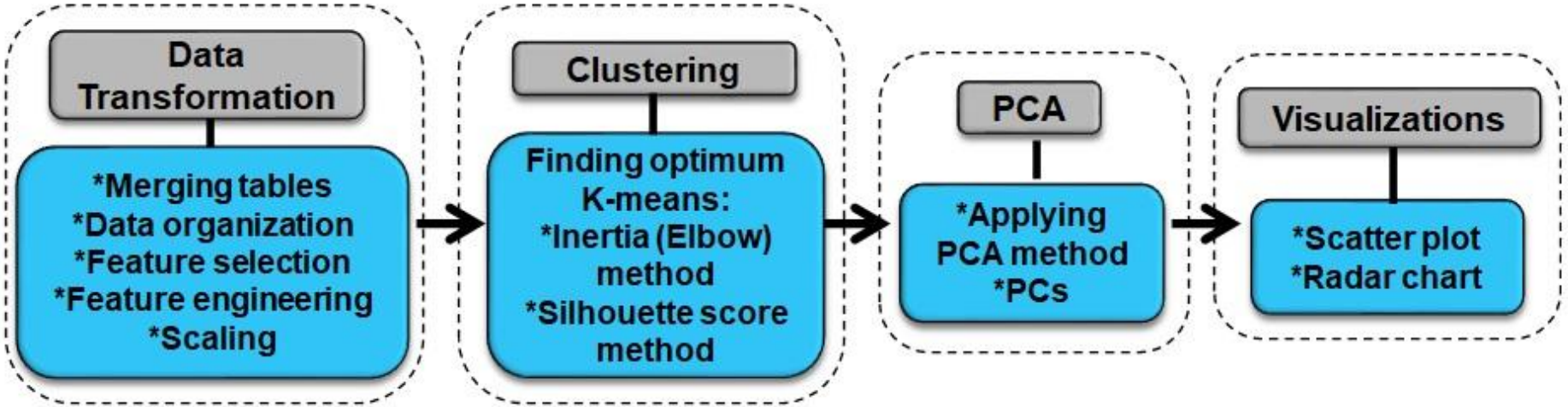
Clustering Project

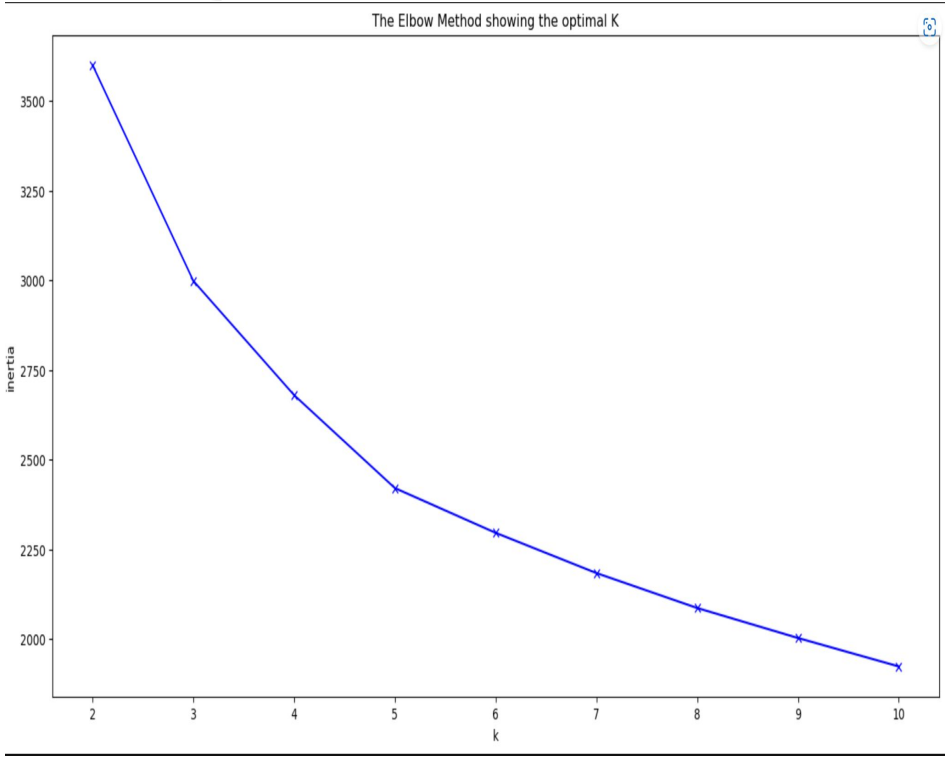
Workflow

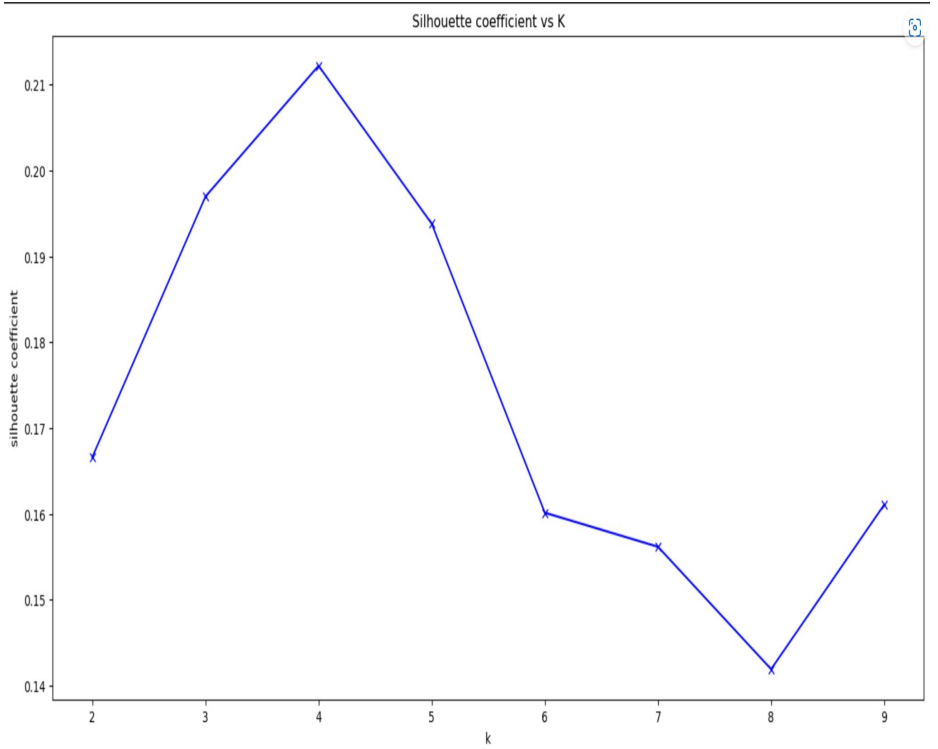

Data Transformation
*Merging tables
*Data organization
*Feature selection
*Feature engineering
*Scaling
Clustering
Finding optimum K-means:
*Inertia (Elbow) method
*Silhouette score method
PCA
*Applying PCA method
*PCs
Visualizations
*Scatter plot
*Radar chart

