Contents

Tournament Analysis

Mean Scores and Coefficient of Variation

The tournament is run by considering different number of dancers and friends. Mean scores and coefficient of variation (standard deviation / mean) of for various teams is shown in figure 1.

		mean score									coefficient of variation								
group		g1	g2	g3	g4	g5	g6	g7	g8	g9	g1	g2	g3	g4	g5	g6	g7	g8	g9
dancers	friends_ratio																		
100	0%	9786	10206	4500	10230	10038	10245	10235	10127	10222	1%	0%	0%	0%	1%	0%	0%	1%	0%
100	25%	9874	10248	4878	10256	10069	10247	10229	10123	10216	1%	0%	1%	0%	2%	0%	0%	1%	0%
100	50%	9912	10262	5416	10246	10071	10250	10239	10199	10271	1%	0%	1%	0%	1%	0%	0%	0%	0%
100	75%	9917	10246	6010	10272	10191	10281	10242	10210	10251	1%	0%	1%	0%	1%	0%	0%	1%	0%
200	0%	8902	9636	5145	9669	9243	9678	8690	9311	9649	2%	0%	0%	1%	1%	1%	32%	1%	1%
200	25%	8989	9671	5414	9691	8487	9698	9666	9359	9664	1%	0%	0%	1%	31%	0%	0%	1%	0%
200	50%	8937	9708	5838	9677	8499	9735	8790	9446	9689	1%	1%	0%	1%	30%	0%	32%	1%	0%
200	75%	9085	9696	6292	9728	8591	9780	9726	9457	9716	1%	1%	0%	0%	29%	0%	0%	1%	1%
400	0%	6946	8497	5145	8467	7362	8601	8535	7896	8427	3%	1%	0%	1%	6%	1%	1%	2%	1%
400	25%	7123	8476	5378	8519	6738	8578	8587	8069	8517	2%	1%	0%	1%	27%	1%	1%	3%	1%
400	50%	7144	8571	5766	8613	7570	8616	8628	8071	8599	2%	1%	0%	1%	6%	1%	1%	2%	1%
400	75%	7316	8621	6204	8676	6879	8699	8691	8115	8632	3%	1%	1%	1%	25%	1%	1%	2%	0%
800	0%	5103	6114	5145	6178	5145	6350	5624	5022	6056	0%	1%	0%	2%	0%	2%	26%	5%	1%
800	25%	5487	6200	5359	6272	5506	6407	5307	5298	6181	1%	1%	0%	3%	1%	1%	54%	5%	2%
800	50%	5968	6365	5733	6360	5985	6424	4971	5325	6311	1%	2%	0%	2%	1%	2%	60%	5%	1%
800	75%	4909	6427	6167	6460	6473	6617	6485	5590	6508	92%	2%	0%	1%	0%	1%	1%	4%	1%
1600	0%	5103	4905	5145	5085	5145	5076	5130	2004	5130	0%	0%	0%	0%	0%	0%	0%	0%	0%
1600	25%	5441	5114	5319	5313	5488	5310	5503	2072	5318	0%	0%	1%	1%	0%	0%	0%	1%	0%
1600	50%	4470	5536	5677	5650	5976	5656	5998	2068	5685	98%	1%	1%	1%	1%	1%	1%	1%	0%
1600	75%	3364	3065	6111	5988	6434	6023	6412	2084	6095	179%	95%	1%	1%	0%	1%	1%	1%	1%
3200	0%	1224	2502	0	1554	1928	1319	2052	0	1328	0%	0%		0%	1%	0%	0%		0%
3200	25%	1226	2538	0	1581	2108	1419	2092	0	1420	0%	0%		1%	1%	0%	1%		1%
3200	50%	1297	2533	0	1677	2291	1528	2316	0	1527	1%	0%		1%	1%	0%	1%		1%
3200	75%	1403	2534	0	1819	2480	1633	2620	0	1640	1%	0%		1%	1%	0%	1%		1%
6400	0%	456	837	0	756	942	589	1023	0	584	0%	0%		0%	0%	1%	0%		0%
6400	25%	456	916	0	757	988	637	1029	0	609	0%	1%		1%	1%	0%	1%		0%
6400	50%	456	986	0	804	1051	686	1084	0	655	0%	0%		1%	2%	0%	1%		0%
12800	25%	153	416	0	153	258	257	375	0	255	0%	0%		0%	0%	0%	0%		0%
12800	50%	153	412	0	162	284	277	390	0	275	0%	0%		0%	0%	0%	0%		0%
25600	0%	-880	0	0	-7944	84	60	78	0	60	0%			0%	0%	0%	0%		0%

