.88	41.91	C

7. Plot the final cluster in a scatter plot (15 pts)