Mini Project

By

Ali Faisal Raza,
Dimitri Salavci

Agenda

- Methodology

- Customer Demography:
  - Prepared Final Dataset
  - Determine Optimal Number of Clusters (Elbow Chart)
  - Radar Chart
  - PCA Chart

- Banking Behavior:
  - Prepared Final Dataset
  - Determine Optimal Number of Clusters (Elbow Chart)
  - Radar Chart
  - PCA Chart

Methodology

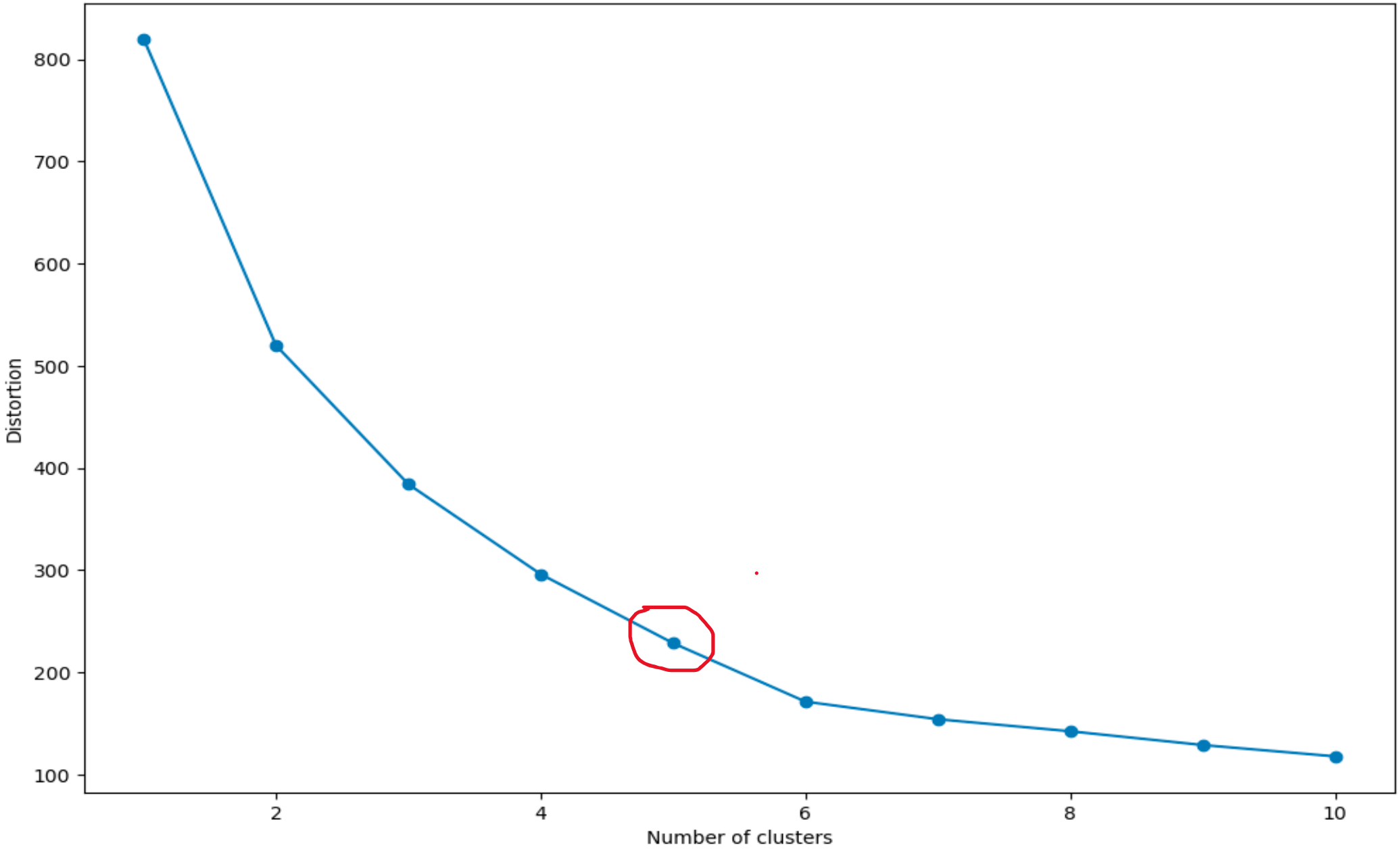
- Apply transformations on observations (scaling and adopting dummy variables).
- Plot elbow graph to determine the optimal number of clusters.
- Apply KMeans algorithm to divide observations into clusters.
- Prepare radar chart to show clusters together for comparison
- Perform PCA analysis to determine 2 principal components for observations.
- Prepare PCA chart to show the two principal components.