Figure 1: Mean Scores and Coefficient of Variation

We note that the variation for configurations is generally pretty low ($< 1\%$). For our group, g2, we notice a huge variation of 95% for 1600 dancers and 75% friend density. On further investigating we notice the following distribution for the given bucket:

```
df1[(df1.dancers == 1600)&(df1.friends == 1200)&(df1.group == 'g2')]
```

	dancers	friends	group	score
1791	1600	1200	g2	6024
1792	1600	1200	g2	141
1793	1600	1200	g2	5925
1794	1600	1200	g2	141
1795	1600	1200	g2	141
1796	1600	1200	g2	5967
1797	1600	1200	g2	141
1798	1600	1200	g2	141
1799	1600	1200	g2	6017
1800	1600	1200	g2	6016

We try to estimate the friend density for our Medium Strategy and based on it, we dance with partners for longer duration of time if we think we have more friends. This gamble does not seem to pay off. We get a competitive score of 5989.80 in half the scenarios and a meagre 141 for the rest. We can interpret the result it to state that there is still a 50% chance that the worst case dancer will not be able to find friends and dance with them. Because of the uncertainity involved, we would have been better off just ignoring friends and trying to ensure that every dancer dances.

We also notice that other players like g1 and g7 also see a huge variation because of claustrophobia. Other high variations are due to single outliers (on the worse side) in data.

```
df1[df1.score < 0]
```

	dancers	friends	group	score
1360	800	200	g7	-3230
1437	800	400	g7	-2776
1468	800	600	g1	-8686
1702	1600	800	g1	-8688
1781	1600	1200	g1	-8687
1782	1600	1200	g1	-8686
2388	25600	0	g1	-880
2391	25600	0	g4	-7944

Team wise Scores

After noting that generally the variations are low and that scores increasing with ratio of friends as a general trend, we decide to take an average across number

of friends to evaluate the performance of each team. We also categorize the tournament in three categories:

- Small
  - for $d <= 800$
  - It is possible to systematically find soulmates
- Medium
  - for $800 < d <= 1600$
  - It is possible for all dancers to dance at the same time on the dance floor
- Large
  - for $1600 < d$
  - We require some sort of scheduling to ensure each of the dancer gets to dance
  - none of the dancers should suffer claustrophobia

	score									
group dancers	g2	g7	g6	g9	g5	g1	g4	g8	g3	Average if gt 0
100 dancers	10240	10236	10256	10240	10092	9872	10251	10165	5201	9617
200 dancers	9678	9218	9723	9679	8705	8978	9691	9393	5672	8971
400 dancers	8541	8610	8624	8544	7137	7132	8569	8038	5623	7869
800 dancers	6276	5597	6450	6264	5777	5367	6318	5309	5601	5884
1600 dancers	4655	5761	5516	5557	5761	4594	5509	2057	5563	4997
3200 dancers	2527	2270	1474	1479	2202	1287	1658	0	0	1842
6400 dancers	928	1049	647	615	993	456	776	0	0	781
12800 dancers	414	383	267	265	271	153	158	0	0	273
25600 dancers	0	78	60	60	84	-880	-7944	0	0	71
Average	**4807**	**4800**	**4780**	**4745**	**4558**	**4107**	**3887**	**3885**	**3073**	

Figure 2: Scores of teams per number of dancers

Figure 2 shows the results for each of the configuration. We color each row with a heat map, blue being better score. **Our group (g2) scores the best Average**, closely followed by g7. Next best score is by g6 and g9, which share a rather noticable correlation! We notice the following kinds of teams:

