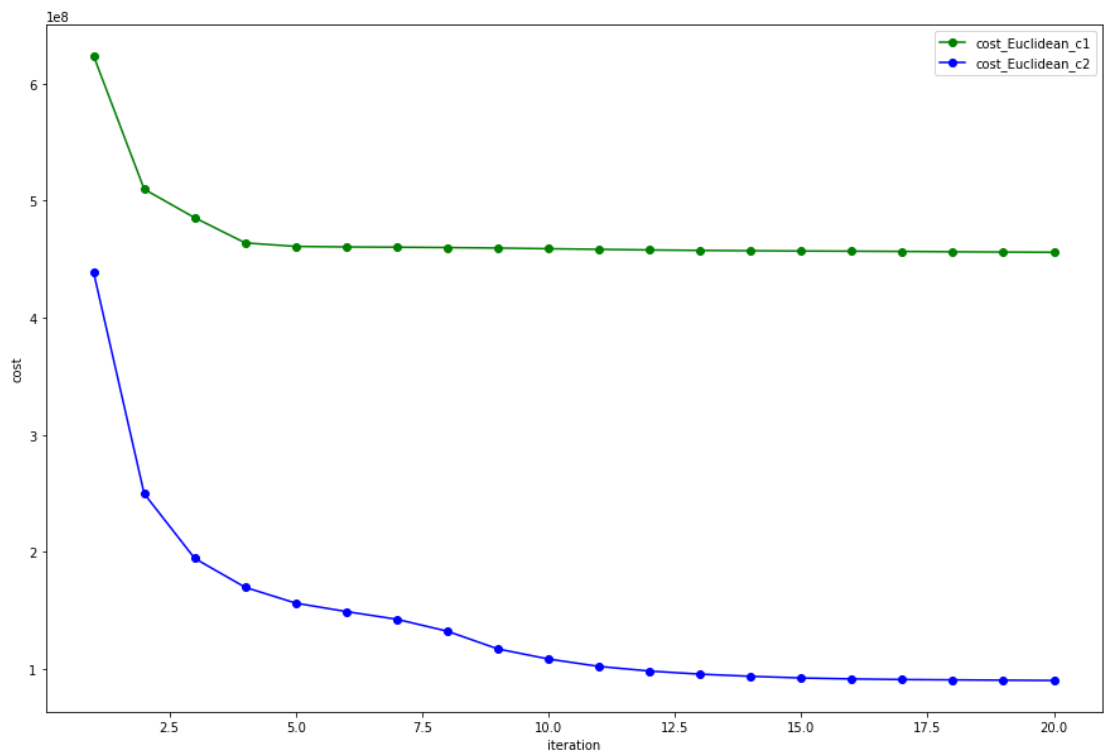


# Kmeans Report

110062643 嚴翎愷

## Euclidean Distance

- 1) A plot of cost vs. iteration for 2 initialization strategies(c1 and c2) for Euclidean distance



- 2) Percentage improvement values and your explanation for Euclidean distance

C1: 26.885383292518288 ; C2: 79.437750291599

**Reason:** 可以發現選取 C2 作為 initial centroid 比較適合。因為 C1 是取前 10 個點視為 centroid，這 10 個點可能原本就屬於同個 cluster，所以效果比較不好。但 C2 的取法所形成的 cluster 可以比較分散，故分群的效果比較好。

3) The Euclidean and Manhattan Distances for all pairs of centroids, with 2 initialization strategies:

i. C1 with Euclidean Distance

Euclidean	0	1	2	3	4	5	6	7	8	9
0	0	646.931	1615.852	167.15	99.546	1038.827	346.719	220.902	142.439	3836.907
1	646.931	0	975.32	814.076	746.336	412.076	307.669	867.823	504.634	3195.924
2	1615.852	975.32	0	1782.203	1715.253	669.89	1282.771	1835.64	1474.945	2294.58
3	167.15	814.076	1782.203	0	67.912	1204.078	512.612	53.79	309.506	4002.689
4	99.546	746.336	1715.253	67.912	0	1136.327	444.731	121.634	241.73	3934.872
5	1038.827	412.076	669.89	1204.078	1136.327	0	692.158	1257.45	897.659	2798.801
6	346.719	307.669	1282.771	512.612	444.731	692.158	0	566.202	205.75	3490.259
7	220.902	867.823	1835.64	53.79	121.634	1257.45	566.202	0	363.263	4056.136
8	142.439	504.634	1474.945	309.506	241.73	897.659	205.75	363.263	0	3695.114
9	3836.907	3195.924	2294.58	4002.689	3934.872	2798.801	3490.259	4056.136	3695.114	0