Demographics

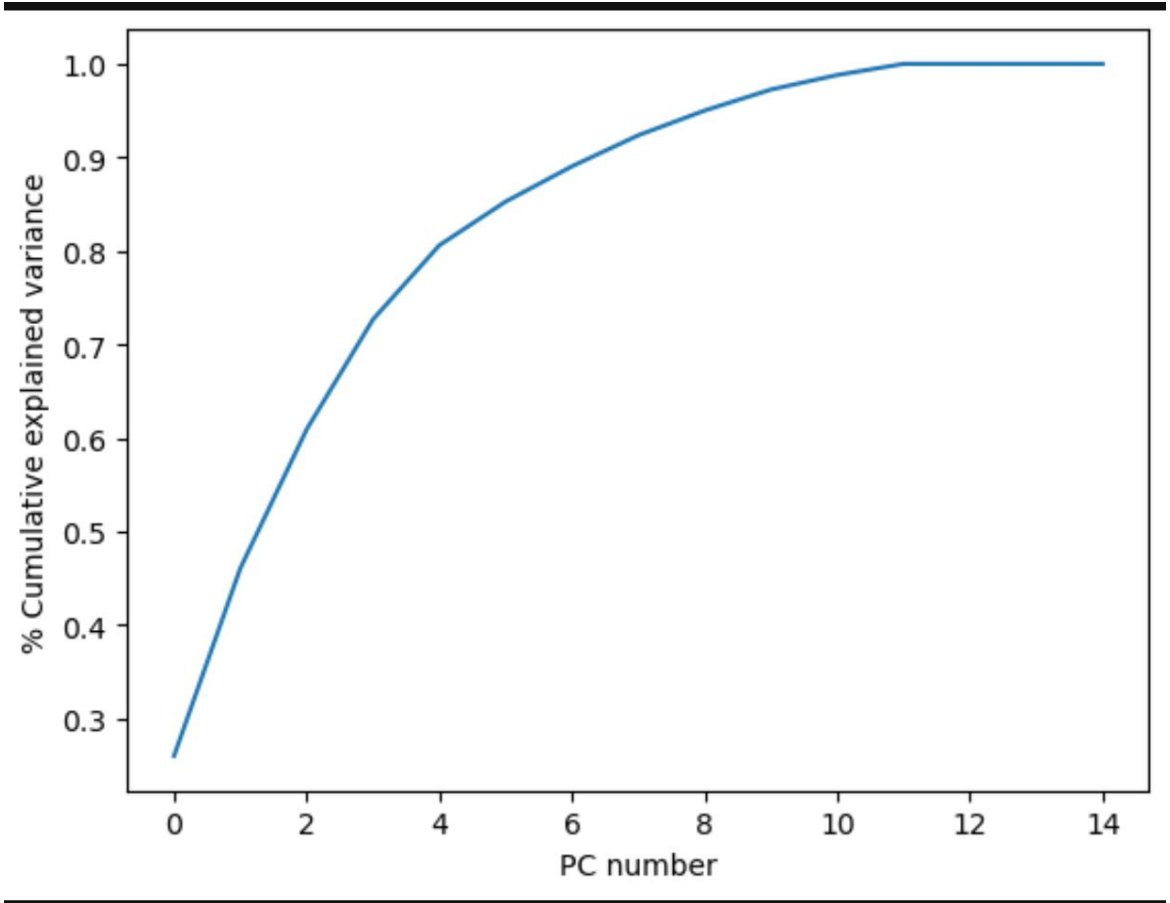
	income	age	years_with_bank	nbr_children	gender	marital_status	region
0	26150	46	5	1	M	2	5
1	6605	71	1	0	M	2	5
2	18548	38	8	0	F	1	5
3	47668	54	3	0	F	1	5
4	44554	59	9	2	F	4	1
...	...	...	...	...	...	...	...
742	14795	36	6	1	F	4	5
743	26387	56	6	1	M	2	1
744	61300	50	0	2	M	2	4
745	15100	37	7	0	F	2	1
746	17875	18	3	0	F	1	5