- g1 and g7
  - score well on all categories
  - thus score highest and second higher respectively
- g6, g9, g4
  - score well on small category
- g5
  - scores well on Large category

– but fails to perform on Small

We normalize the scores based on an average of positive scores in each category in 3 and plot a curve to see the development of scores with number of dancers for the top 5 teams as shown in figure 4.

	relative score									
group dancers	g2	g7	g6	g9	g5	g1	g4	g8	g3	
100 dancers	1.06	1.06	1.07	1.06	1.05	1.03	1.07	1.06	0.54	1.00
200 dancers	1.08	1.03	1.08	1.08	0.97	1.00	1.08	1.05	0.63	1.00
400 dancers	1.09	1.09	1.10	1.09	0.91	0.91	1.09	1.02	0.71	1.00
800 dancers	1.07	0.95	1.10	1.06	0.98	0.91	1.07	0.90	0.95	1.00
1600 dancers	0.93	1.15	1.10	1.11	1.15	0.92	1.10	0.41	1.11	1.00
3200 dancers	1.37	1.23	0.80	0.80	1.20	0.70	0.90	0.00	0.00	1.00
6400 dancers	1.19	1.34	0.83	0.79	1.27	0.58	0.99	0.00	0.00	1.00
12800 dancers	1.52	1.40	0.98	0.97	0.99	0.56	0.58	0.00	0.00	1.00
25600 dancers	0.00	1.11	0.85	0.85	1.19	-12.48	-112.68	0.00	0.00	1.00

Figure 3: Relative scores of teams per number of dancers

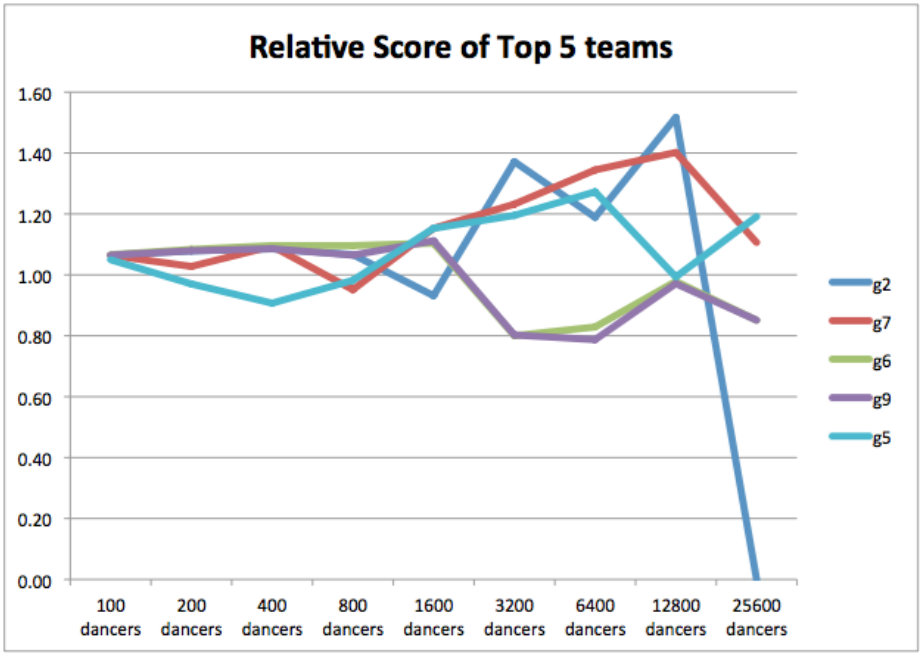

Figure 4: Plot of relative scores of top 5 teams

As you may notice, the strength of our group is the relatively stronger scores in Large category. But to decide an ordering, we still follow a simple average, which seems to favor Small category. Even so, we achieve the highest overall average, which is testimony to our performance across categories.

For Large category, our strategy has step behavior. Where as g7 has a more continuous looking behavior. We present the values of number of dancers vs maximum dancers without creating claustrophobia.