ii. C1 with Manhattan Distance

Manhattan	0	1	2	3	4	5	6	7	8	9
0	0	779.397	2102.865	204.523	125.597	1100.833	374.89	272.935	171.365	4170.305
1	779.397	0	1327.584	983.02	904.37	490.928	406.701	1050.916	609.749	3396.42
2	2102.865	1327.584	0	2306.38	2227.556	1005.293	1731.064	2374.545	1934.087	2513.423
3	204.523	983.02	2306.38	0	79.402	1303.896	577.402	69.59	375.248	4372.789
4	125.597	904.37	2227.556	79.402	0	1225.352	499.158	147.866	296.255	4294.953
5	1100.833	490.928	1005.293	1303.896	1225.352	0	728.924	1372.092	935.885	3072.889
6	374.89	406.701	1731.064	577.402	499.158	728.924	0	645.77	212.181	3797.899
7	272.935	1050.916	2374.545	69.59	147.866	1372.092	645.77	0	443.498	4440.72
8	171.365	609.749	1934.087	375.248	296.255	935.885	212.181	443.498	0	4001.038
9	4170.305	3396.42	2513.423	4372.789	4294.953	3072.889	3797.899	4440.72	4001.038	0

iii. C2 with Euclidean Distance

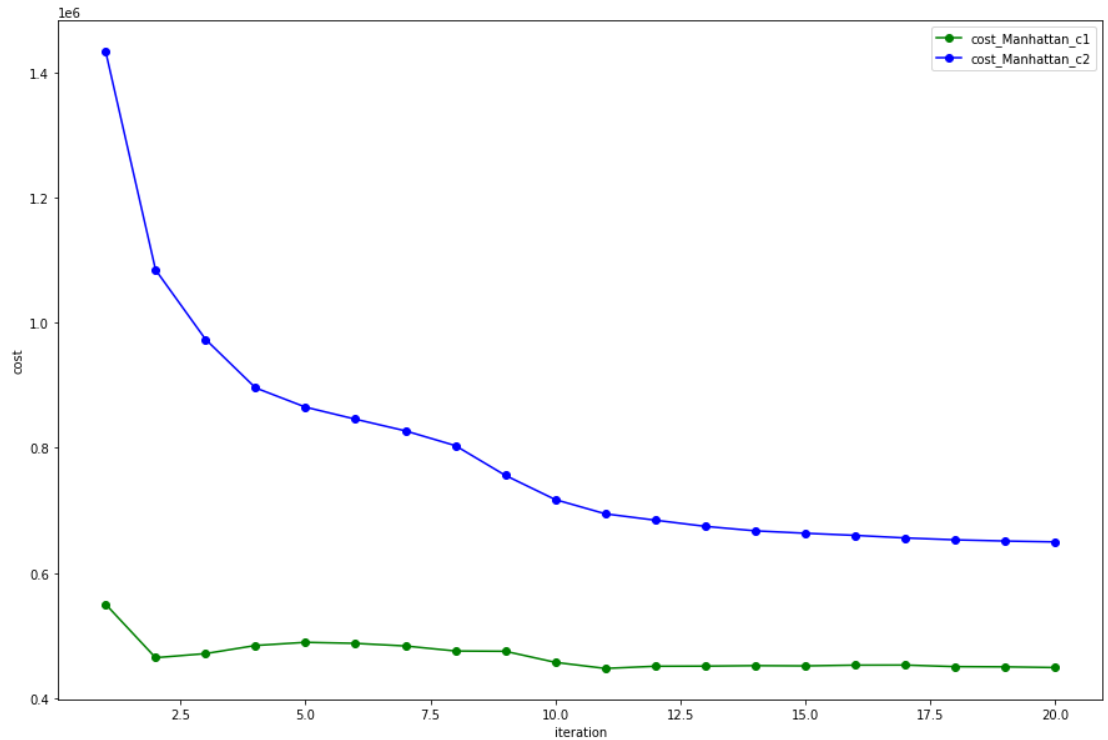
Euclidean	0	1	2	3	4	5	6	7	8	9
0	0	1100.859	2105.443	402.891	3169.004	1924.624	9045.32	15760.122	14110.834	5567.685
1	1100.859	0	1010.198	698.488	2085.461	1182.864	7957.776	14682.451	13208.003	4492.458
2	2105.443	1010.198	0	1702.793	1080.535	1313.327	6947.821	13674.708	12508.957	3488.159
3	402.891	698.488	1702.793	0	2768.608	1615.788	8644.807	15362.418	13786.484	5169.937
4	3169.004	2085.461	1080.535	2768.608	0	2153.771	5876.33	12597.04	11938.376	2407.919
5	1924.624	1182.864	1313.327	1615.788	2153.771	0	7718.222	14455.119	12233.96	4404.563
6	9045.32	7957.776	6947.821	8644.807	5876.33	7718.222	0	6743.884	9545.879	3494.222
7	15760.122	14682.451	13674.708	15362.418	12597.04	14455.119	6743.884	0	11524.506	10192.525
8	14110.834	13208.003	12508.957	13786.484	11938.376	12233.96	9545.879	11524.506	0	10883.382
9	5567.685	4492.458	3488.159	5169.937	2407.919	4404.563	3494.222	10192.525	10883.382	0