Demographics

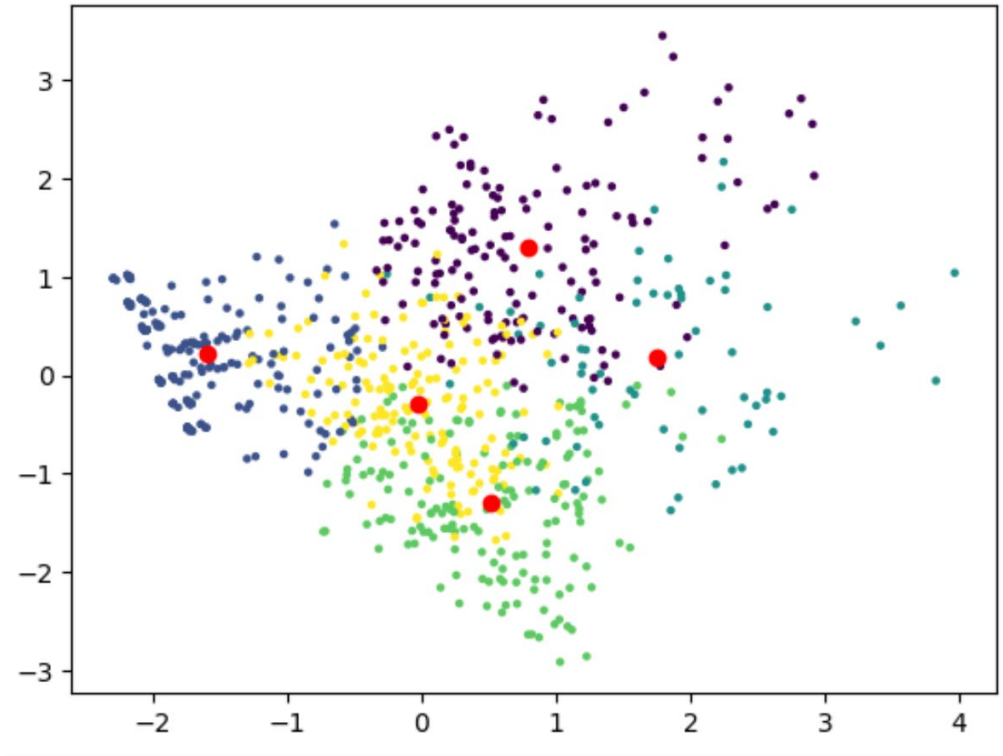


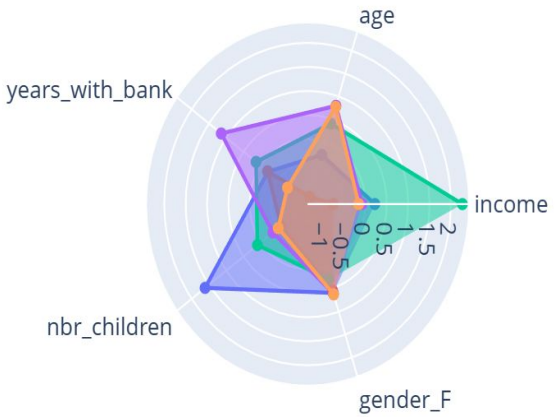


Demographics



Demographics

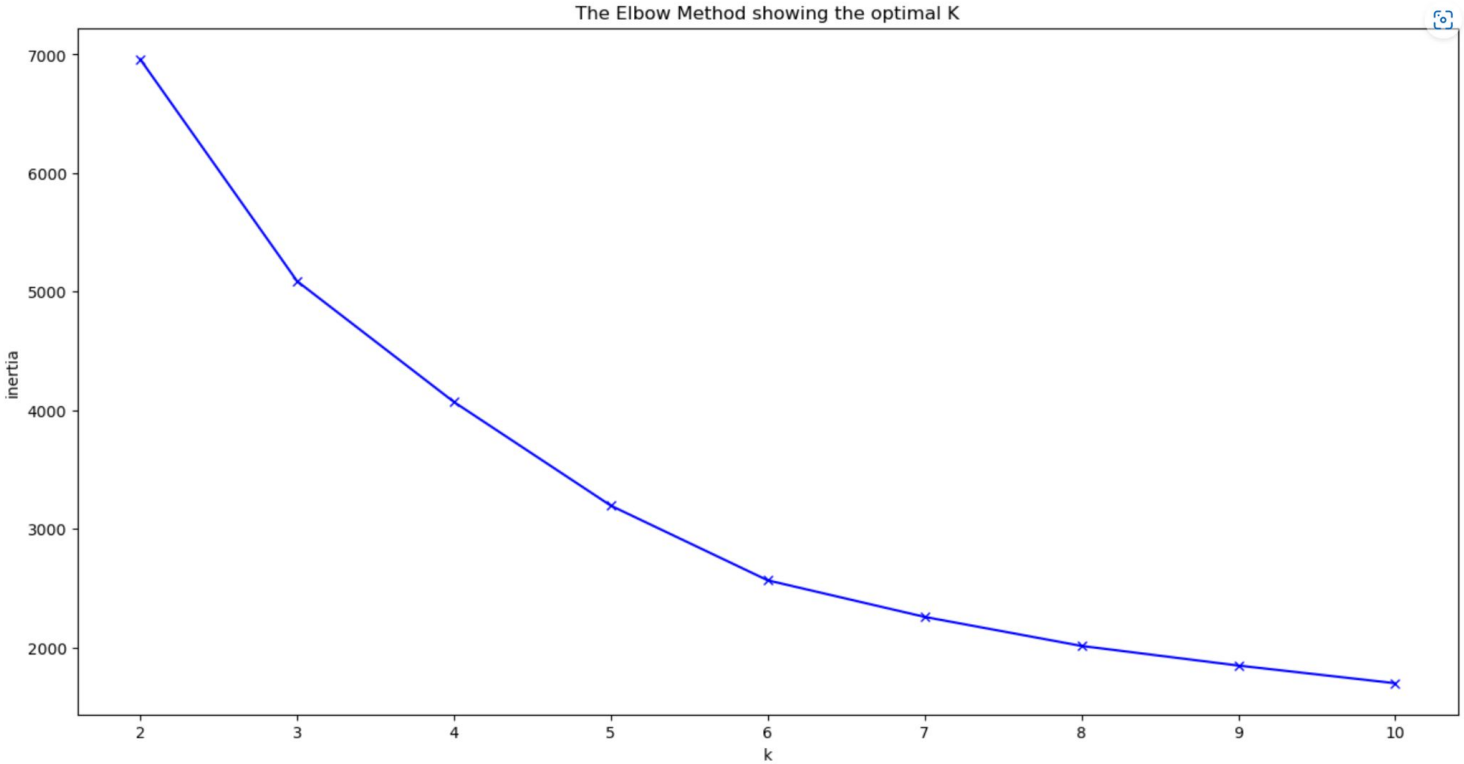




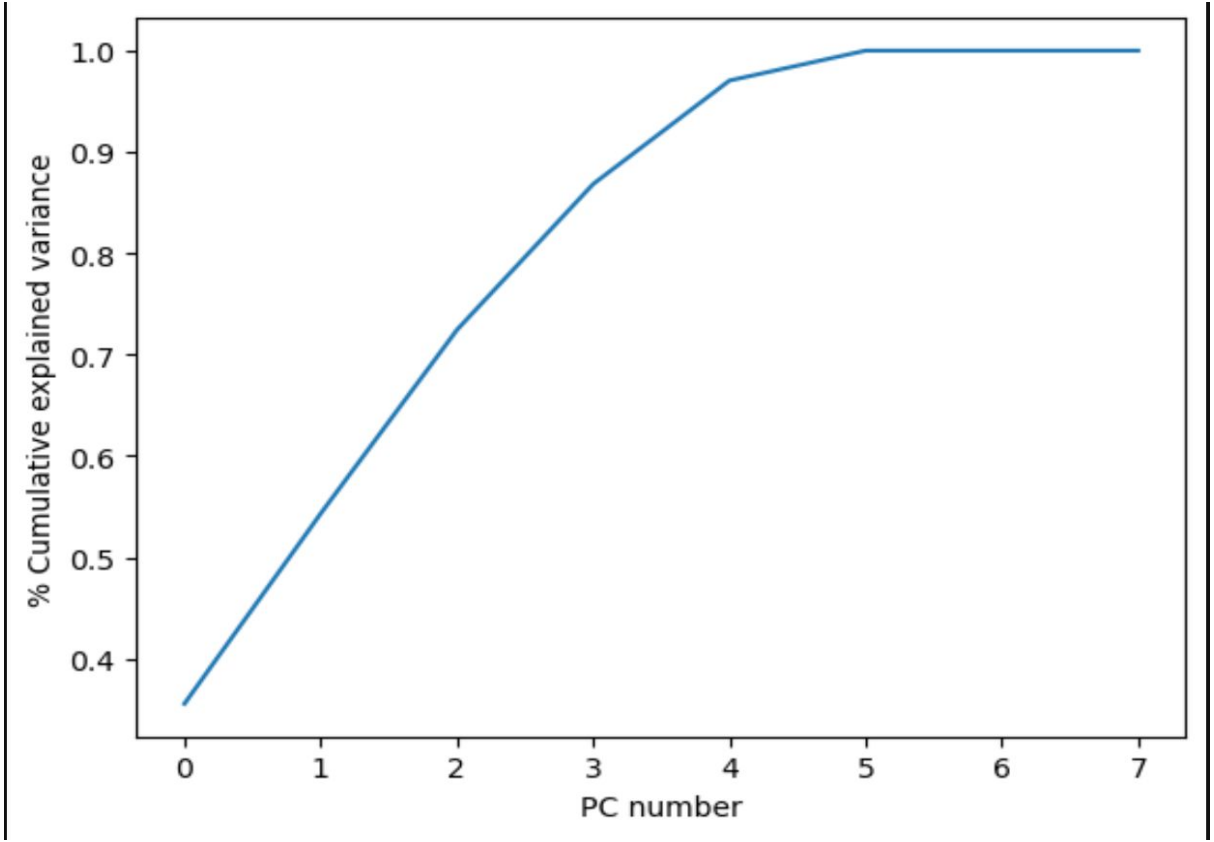
Banking Behaviour

acct_type	starting_balance	ending_balance	balance_change	num_transactions	avg_transaction_amt
SV	1430.22	284.58	-1145.64	40	-28.641000
CC	266.34	496.15	229.81	55	-4.110182
CC	55.90	1000.00	944.10	41	-19.863659
CK	11017.13	968.46	-10048.67	99	-101.501717
CC	849.37	462.28	-387.09	43	9.280233
...	...	...	...	...	...
CK	428.90	626.79	197.89	99	1.998889
CC	632.75	1400.00	767.25	49	-14.251837
CC	2810.91	3000.00	189.09	38	5.022105
SV	784.53	622.46	-162.07	21	-7.717619
CK	3649.18	107.33	-3541.85	99	-35.776263

