Machine Learning techniques to Model Data Intensive Application Performance

A. Battistello P. Ferretti

Contents

1 Fixed Datasize, all the features **6**
1.1 Query R1 . . . 6
1.1.1 Query R1 — Datasize 250GB . . . 6
1.1.2 Query R1 — Datasize 500GB . . . 7
1.1.3 Query R1 — Datasize 750GB . . . 8
1.1.4 Query R1 — Datasize 1000GB . . . 9
1.2 Query R2 . . . 10
1.2.1 Query R2 — Datasize 250GB . . . 10
1.2.2 Query R2 — Datasize 500GB . . . 11
1.2.3 Query R2 — Datasize 750GB . . . 12
1.2.4 Query R2 — Datasize 1000GB . . . 13
1.3 Query R3 . . . 14
1.3.1 Query R3 — Datasize 250GB . . . 14
1.3.2 Query R3 — Datasize 500GB . . . 15
1.3.3 Query R3 — Datasize 750GB . . . 16
1.3.4 Query R3 — Datasize 1000GB . . . 17
1.4 Query R4 . . . 18
1.4.1 Query R4 — Datasize 250GB . . . 18
1.4.2 Query R4 — Datasize 500GB . . . 19
1.4.3 Query R4 — Datasize 750GB . . . 20
1.4.4 Query R4 — Datasize 1000GB . . . 21
1.5 Query R5 . . . 22
1.5.1 Query R5 — Datasize 250GB . . . 22
1.5.2 Query R5 — Datasize 500GB . . . 23
1.5.3 Query R5 — Datasize 750GB . . . 24
1.5.4 Query R5 — Datasize 1000GB . . . 25

2 Fixed Datasize, linear *ncores*, only *ncores* **26**
2.1 Query R1 . . . 26
2.1.1 Query R1 — Datasize 250GB . . . 26
2.1.2 Query R1 — Datasize 500GB . . . 27
2.1.3 Query R1 — Datasize 750GB . . . 28

2.1.4 Query R1 — Datasize 1000GB 29
2.2 Query R2 30
2.2.1 Query R2 — Datasize 250GB 30
2.2.2 Query R2 — Datasize 500GB 31
2.2.3 Query R2 — Datasize 750GB 32
2.2.4 Query R2 — Datasize 1000GB 33
2.3 Query R3 34
2.3.1 Query R3 — Datasize 250GB 34
2.3.2 Query R3 — Datasize 500GB 35
2.3.3 Query R3 — Datasize 750GB 36
2.3.4 Query R3 — Datasize 1000GB 37
2.4 Query R4 38
2.4.1 Query R4 — Datasize 250GB 38
2.4.2 Query R4 — Datasize 500GB 39
2.4.3 Query R4 — Datasize 750GB 40
2.4.4 Query R4 — Datasize 1000GB 41
2.5 Query R5 42
2.5.1 Query R5 — Datasize 250GB 42
2.5.2 Query R5 — Datasize 500GB 43
2.5.3 Query R5 — Datasize 750GB 44
2.5.4 Query R5 — Datasize 1000GB 45
2.5.5 Query R5 — Datasize 1000GB 45

3 Fixed Datasize, $ncores^{-1}$, all the features 46
3.1 Query R1 46
3.1.1 Query R1 — Datasize 250GB 46
3.1.2 Query R1 — Datasize 500GB 47
3.1.3 Query R1 — Datasize 750GB 48
3.1.4 Query R1 — Datasize 1000GB 49
3.2 Query R2 50
3.2.1 Query R2 — Datasize 250GB 50
3.2.2 Query R2 — Datasize 500GB 51
3.2.3 Query R2 — Datasize 750GB 52
3.2.4 Query R2 — Datasize 1000GB 53
3.3 Query R3 54
3.3.1 Query R3 — Datasize 250GB 54
3.3.2 Query R3 — Datasize 500GB 55
3.3.3 Query R3 — Datasize 750GB 56
3.3.4 Query R3 — Datasize 1000GB 57
3.4 Query R4 58
3.4.1 Query R4 — Datasize 250GB 58
3.4.2 Query R4 — Datasize 500GB 59
3.4.3 Query R4 — Datasize 750GB 60
3.4.4 Query R4 — Datasize 1000GB 61

3.5 Query R5 . . . 62
3.5.1 Query R5 — Datasize 250GB . . . 62
3.5.2 Query R5 — Datasize 500GB . . . 63
3.5.3 Query R5 — Datasize 750GB . . . 64
3.5.4 Query R5 — Datasize 1000GB . . . 65

4 Fixed Datasize, $ncores^{-1}$, only $ncores$ 66
4.1 Query R1 . . . 66
4.1.1 Query R1 — Datasize 250GB . . . 66
4.1.2 Query R1 — Datasize 500GB . . . 67
4.1.3 Query R1 — Datasize 750GB . . . 68
4.1.4 Query R1 — Datasize 1000GB . . . 69
4.2 Query R2 . . . 70
4.2.1 Query R2 — Datasize 250GB . . . 70
4.2.2 Query R2 — Datasize 500GB . . . 71
4.2.3 Query R2 — Datasize 750GB . . . 72
4.2.4 Query R2 — Datasize 1000GB . . . 73
4.3 Query R3 . . . 74
4.3.1 Query R3 — Datasize 250GB . . . 74
4.3.2 Query R3 — Datasize 500GB . . . 75
4.3.3 Query R3 — Datasize 750GB . . . 76
4.3.4 Query R3 — Datasize 1000GB . . . 77
4.4 Query R4 . . . 78
4.4.1 Query R4 — Datasize 250GB . . . 78
4.4.2 Query R4 — Datasize 500GB . . . 79
4.4.3 Query R4 — Datasize 750GB . . . 80
4.4.4 Query R4 — Datasize 1000GB . . . 81
4.5 Query R5 . . . 82
4.5.1 Query R5 — Datasize 250GB . . . 82
4.5.2 Query R5 — Datasize 500GB . . . 83
4.5.3 Query R5 — Datasize 750GB . . . 84
4.5.4 Query R5 — Datasize 1000GB . . . 85

5 Query comparison 86
5.1 R5 → R2 (Fast - Fast) . . . 86
5.1.1 R5 → R2 — Datasize 250GB . . . 86
5.1.2 R5 → R2 — Datasize 500GB . . . 87
5.1.3 R5 → R2 — Datasize 750GB . . . 88
5.1.4 R5 → R2 — Datasize 1000GB . . . 89
5.2 R2 → R5 (Fast - Fast) . . . 90
5.2.1 R2 → R5 — Datasize 250GB . . . 90
5.2.2 R2 → R5 — Datasize 500GB . . . 91
5.2.3 R2 → R5 — Datasize 750GB . . . 92
5.2.4 R2 → R5 — Datasize 1000GB . . . 93

5.3 R3 → R4 (Slow - Slow) . . . 94
5.3.1 R3 → R4 — Datasize 250GB . . . 94
5.3.2 R3 → R4 — Datasize 500GB . . . 95
5.3.3 R3 → R4 — Datasize 750GB . . . 96
5.3.4 R3 → R4 — Datasize 1000GB . . . 97
5.4 R4 → R3 (Slow - Slow) . . . 98
5.4.1 R4 → R3 — Datasize 250GB . . . 98
5.4.2 R4 → R3 — Datasize 500GB . . . 99
5.4.3 R4 → R3 — Datasize 750GB . . . 100
5.4.4 R4 → R3 — Datasize 1000GB . . . 101
5.5 R1, R2, R4 → R3 (Fast and Slow - Average) . . . 102
5.5.1 R1,R2,R4 → R3 — Datasize 250GB . . . 102
5.5.2 R1,R2,R4 → R3 — Datasize 500GB . . . 103
5.5.3 R1,R2,R4 → R3 — Datasize 750GB . . . 104
5.5.4 R1,R2,R4 → R3 — Datasize 1000GB . . . 105
5.6 R2, R5 → R3 (Fast - Slow) . . . 106
5.6.1 R2,R5 → R3 — Datasize 250GB . . . 106
5.6.2 R2,R5 → R3 — Datasize 500GB . . . 107
5.6.3 R2,R5 → R3 — Datasize 750GB . . . 108
5.6.4 R2,R5 → R3 — Datasize 1000GB . . . 109
5.7 R3, R4 → R2 (Fast - Slow) . . . 110
5.7.1 R3,R4 → R2 — Datasize 250GB . . . 110
5.7.2 R3,R4 → R2 — Datasize 500GB . . . 111
5.7.3 R3,R4 → R2 — Datasize 750GB . . . 112
5.7.4 R3,R4 → R2 — Datasize 1000GB . . . 113

6 Fixed Datasize, all the features, 115
6.1 Query Q2 . . . 115
6.1.1 Query Q2 — Datasize 40GB . . . 115
6.1.2 Query Q2 — Datasize 50GB . . . 116
6.2 Query Q3 . . . 117
6.2.1 Query Q3 — Datasize 50GB . . . 117
6.3 Query Q4 . . . 118
6.3.1 Query Q4 — Datasize 50GB . . . 118

7 Fixed Datasize, only nCores, 119
7.1 Query Q2 . . . 119
7.1.1 Query Q2 — Datasize 40GB . . . 119
7.1.2 Query Q2 — Datasize 50GB . . . 120
7.2 Query Q3 . . . 121
7.2.1 Query Q3 — Datasize 50GB . . . 121
7.3 Query Q4 . . . 122
7.3.1 Query Q4 — Datasize 50GB . . . 122