iv. C2 with Manhattan Distance

Manhattan	0	1	2	3	4	5	6	7	8	9
0	0	1311.039	2369.412	471.266	3349.657	3088.054	9533.171	15772.615	20215.646	5604.2
1	1311.039	0	1068.94	840.723	2137.788	1781.823	8228.355	14909.17	18912.605	4696.975
2	2369.412	1068.94	0	1901.209	1176.45	2162.802	7168.733	13950.576	17851.807	3737.707
3	471.266	840.723	1901.209	0	2883.735	2619.811	9065.404	15434.46	19748.936	5221.253
4	3349.657	2137.788	1176.45	2883.735	0	3337.746	6190.679	12776.883	16873.244	2564.171
5	3088.054	1781.823	2162.802	2619.811	3337.746	0	8896.389	16105.347	17509.903	5893.07
6	9533.171	8228.355	7168.733	9065.404	6190.679	8896.389	0	7219.197	10690.484	3935.293
7	15772.615	14909.17	13950.576	15434.46	12776.883	16105.347	7219.197	0	16003.499	10221.031
8	20215.646	18912.605	17851.807	19748.936	16873.244	17509.903	10690.484	16003.499	0	14613.552
9	5604.2	4696.975	3737.707	5221.253	2564.171	5893.07	3935.293	10221.031	14613.552	0

## Manhattan Distance

- 1) A plot of cost vs. iteration for 2 initialization strategies (c1 and c2) for Manhattan distance



- 2) Percentage improvement values and your explanation for Euclidean distance

C1: 18.393840107083165; C2: 54.68569434813371

Reason: 由上圖發現，在 Manhattan Distance 下，C2 的 cost 明顯大於 C1，而在 improvement 上面，相比 Euclidean Distance，在 Manhattan Distance 上選取 C2 的改善幅度沒有 Euclidean 大。

### 3) The Euclidean and Manhattan Distances for all pairs of centroids, with 2 initialization strategies:

#### i. C1 with Euclidean Distance

Euclidean	0	1	2	3	4	5	6	7	8	9
0	0	685.155	1407.269	236.438	146.795	272.536	2898.578	249.613	1391.422	10626.352
1	685.155	0	724.882	921.176	831.594	415.77	2214.964	531.301	828.932	9943.77
2	1407.269	724.882	0	1641.914	1552.742	1135.224	1491.357	1249.906	709.408	9236.84
3	236.438	921.176	1641.914	0	89.666	506.763	3133.245	457.816	1613.363	10862.753
4	146.795	831.594	1552.742	89.666	0	417.519	3044.096	375.438	1529.086	10773.141
5	272.536	415.77	1135.224	506.763	417.519	0	2626.579	221.316	1170.533	10359.504
6	2898.578	2214.964	1491.357	3133.245	3044.096	2626.579	0	2732.841	1812.455	7767.946
7	249.613	531.301	1249.906	457.816	375.438	221.316	2732.841	0	1155.797	10431.976
8	1391.422	828.932	709.408	1613.363	1529.086	1170.533	1812.455	1155.797	0	9340.275
9	10626.352	9943.77	9236.84	10862.753	10773.141	10359.504	7767.946	10431.976	9340.275	0