Customer Demography – Final Dataset

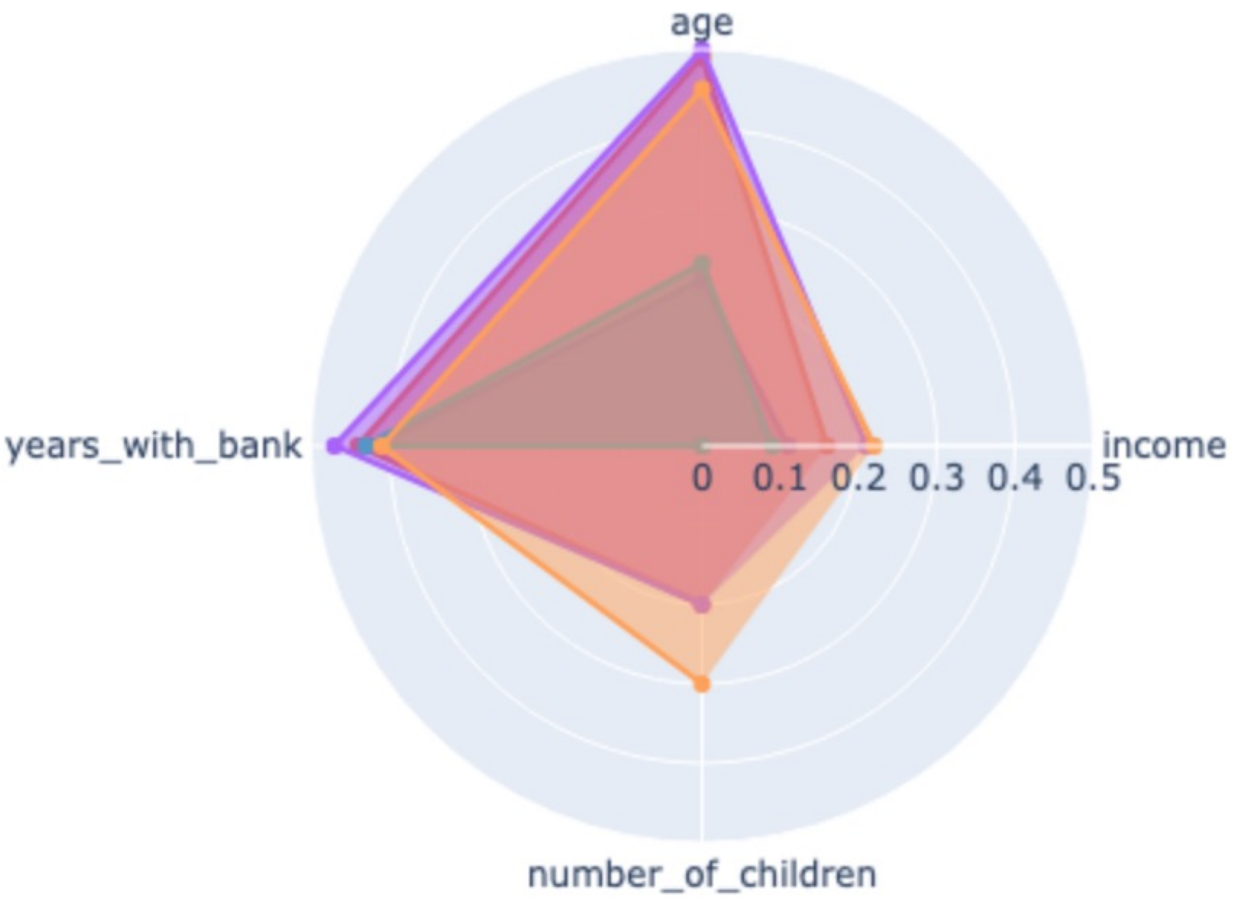
	income	age	years_with_bank	nbr_children	gender_num	marital_status_1	marital_status_2	marital_status_3	marital_status_4
0	0.181399	0.434211	0.555556	0.2	1.0	0.0	1.0	0.0	0.0
1	0.045818	0.763158	0.111111	0.0	1.0	0.0	1.0	0.0	0.0
2	0.128665	0.328947	0.888889	0.0	0.0	1.0	0.0	0.0	0.0
3	0.330667	0.539474	0.333333	0.0	0.0	1.0	0.0	0.0	0.0
4	0.309066	0.605263	1.000000	0.4	0.0	0.0	0.0	0.0	1.0
...	...	...	...	...	...	...	...	...	...
742	0.102631	0.302632	0.666667	0.2	0.0	0.0	0.0	0.0	1.0
743	0.183043	0.565789	0.666667	0.2	1.0	0.0	1.0	0.0	0.0
744	0.425231	0.486842	0.000000	0.4	1.0	0.0	1.0	0.0	0.0
745	0.104747	0.315789	0.777778	0.0	0.0	0.0	1.0	0.0	0.0
746	0.123997	0.065789	0.333333	0.0	0.0	1.0	0.0	0.0	0.0