8 Fixed Datasize, all features with 1/nCores, 123
8.1 Query Q2 . . . 123
8.1.1 Query Q2 — Datasize 40GB . . . 123
8.1.2 Query Q2 — Datasize 50GB . . . 124
8.2 Query Q3 . . . 125
8.2.1 Query Q3 — Datasize 50GB . . . 125
8.3 Query Q4 . . . 126
8.3.1 Query Q4 — Datasize 50GB . . . 126

9 Fixed Datasize, only 1/nCores, 127
9.1 Query Q2 . . . 127
9.1.1 Query Q2 — Datasize 40GB . . . 127
9.1.2 Query Q2 — Datasize 50GB . . . 128
9.2 Query Q3 . . . 129
9.2.1 Query Q3 — Datasize 50GB . . . 129
9.3 Query Q4 . . . 130
9.3.1 Query Q4 — Datasize 50GB . . . 130

1 Fixed Datasize, all the features

1.1 Query R1

1.1.1 Query R1 — Datasize 250GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.3041	0.9591	2480	0.0329
Linear SVR	0.3418	0.9830	2820	0.0375
Polynomial SVR (2)	0.7566	0.7509	10643	0.2065
Polynomial SVR (3)	4.2307	0.7818	23208	0.2408
Polynomial SVR (4)	1.1959	0.4622	11066	0.1907
Polynomial SVR (6)	5.3772	0.4964	37535	0.4284
Gaussian SVR	1.2070	0.4548	7808	0.0864

Table 1: Results for R1-250

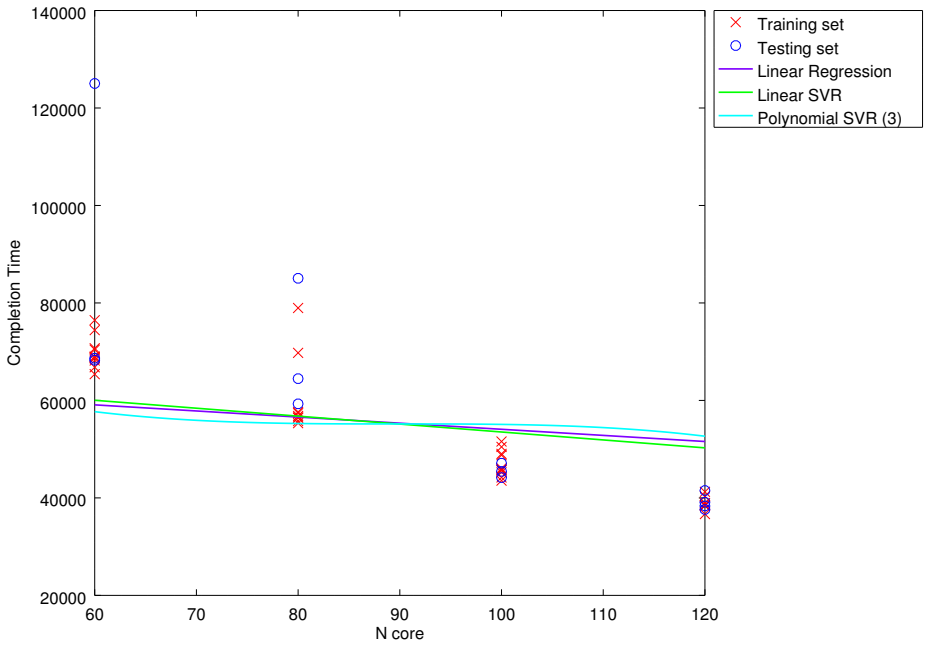

Figure 1: Completion time vs ncores for query R1 with datasize 250GB

1.1.2 Query R1 — Datasize 500GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.0905	0.9954	5714	0.0293
Linear SVR	0.0994	0.9979	6740	0.0356
Polynomial SVR (2)	0.4862	0.9173	30727	0.1741
Polynomial SVR (3)	0.2537	0.9662	13302	0.0643
Polynomial SVR (4)	0.6772	0.8746	42865	0.2285
Polynomial SVR (6)	1.3799	0.6979	67886	0.3128
Gaussian SVR	0.3571	0.9779	17589	0.0756

Table 2: Results for R1-500

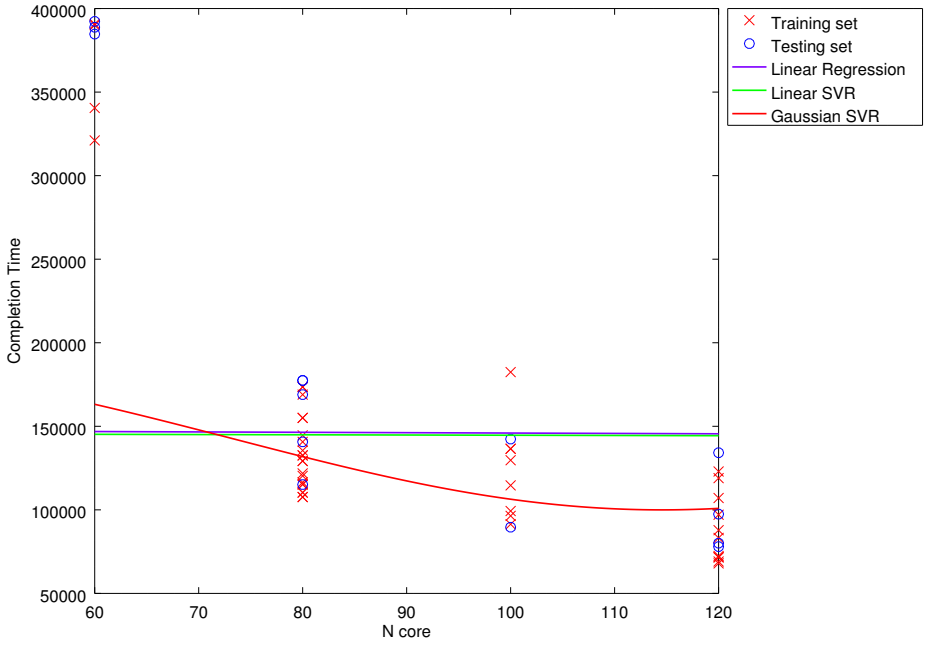

Figure 2: Completion time vs ncores for query R1 with datasize 500GB

1.1.3 Query R1 — Datasize 750GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.1813	0.9216	12361	0.0474
Linear SVR	0.1727	0.9390	11060	0.0413
Polynomial SVR (2)	0.6893	0.1868	47136	0.1980
Polynomial SVR (3)	0.4048	0.6131	27687	0.1097
Polynomial SVR (4)	0.5559	0.2879	39023	0.1624
Polynomial SVR (6)	0.5809	0.2267	41711	0.1717
Gaussian SVR	0.1855	0.9592	12583	0.0531

Table 3: Results for R1-750

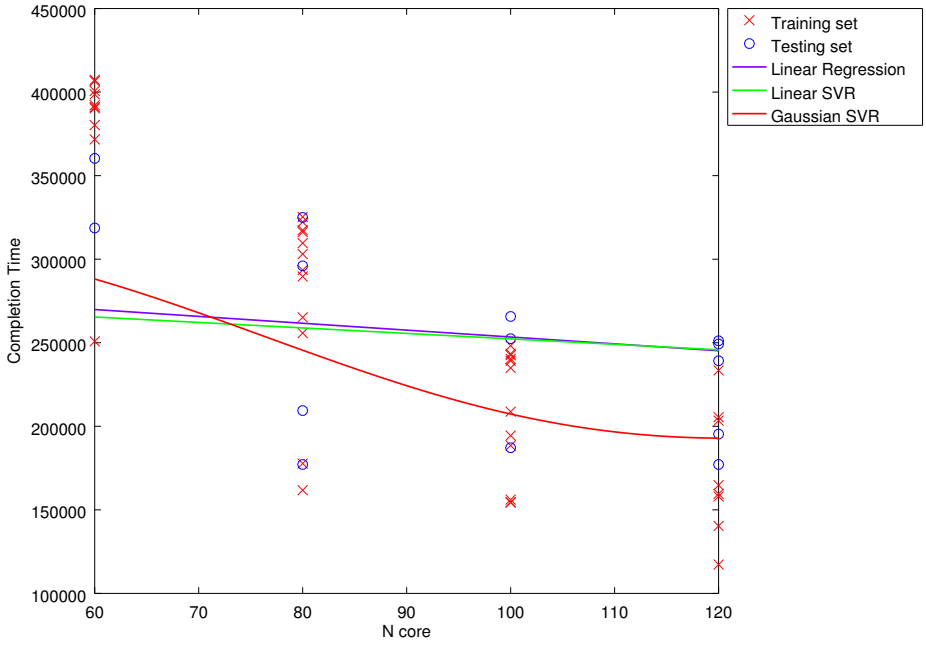

Figure 3: Completion time vs ncores for query R1 with datasize 750GB

1.1.4 Query R1 — Datasize 1000GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.1784	0.9416	12742	0.0332
Linear SVR	0.1914	0.9451	13208	0.0343
Polynomial SVR (2)	0.6728	0.4426	51407	0.1411
Polynomial SVR (3)	0.4131	0.7854	31112	0.0836
Polynomial SVR (4)	1.0215	0.0334	55894	0.1605
Polynomial SVR (6)	1.1153	0.0093	60608	0.1735
Gaussian SVR	0.3595	0.8131	15070	0.0441