#### ii. C1 with Manhattan Distance

Manhattan	0	1	2	3	4	5	6	7	8	9
0	0	775.805	1500.84	287.351	177.255	278.003	3104.116	382.469	2028.739	12695.389
1	775.805	0	732.704	1061.783	952.467	499.653	2336.037	652.125	1255.347	11924.08
2	1500.84	732.704	0	1786.572	1677.155	1224.764	1605.27	1376.942	1006.368	11196.787
3	287.351	1061.783	1786.572	0	110.47	563.475	3388.744	667.483	2314.427	12978.891
4	177.255	952.467	1677.155	110.47	0	454.244	3279.855	558.134	2204.8	12870.996
5	278.003	499.653	1224.764	563.475	454.244	0	2828.265	335.664	1753.198	12419.434
6	3104.116	2336.037	1605.27	3388.744	3279.855	2828.265	0	2777.742	2380.461	9597.441
7	382.469	652.125	1376.942	667.483	558.134	335.664	2777.742	0	1653.445	12322.898
8	2028.739	1255.347	1006.368	2314.427	2204.8	1753.198	2380.461	1653.445	0	10775.939
9	12695.389	11924.08	11196.787	12978.891	12870.996	12419.434	9597.441	12322.898	10775.939	0

#### iii. C2 with Euclidean Distance

Euclidean	0	1	2	3	4	5	6	7	8	9
0	0	514.627	1571.243	1338.161	3022.661	2006.703	9032.333	15747.234	14100.145	5554.787
1	514.627	0	1081.379	827.841	2511.459	1637.729	8521.198	15239.877	13684.607	5047.516
2	1571.243	1081.379	0	566.551	1649.389	910.994	7588.405	14328.226	12643.986	4167.637
3	1338.161	827.841	566.551	0	1684.516	1405.109	7694.277	14412.057	13125.351	4219.761
4	3022.661	2511.459	1649.389	1684.516	0	2124.263	6009.82	12731.398	12006.395	2542.569
5	2006.703	1637.729	910.994	1405.109	2124.263	0	7742.628	14474.554	12167.794	4452.972
6	9032.333	8521.198	7588.405	7694.277	6009.82	7742.628	0	6743.884	9545.879	3494.222
7	15747.234	15239.877	14328.226	14412.057	12731.398	14474.554	6743.884	0	11524.506	10192.525
8	14100.145	13684.607	12643.986	13125.351	12006.395	12167.794	9545.879	11524.506	0	10883.382
9	5554.787	5047.516	4167.637	4219.761	2542.569	4452.972	3494.222	10192.525	10883.382	0

#### iv. C2 with Manhattan Distance

Manhattan	0	1	2	3	4	5	6	7	8	9
0	0	602.955	2102.554	1430.209	3211.456	3281.488	9517.668	15757.691	20200.259	5588.854
1	602.955	0	1500.825	833.43	2613.997	2682.569	8918.813	15335.957	19602.263	5123.067
2	2102.554	1500.825	0	674.828	2062.251	1358.796	7771.222	14980.056	18111.885	4768.923
3	1430.209	833.43	674.828	0	1784.512	1855.58	8090.51	14506.486	18775.121	4293.502
4	3211.456	2613.997	2062.251	1784.512	0	3413.036	6312.53	12922.931	16995.134	2710.057
5	3281.488	2682.569	1358.796	1855.58	3413.036	0	9116.025	16325.271	17521.518	6110.832
6	9517.668	8918.813	7771.222	8090.51	6312.53	9116.025	0	7219.197	10690.484	3935.293
7	15757.691	15335.957	14980.056	14506.486	12922.931	16325.271	7219.197	0	16003.499	10221.031
8	20200.259	19602.263	18111.885	18775.121	16995.134	17521.518	10690.484	16003.499	0	14613.552
9	5588.854	5123.067	4768.923	4293.502	2710.057	6110.832	3935.293	10221.031	14613.552	0