747 rows × 9 columns

Customer Demography – Elbow Chart

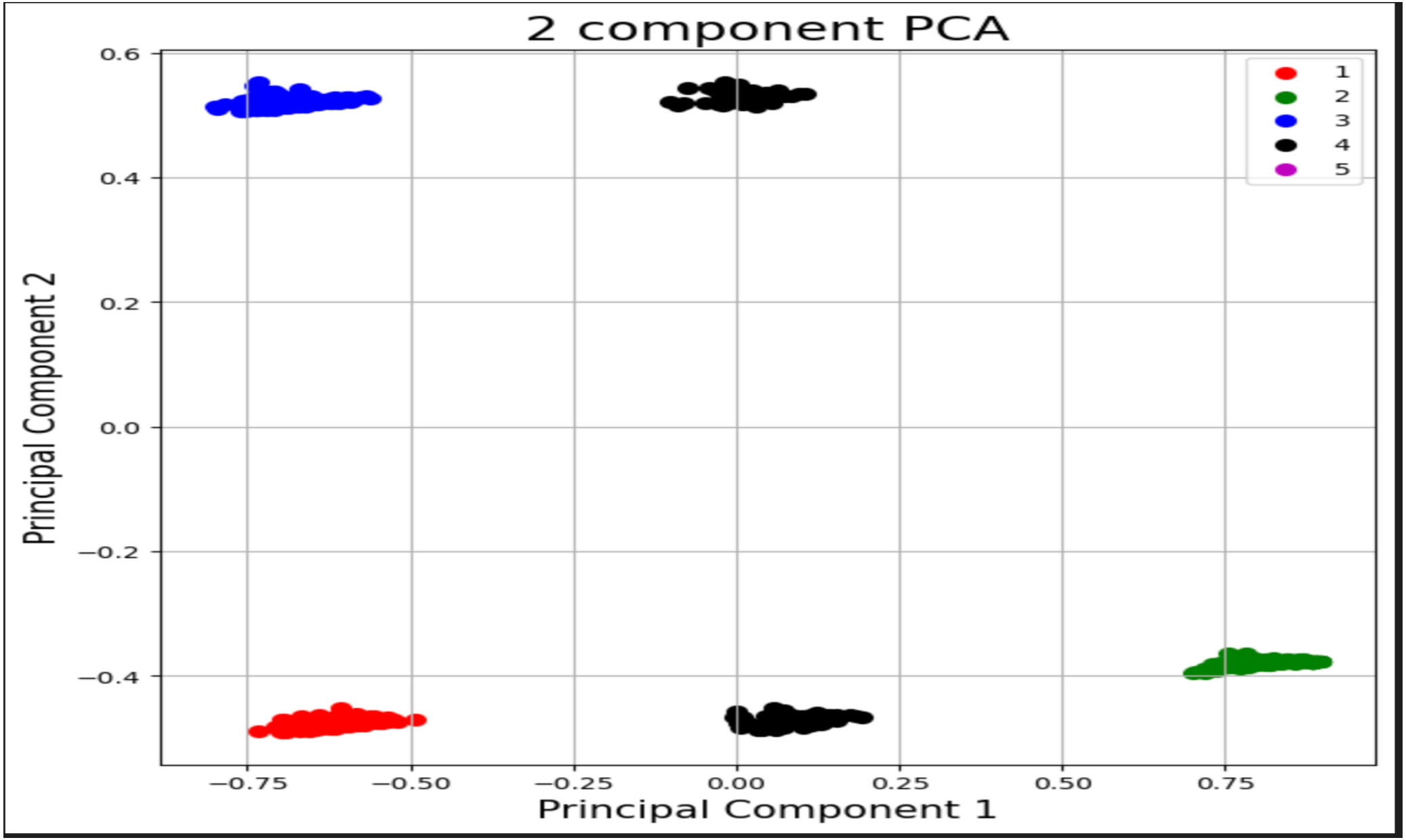


Customer Demography – Radar Chart



cluster	income	age	years_with_bank	nbr_children
0	0.11	0.21	0.42	0.0
1	0.16	0.49	0.44	0.2
2	0.09	0.23	0.43	0.0
3	0.21	0.50	0.47	0.2
4	0.22	0.45	0.41	0.3