Table 4: Results for R1-1000

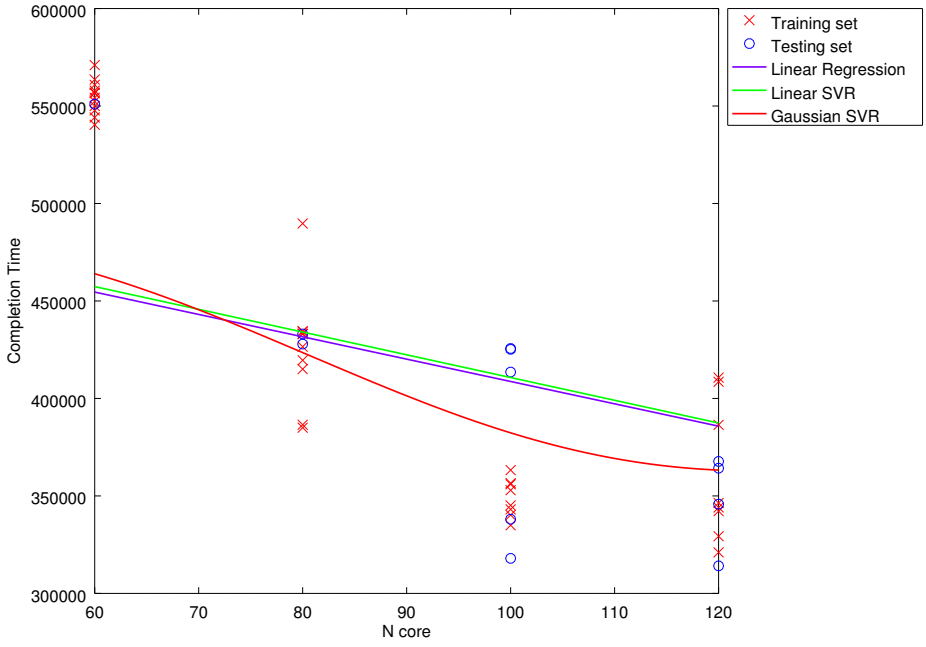

Figure 4: Completion time vs ncores for query R1 with datasize 1000GB

1.2 Query R2

1.2.1 Query R2 — Datasize 250GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.2914	0.9080	609	0.0073
Linear SVR	0.2882	0.9215	622	0.0075
Polynomial SVR (2)	1.0548	0.3165	2442	0.0303
Polynomial SVR (3)	0.5157	0.8003	1189	0.0149
Polynomial SVR (4)	0.9650	0.1946	2218	0.0276
Polynomial SVR (6)	1.0183	0.1074	2319	0.0289
Gaussian SVR	0.4247	0.8933	1021	0.0126

Table 5: Results for R2-250

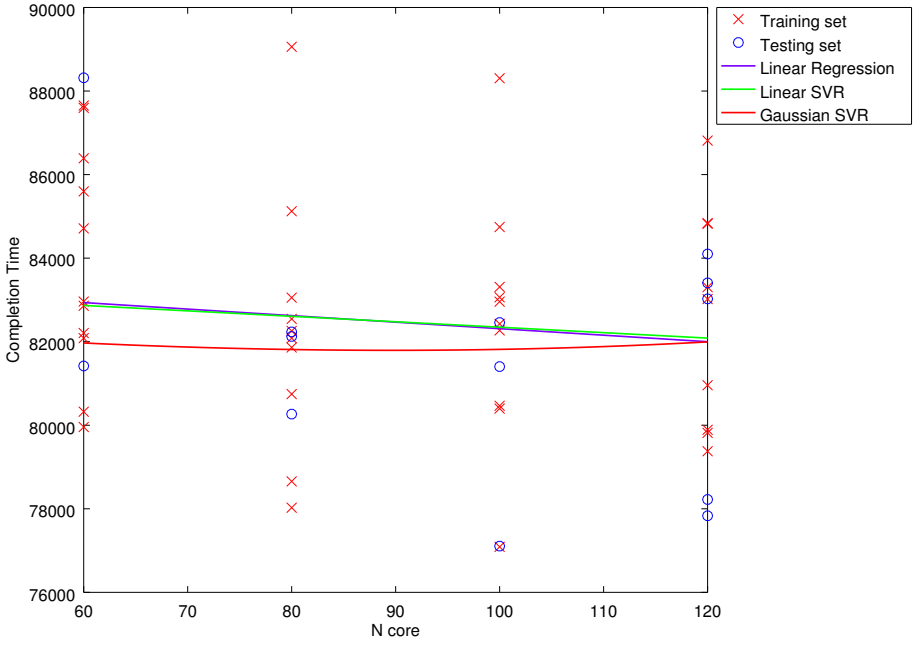

Figure 5: Completion time vs ncores for query R2 with datasize 250GB

1.2.2 Query R2 — Datasize 500GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.1427	0.9539	278	0.0039
Linear SVR	0.1346	0.9662	279	0.0039
Polynomial SVR (2)	0.6942	0.2101	1256	0.0178
Polynomial SVR (3)	0.3123	0.8255	643	0.0089
Polynomial SVR (4)	0.6433	0.5642	1228	0.0174
Polynomial SVR (6)	0.7042	0.4659	1277	0.0181
Gaussian SVR	0.2514	0.8786	574	0.0079

Table 6: Results for R2-500

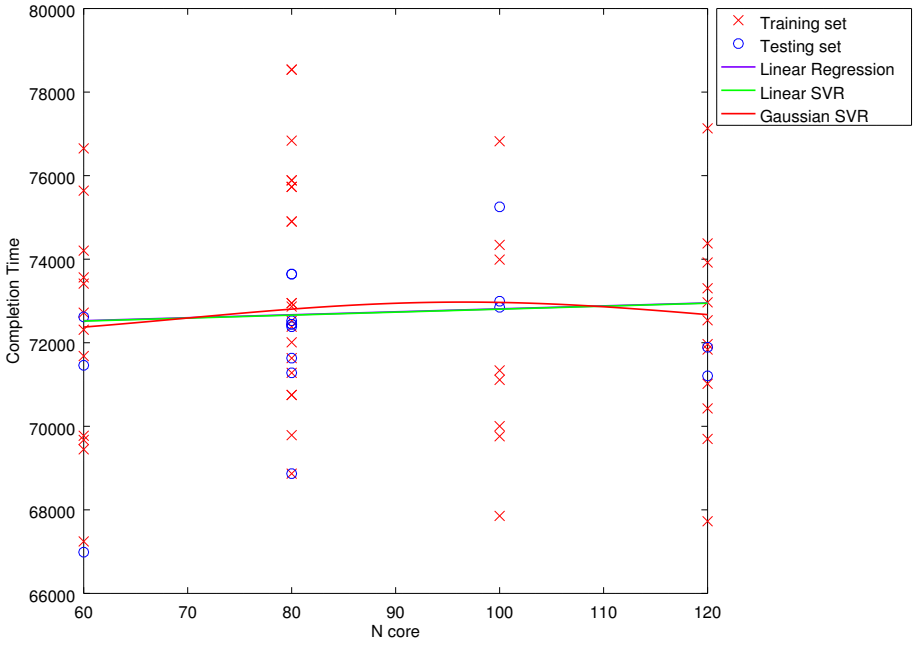

Figure 6: Completion time vs ncores for query R2 with datasize 500GB

1.2.3 Query R2 — Datasize 750GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.2345	0.9596	469	0.0059
Linear SVR	0.2177	0.9748	445	0.0056
Polynomial SVR (2)	1.2193	0.0000	2315	0.0292
Polynomial SVR (3)	0.4706	0.8904	896	0.0114
Polynomial SVR (4)	1.2193	0.0000	2315	0.0292
Polynomial SVR (6)	1.2193	0.0000	2315	0.0292
Gaussian SVR	0.5978	0.8044	1099	0.0138

Table 7: Results for R2-750

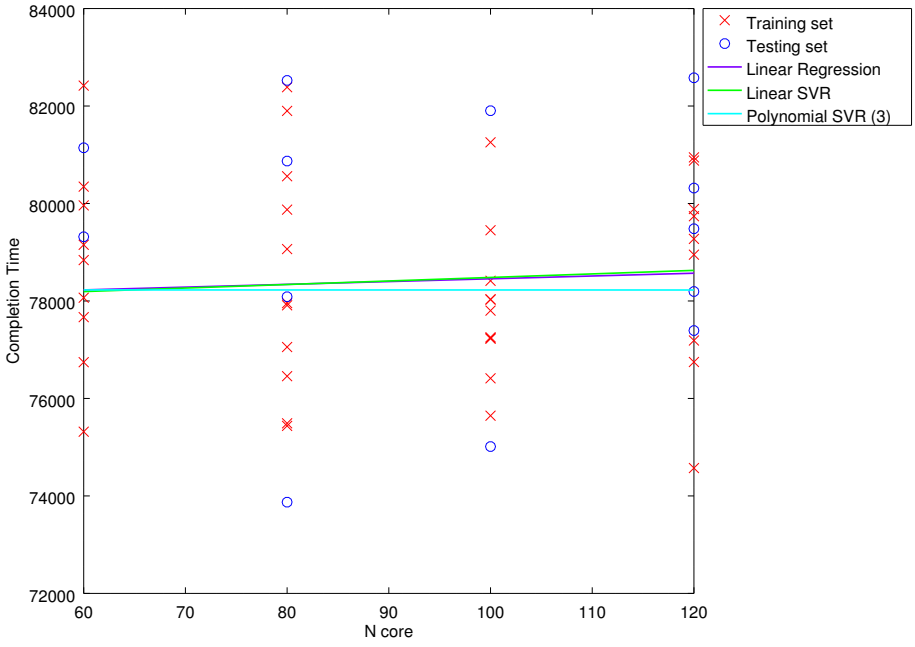

Figure 7: Completion time vs ncores for query R2 with datasize 750GB

1.2.4 Query R2 — Datasize 1000GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.0529	0.9971	21336	0.0227
Linear SVR	0.0872	0.9939	39397	0.0449
Polynomial SVR (2)	0.7767	0.4473	363910	0.4851
Polynomial SVR (3)	0.5370	0.8054	210031	0.2266
Polynomial SVR (4)	0.9436	0.5587	387439	0.4181
Polynomial SVR (6)	2.3977	0.5408	663955	0.6566
Gaussian SVR	0.2309	0.9596	94043	0.1177