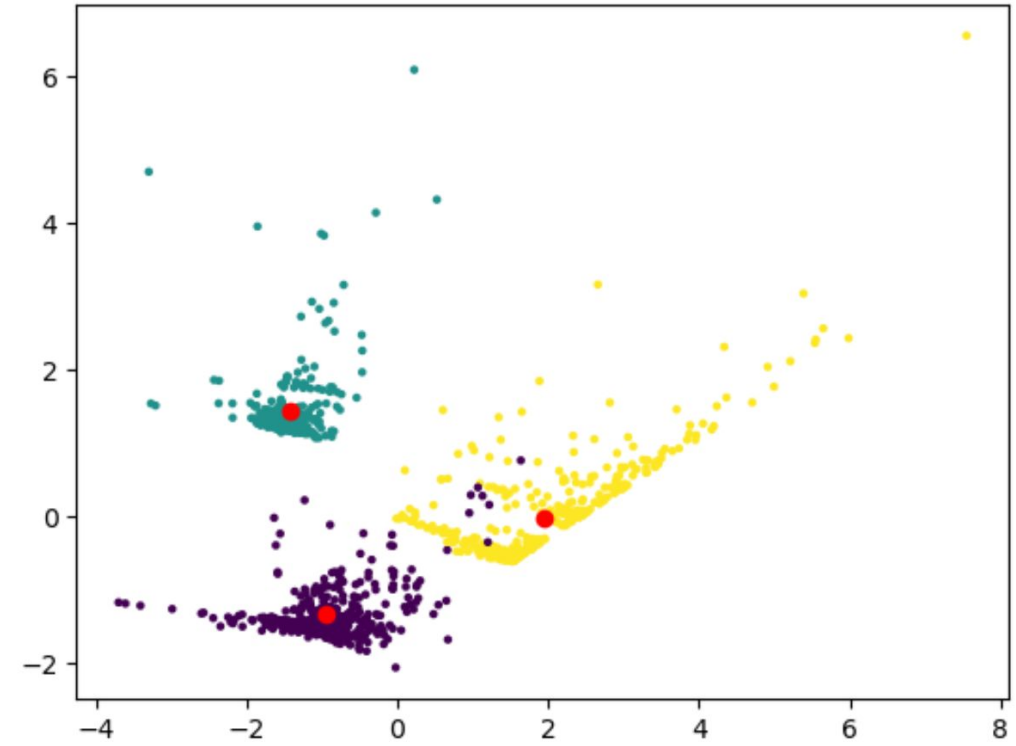
Banking Behaviour

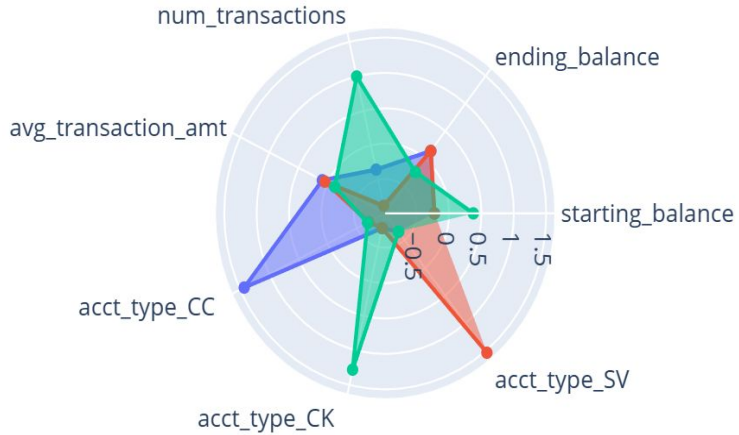


Banking Behaviour



Banking Behaviour





Future Goals

- Try other forms of clustering
  - Hierarchical Clustering
  - DB Scan

- Visualize in 2D how our clusters are evolving in each iteration of KMeans

Thank You for listening!