Customer Demography – 2 Component PCA

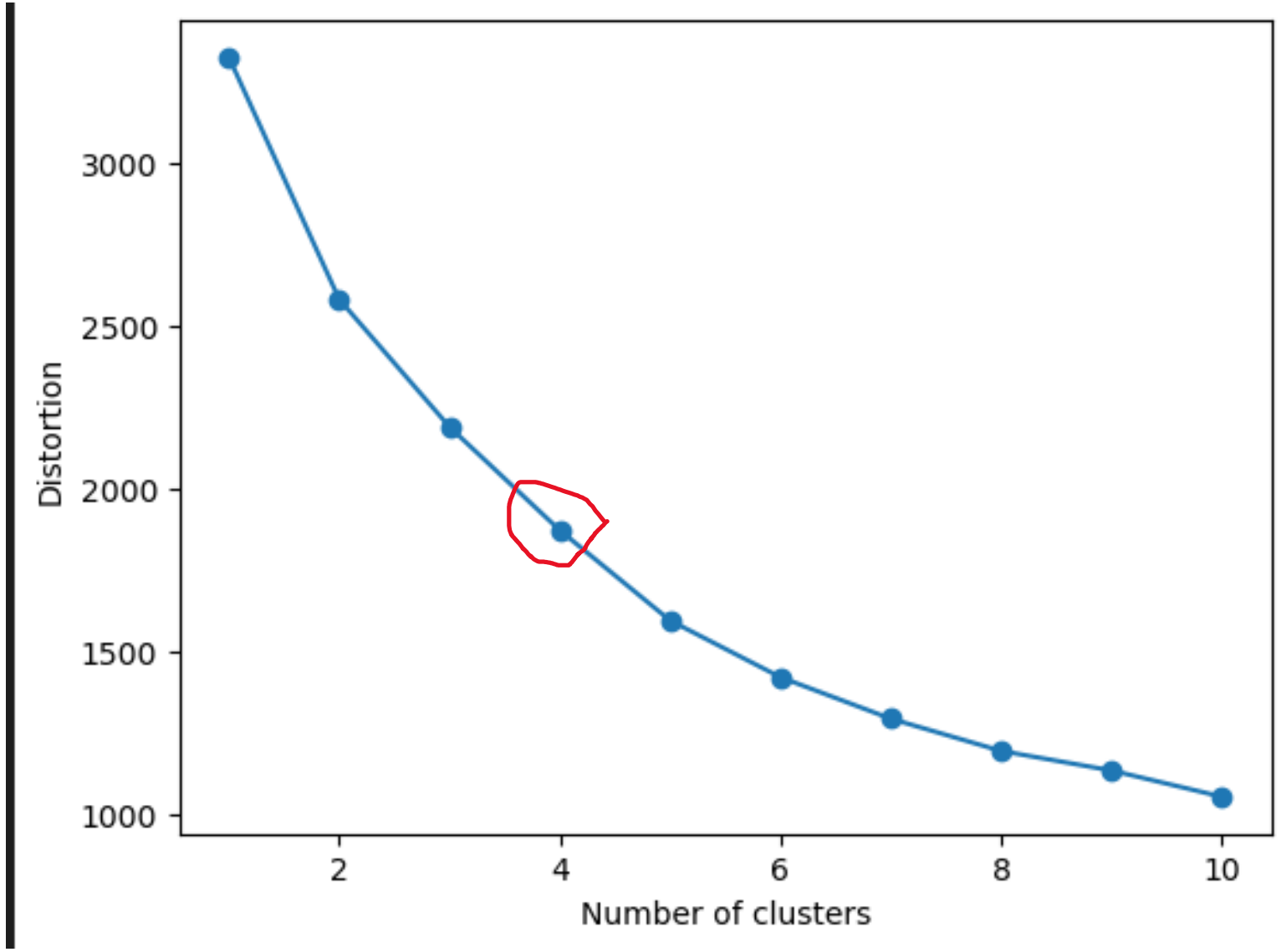


Banking Behavior – Final Dataset

	cust_id	total_savings	total_debt	small_trans_count_checking	big_trans_count_checking
0	1362480	480.88	2480.00	151	6
1	1362484	1927.41	2630.28	180	7
2	1362485	0.00	0.00	163	8
3	1362486	-168.24	0.00	50	0
4	1362487	-638.84	1451.77	100	3
...	...	...	...	...	...
660	1363490	34.24	0.00	12	0
661	1363491	-187.63	0.00	17	1
662	1363492	163.72	231.00	135	18
663	1363493	222.43	1500.00	116	2
664	1363495	461.26	0.00	134	1