Table 8: Results for R2-1000

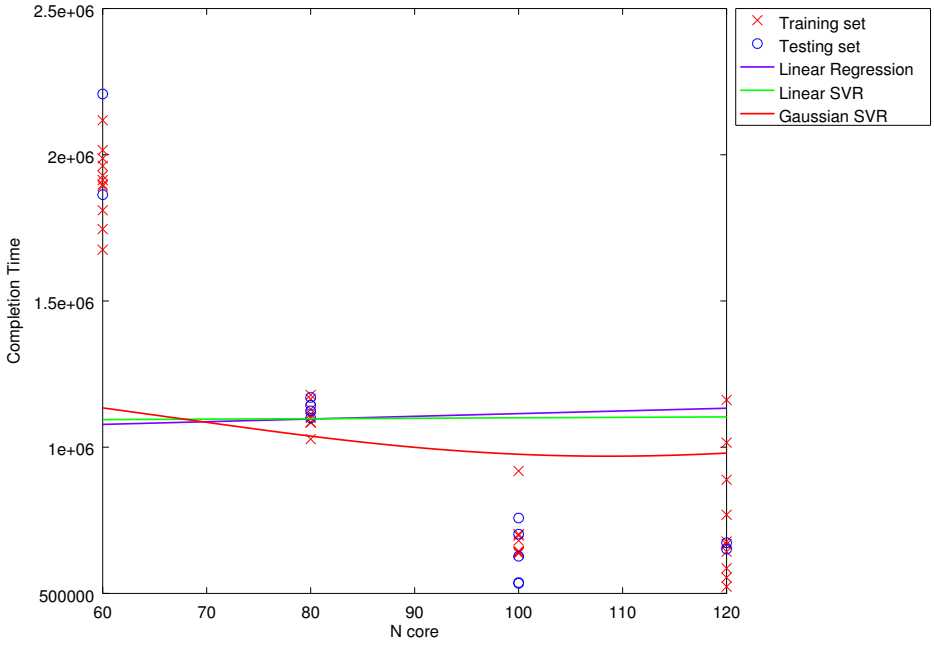

Figure 8: Completion time vs ncores for query R2 with datasize 1000GB

1.3 Query R3

1.3.1 Query R3 — Datasize 250GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.1109	0.9786	5118	0.0293
Linear SVR	0.1106	0.9889	5059	0.0292
Polynomial SVR (2)	0.7172	0.1678	38067	0.2733
Polynomial SVR (3)	0.2449	0.8985	13157	0.0911
Polynomial SVR (4)	0.6752	0.2430	35012	0.2465
Polynomial SVR (6)	0.6198	0.3670	35378	0.2375
Gaussian SVR	0.1868	0.9628	8716	0.0663

Table 9: Results for R3-250

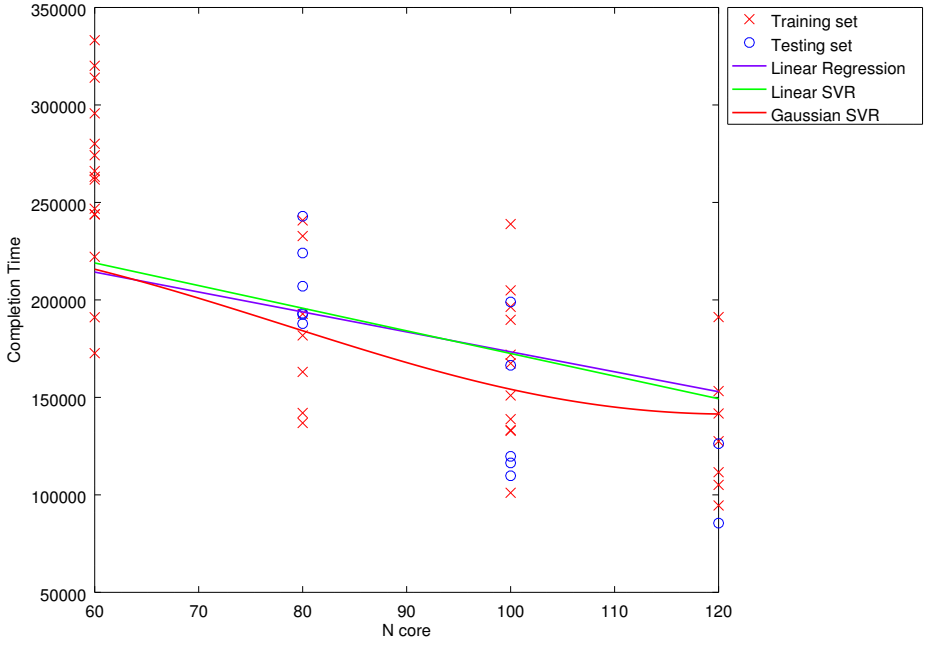

Figure 9: Completion time vs ncores for query R3 with datasize 250GB

1.3.2 Query R3 — Datasize 500GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.0535	0.9963	11098	0.0224
Linear SVR	0.0606	0.9955	13859	0.0280
Polynomial SVR (2)	0.2930	0.9013	59492	0.1343
Polynomial SVR (3)	0.1548	0.9691	36584	0.0771
Polynomial SVR (4)	0.1646	0.9658	37111	0.0703
Polynomial SVR (6)	0.3303	0.8652	60515	0.1288
Gaussian SVR	0.0698	0.9937	14553	0.0288

Table 10: Results for R3-500

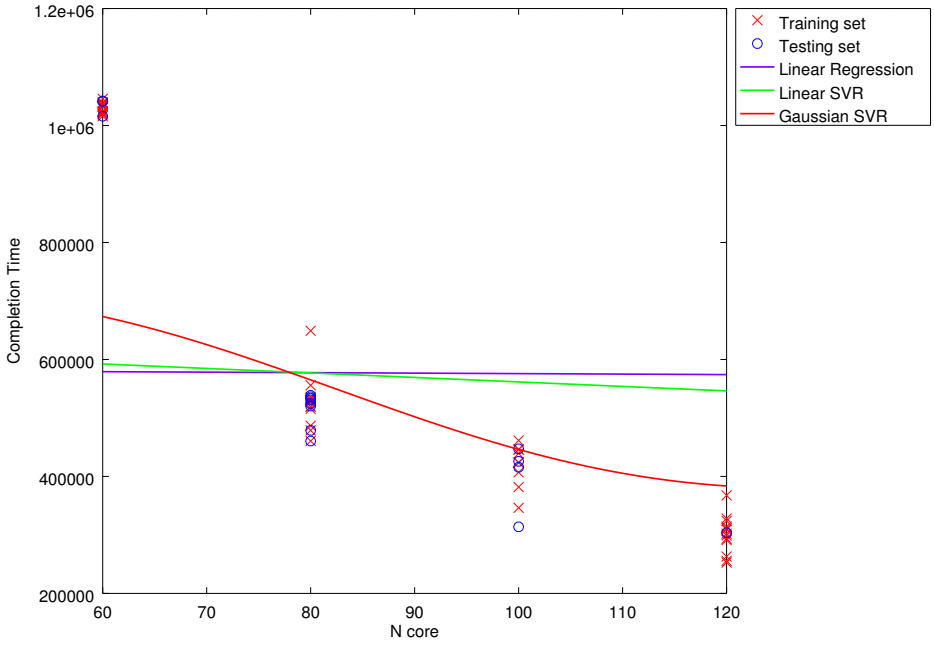

Figure 10: Completion time vs ncores for query R3 with datasize 500GB

1.3.3 Query R3 — Datasize 750GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.0220	0.9996	2722	0.0034
Linear SVR	0.0888	0.9948	13574	0.0177
Polynomial SVR (2)	1.0647	0.4290	145696	0.1758
Polynomial SVR (3)	0.5284	0.8347	73215	0.0972
Polynomial SVR (4)	0.8089	0.6248	113146	0.1515
Polynomial SVR (6)	0.8551	0.4785	114570	0.1545
Gaussian SVR	0.2530	0.9796	34529	0.0434

Table 11: Results for R3-750

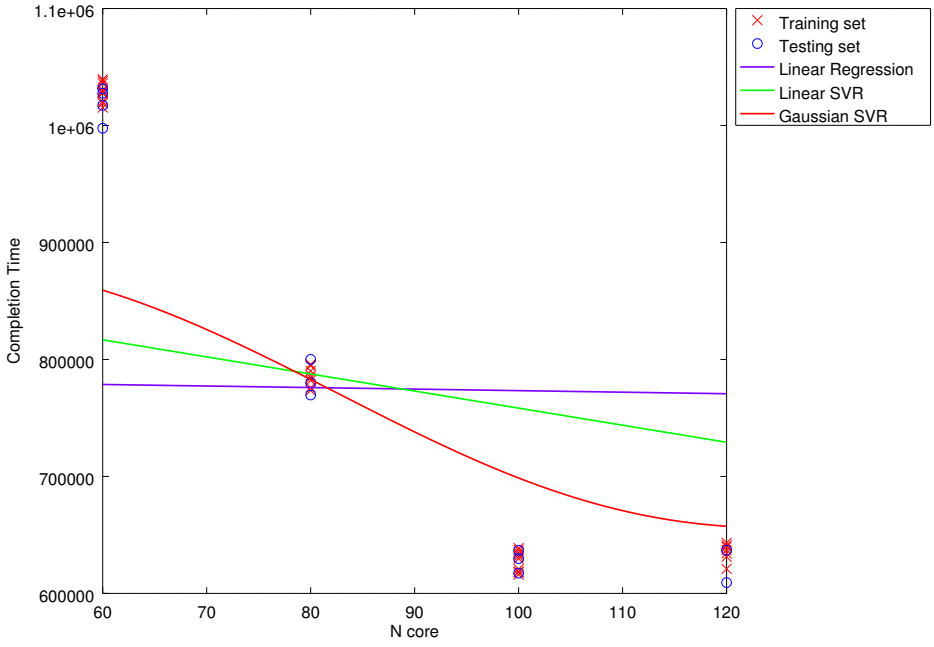

Figure 11: Completion time vs ncores for query R3 with datasize 750GB

1.3.4 Query R3 — Datasize 1000GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.0734	0.9955	9574	0.0085
Linear SVR	0.1067	0.9922	18970	0.0187
Polynomial SVR (2)	0.7430	0.6271	156805	0.1511
Polynomial SVR (3)	0.3859	0.9314	63411	0.0560
Polynomial SVR (4)	0.9817	0.2514	194521	0.1816
Polynomial SVR (6)	1.1743	0.0219	217951	0.2044
Gaussian SVR	0.3360	0.9471	54863	0.0472

Table 12: Results for R3-1000

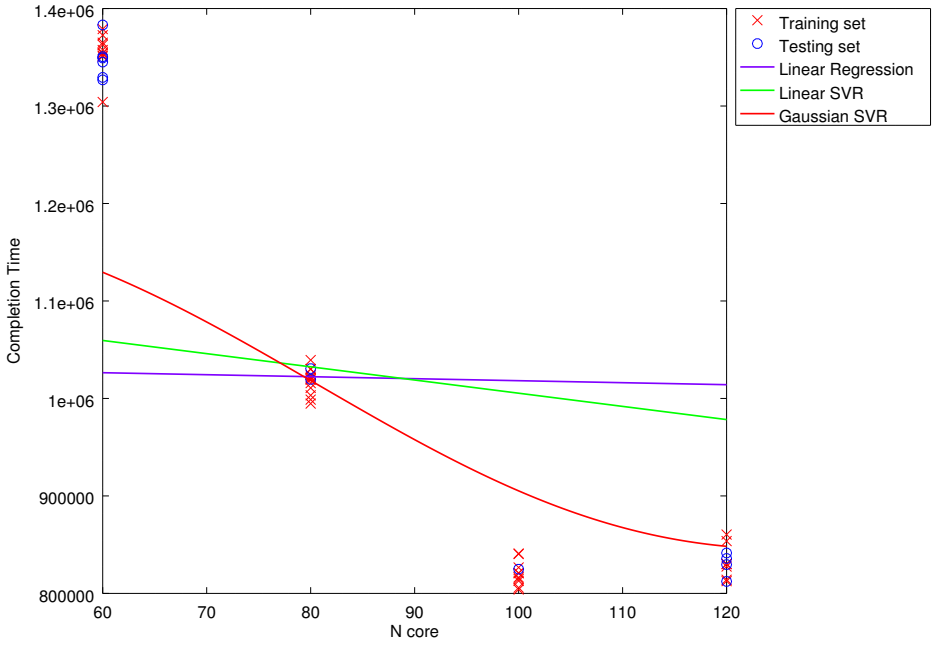

Figure 12: Completion time vs ncores for query R3 with datasize 1000GB

1.4 Query R4

1.4.1 Query R4 — Datasize 250GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.1719	0.9691	5155	0.0313
Linear SVR	0.1665	0.9721	5476	0.0331
Polynomial SVR (2)	0.8901	0.2806	37089	0.2708
Polynomial SVR (3)	0.6349	0.8391	23943	0.1357
Polynomial SVR (4)	0.8922	0.4584	37940	0.2624
Polynomial SVR (6)	0.8339	0.5929	35512	0.2419
Gaussian SVR	0.2832	0.9483	9304	0.0558

Table 13: Results for R4-250

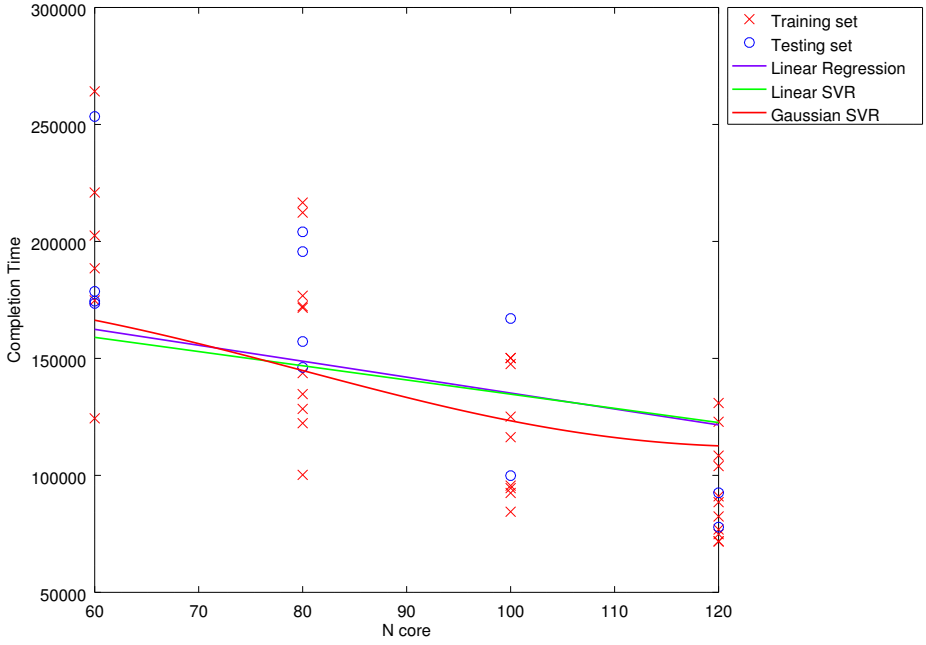

Figure 13: Completion time vs ncores for query R4 with datasize 250GB

1.4.2 Query R4 — Datasize 500GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.0523	0.9941	5530	0.0157
Linear SVR	0.0815	0.9947	13073	0.0334
Polynomial SVR (2)	0.3104	0.8825	40014	0.1194
Polynomial SVR (3)	0.1446	0.9635	19321	0.0472
Polynomial SVR (4)	0.2598	0.9353	35711	0.0904
Polynomial SVR (6)	0.2765	0.8948	36156	0.0950
Gaussian SVR	0.0816	0.9881	10981	0.0270

Table 14: Results for R4-500

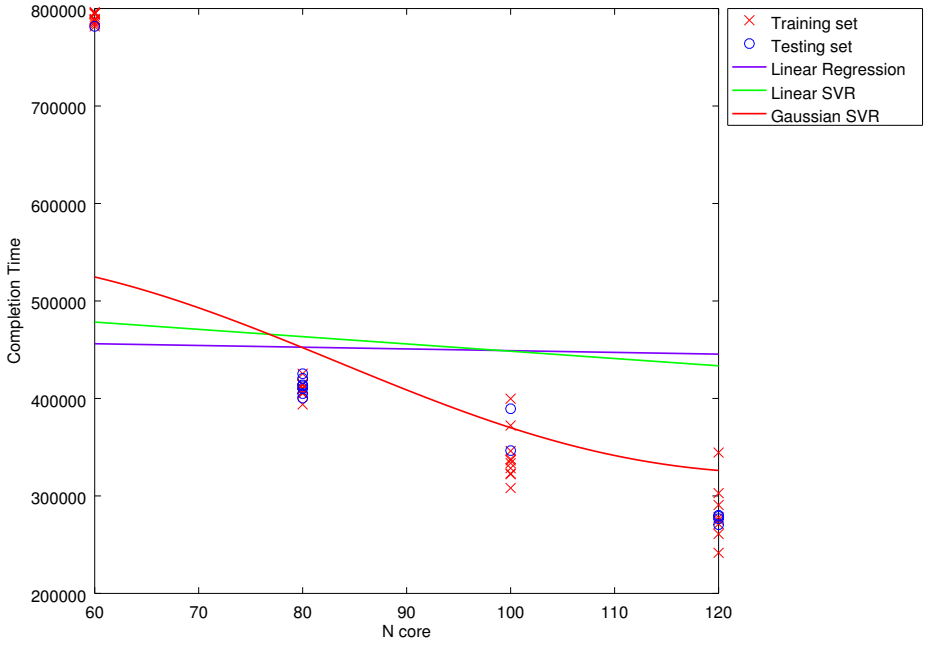

Figure 14: Completion time vs ncores for query R4 with datasize 500GB

1.4.3 Query R4 — Datasize 750GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.0215	0.9996	2345	0.0038
Linear SVR	0.0808	0.9957	9071	0.0162
Polynomial SVR (2)	0.6955	0.7218	85402	0.1413
Polynomial SVR (3)	0.3624	0.9413	39380	0.0615
Polynomial SVR (4)	0.8872	0.5977	94544	0.1441
Polynomial SVR (6)	1.0199	0.3856	106583	0.1602
Gaussian SVR	0.1759	0.9783	14459	0.0248