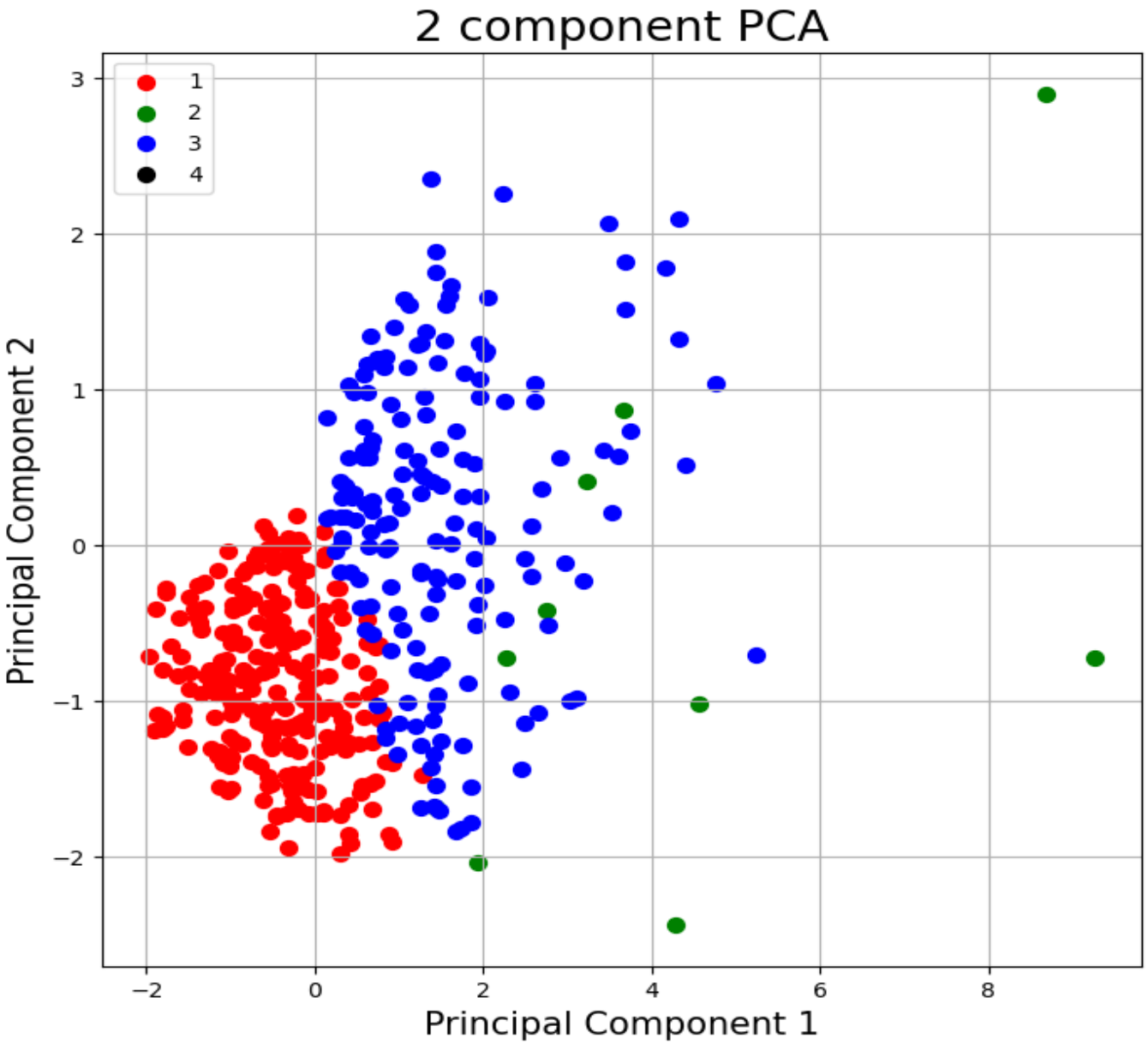
665 rows × 5 columns

Banking Behavior – Elbow Chart



Banking Behavior – Radar Chart

Banking Behavior – 2 Component PCA



Queries?