Table 15: Results for R4-750

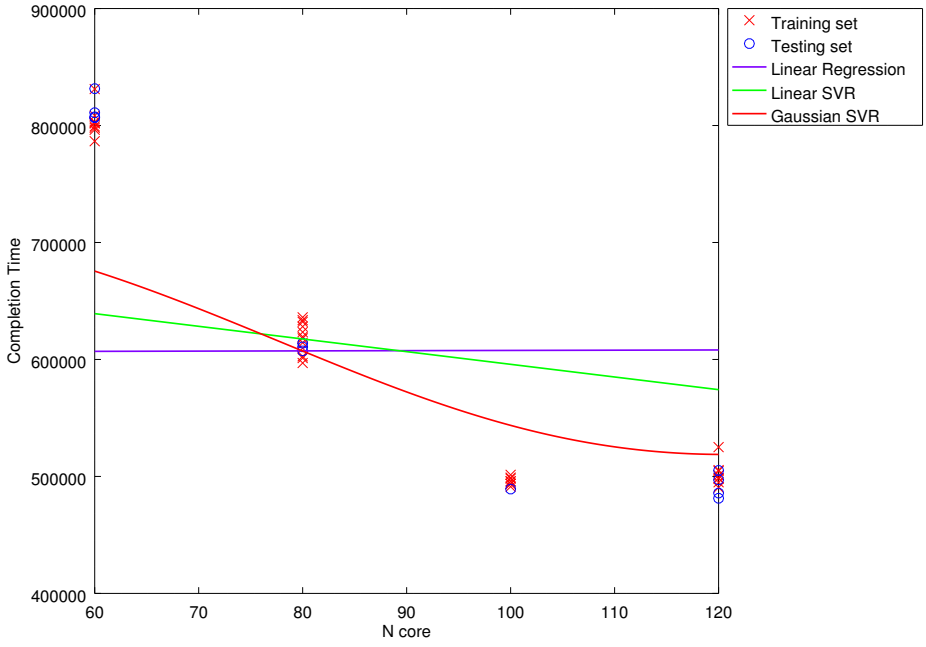

Figure 15: Completion time vs ncores for query R4 with datasize 750GB

1.4.4 Query R4 — Datasize 1000GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.0944	0.9924	56636	0.0333
Linear SVR	0.1024	0.9936	60674	0.0388
Polynomial SVR (2)	0.8395	0.4422	580227	0.3698
Polynomial SVR (3)	0.4507	0.9060	296096	0.2340
Polynomial SVR (4)	0.8560	0.6428	607746	0.4239
Polynomial SVR (6)	0.9475	0.5411	673570	0.4938
Gaussian SVR	0.2280	0.9589	129214	0.0695

Table 16: Results for R4-1000

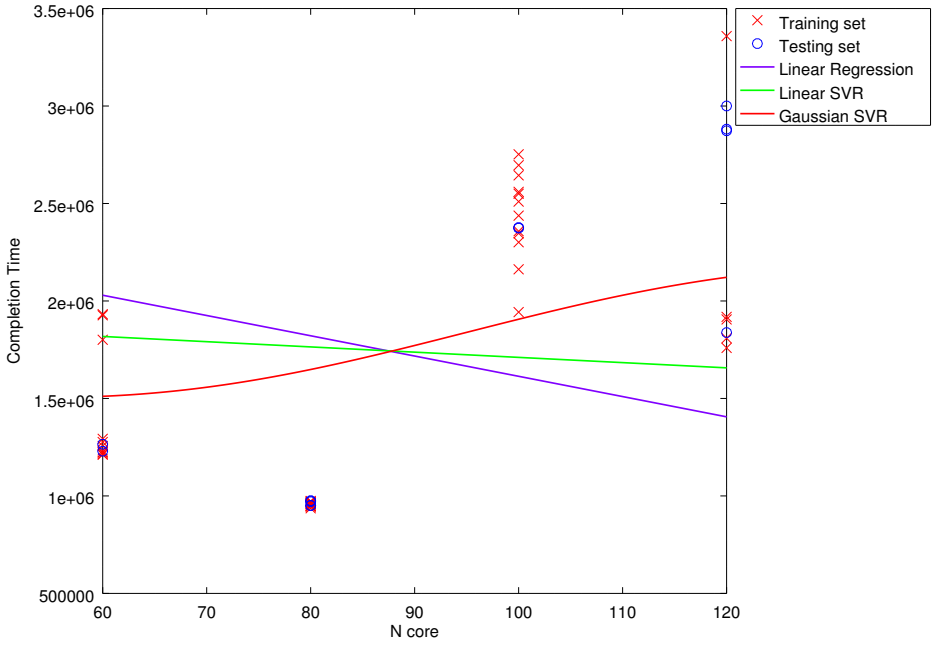

Figure 16: Completion time vs ncores for query R4 with datasize 1000GB

1.5 Query R5

1.5.1 Query R5 — Datasize 250GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	1.1021	-0.0551	713	0.0275
Linear SVR	1.0368	0.3190	639	0.0246
Polynomial SVR (2)	1.1427	0.0277	769	0.0299
Polynomial SVR (3)	1.0482	0.1294	688	0.0266
Polynomial SVR (4)	1.1212	0.0921	728	0.0282
Polynomial SVR (6)	1.1117	0.0337	696	0.0270
Gaussian SVR	0.9581	0.4034	595	0.0229

Table 17: Results for R5-250

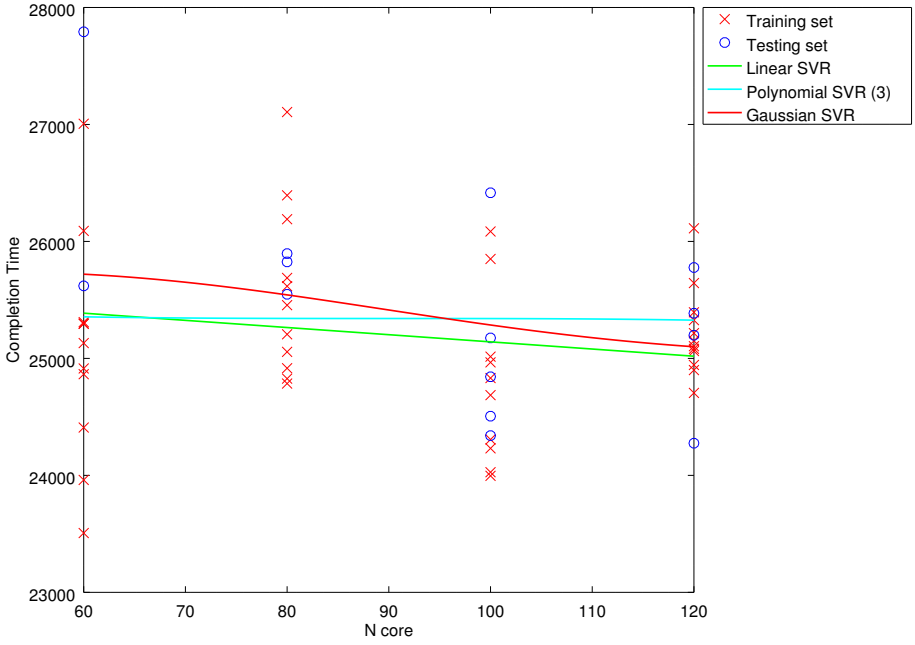

Figure 17: Completion time vs ncores for query R5 with datasize 250GB

1.5.2 Query R5 — Datasize 500GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.5027	0.5247	398	0.0169
Linear SVR	0.4869	0.6065	386	0.0163
Polynomial SVR (2)	0.7132	0.2590	753	0.0318
Polynomial SVR (3)	0.6277	0.4057	524	0.0221
Polynomial SVR (4)	0.8572	0.0574	770	0.0325
Polynomial SVR (6)	0.7629	0.0794	741	0.0314
Gaussian SVR	0.5268	0.5836	433	0.0182

Table 18: Results for R5-500

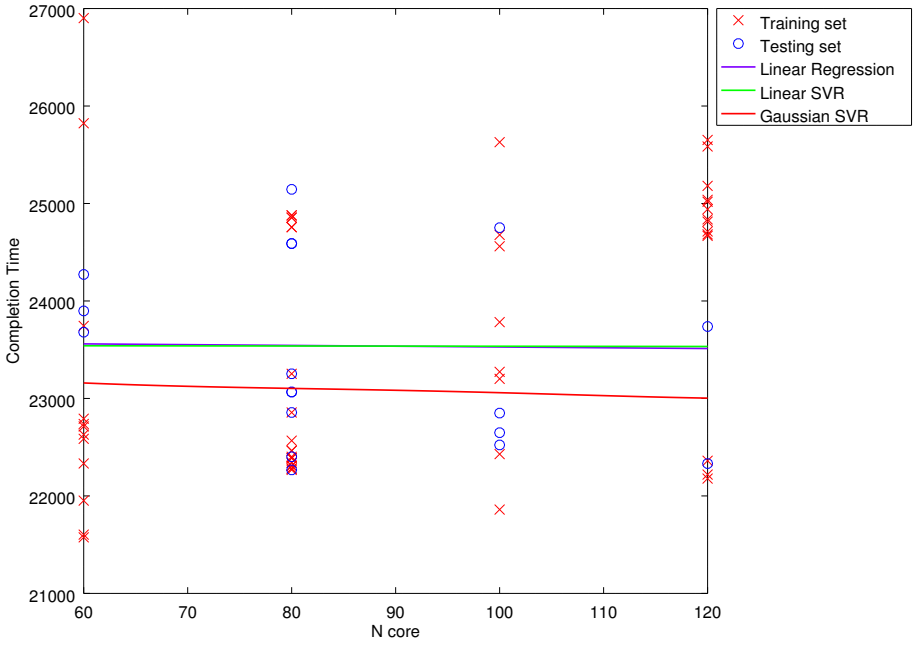

Figure 18: Completion time vs ncores for query R5 with datasize 500GB

1.5.3 Query R5 — Datasize 750GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.5755	0.6749	440	0.0176
Linear SVR	0.6121	0.7036	465	0.0186
Polynomial SVR (2)	1.4732	0.1597	1181	0.0486
Polynomial SVR (3)	0.8230	0.5766	841	0.0343
Polynomial SVR (4)	1.4430	0.0262	1312	0.0534
Polynomial SVR (6)	1.2012	0.0051	1189	0.0484
Gaussian SVR	0.6737	0.5912	521	0.0212

Table 19: Results for R5-750

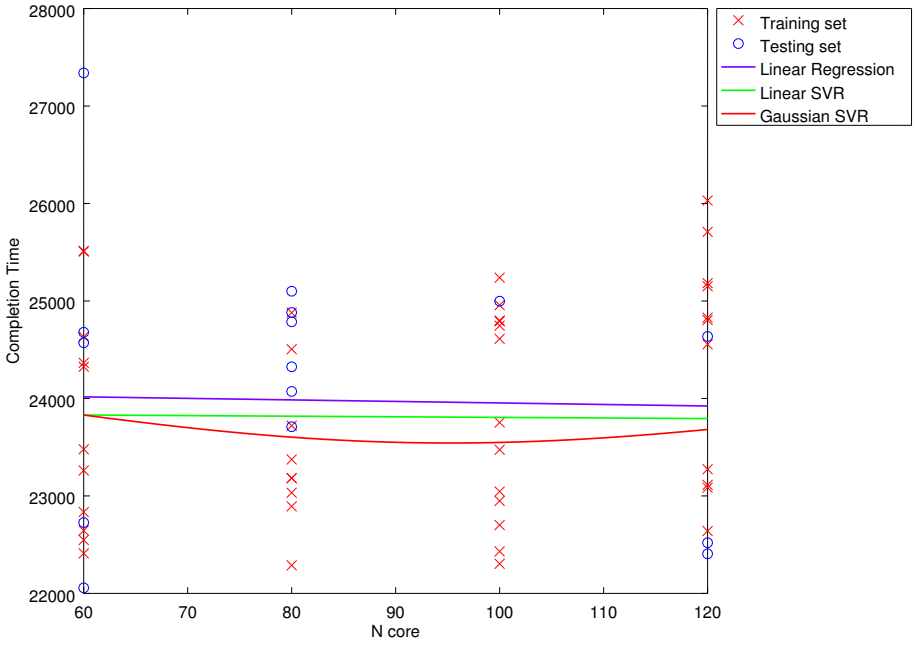

Figure 19: Completion time vs ncores for query R5 with datasize 750GB

1.5.4 Query R5 — Datasize 1000GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.5015	0.1498	1322	0.0338
Linear SVR	0.4946	0.3382	1208	0.0309
Polynomial SVR (2)	0.4199	0.5664	947	0.0237
Polynomial SVR (3)	0.5480	0.2594	1525	0.0388
Polynomial SVR (4)	0.4707	0.3918	1138	0.0287
Polynomial SVR (6)	0.5688	0.3014	1456	0.0366
Gaussian SVR	0.3646	0.5791	820	0.0207

Table 20: Results for R5-1000

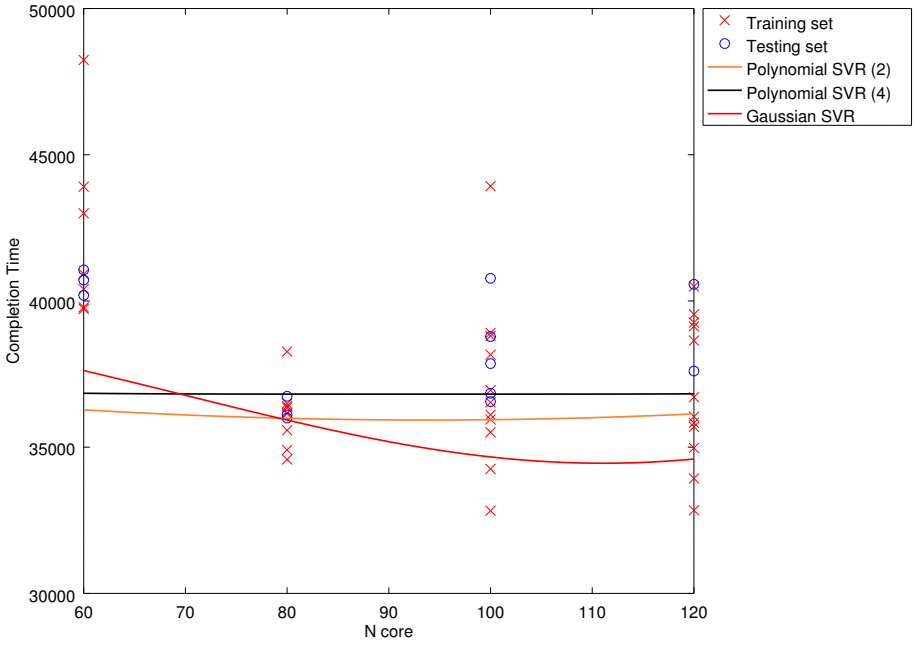

Figure 20: Completion time vs ncores for query R5 with datasize 1000GB

2 Fixed Datasize, linear *ncores*, only *ncores*

2.1 Query R1

2.1.1 Query R1 — Datasize 250GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	1.2075	0.3647	11200	0.1125
Linear SVR	1.2849	0.4436	10934	0.1003
Polynomial SVR (2)	1.5934	0.0830	22694	0.3583
Polynomial SVR (3)	1.3520	0.3512	12494	0.1275
Polynomial SVR (4)	1.5463	0.0938	22056	0.3381
Polynomial SVR (6)	1.5344	0.1080	22003	0.3376
Gaussian SVR	1.2809	0.4200	11773	0.1226

Table 21: Results for R1-250GB, only ncores

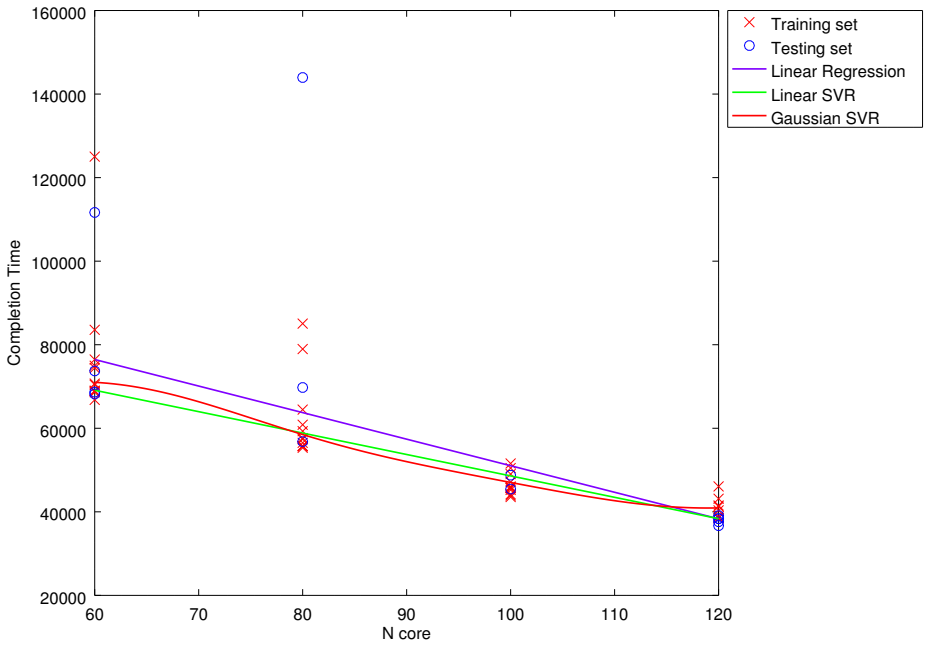

Figure 21: Completion time vs ncores for query R1 with datasize 250GB

2.1.2 Query R1 — Datasize 500GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.6040	0.6338	44332	0.3321
Linear SVR	0.6692	0.6428	48643	0.3623
Polynomial SVR (2)	0.9736	0.0889	62658	0.5003
Polynomial SVR (3)	0.5829	0.6924	40582	0.3111
Polynomial SVR (4)	0.9816	0.1055	66974	0.6097
Polynomial SVR (6)	0.9863	0.1273	67430	0.6322
Gaussian SVR	0.4035	0.9019	30189	0.2597

Table 22: Results for R1-500GB, only ncores

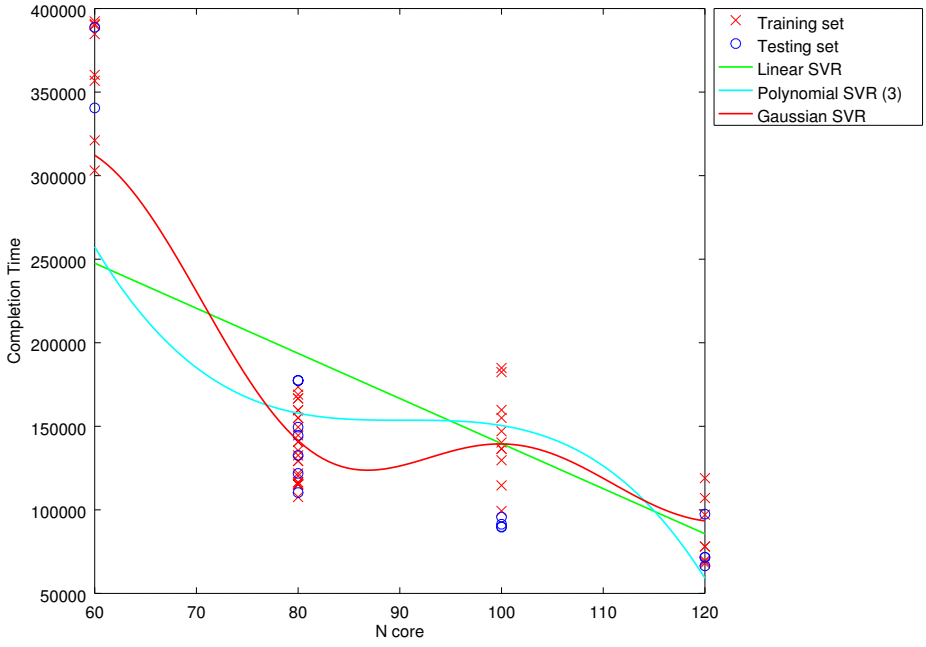

Figure 22: Completion time vs ncores for query R1 with datasize 500GB

2.1.3 Query R1 — Datasize 750GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.6470	0.7384	50197	0.2333
Linear SVR	0.6033	0.8114	44222	0.2349
Polynomial SVR (2)	1.2098	0.1305	97577	0.4832
Polynomial SVR (3)	0.6262	0.7862	46007	0.2396
Polynomial SVR (4)	1.2068	0.1404	97323	0.4831
Polynomial SVR (6)	1.2002	0.1515	96594	0.4889
Gaussian SVR	0.5955	0.8882	43084	0.2591

Table 23: Results for R1-750GB, only ncores

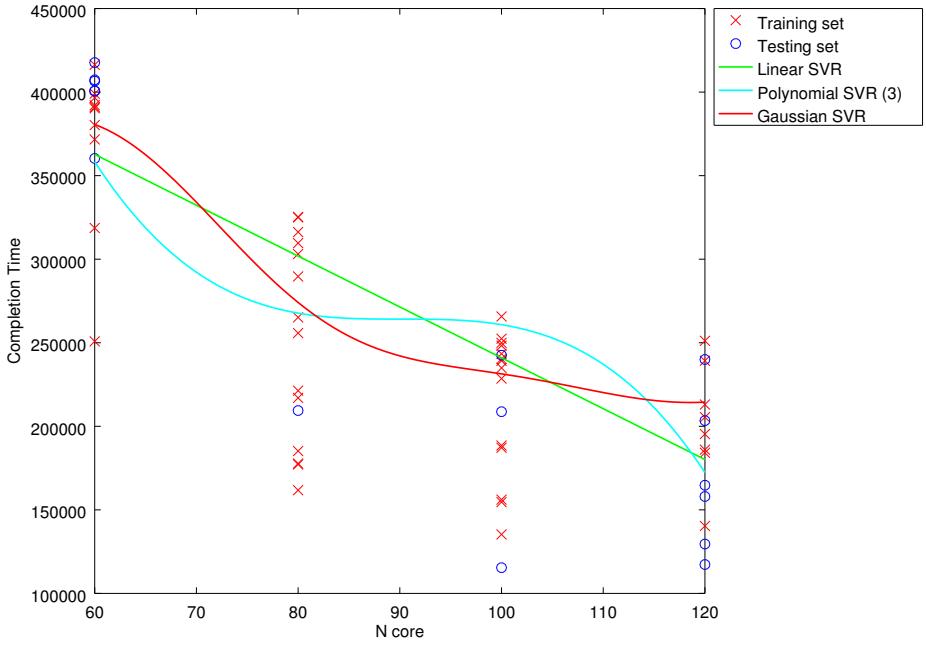

Figure 23: Completion time vs ncores for query R1 with datasize 750GB

2.1.4 Query R1 — Datasize 1000GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.4444	0.8004	38407	0.0906
Linear SVR	0.4607	0.8045	38550	0.0889
Polynomial SVR (2)	0.9613	0.1422	78118	0.1746
Polynomial SVR (3)	0.4834	0.7835	36629	0.0874
Polynomial SVR (4)	0.9672	0.1294	78512	0.1755
Polynomial SVR (6)	0.9746	0.1150	79059	0.1767
Gaussian SVR	0.2943	0.9192	18751	0.0436

Table 24: Results for R1-1000GB, only ncores

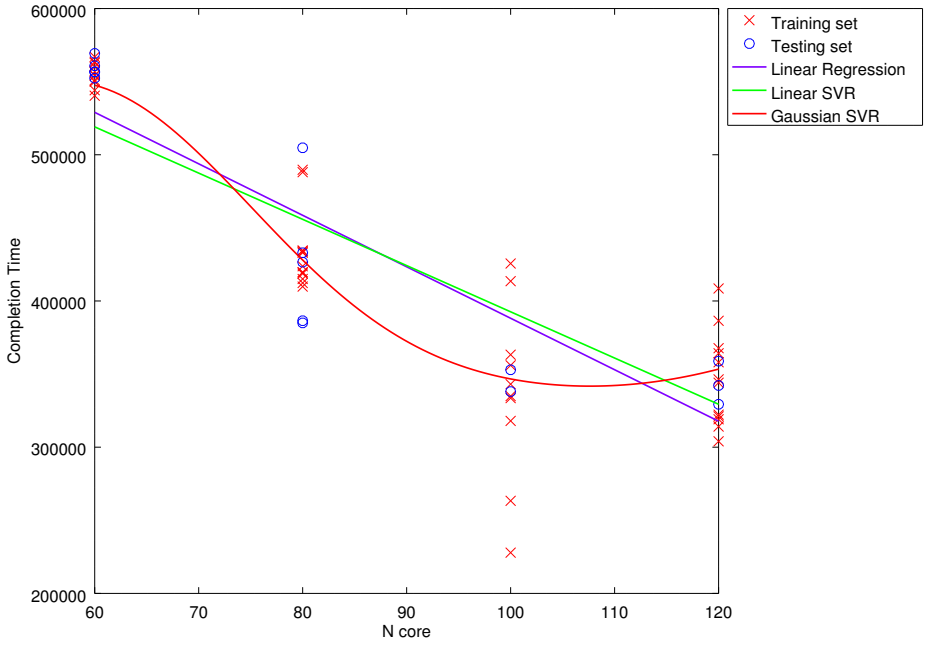

Figure 24: Completion time vs ncores for query R1 with datasize 1000GB

2.2 Query R2

2.2.1 Query R2 — Datasize 250GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.9018	-0.0459	2776	0.0338
Linear SVR	0.9105	0.1154	2801	0.0342
Polynomial SVR (2)	0.9798	0.0107	3015	0.0368
Polynomial SVR (3)	0.9001	0.1108	2702	0.0329
Polynomial SVR (4)	0.9790	0.0144	3009	0.0367
Polynomial SVR (6)	0.9783	0.0199	3003	0.0366
Gaussian SVR	0.9095	0.1187	2800	0.0342

Table 25: Results for R2-250GB, only ncores

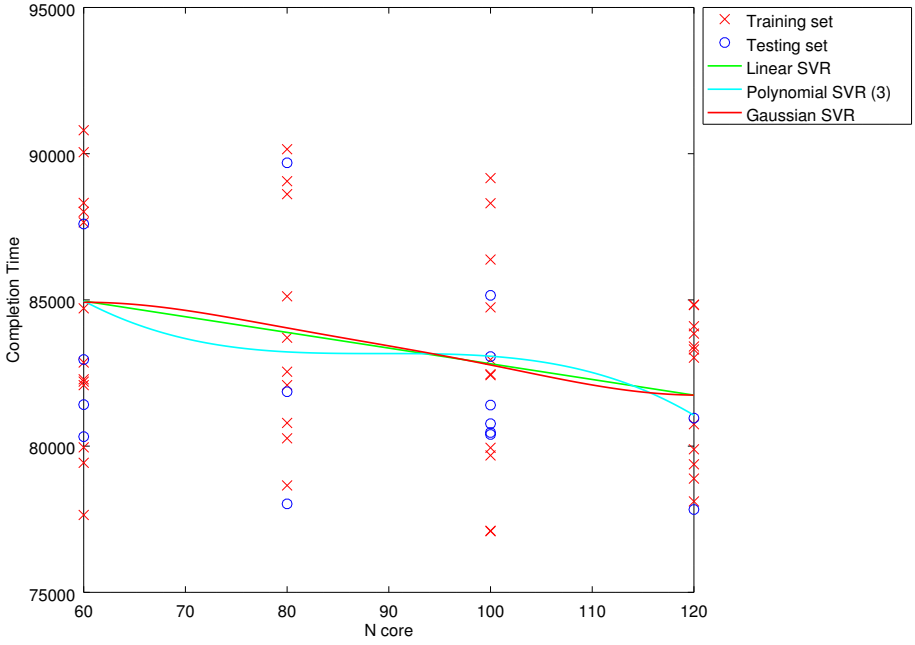

Figure 25: Completion time vs ncores for query R2 with datasize 250GB

2.2.2 Query R2 — Datasize 500GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	1.2116	-0.0047	2935	0.0404
Linear SVR	1.2260	0.0144	2980	0.0410
Polynomial SVR (2)	1.2156	0.1347	2995	0.0413
Polynomial SVR (3)	1.2654	0.0000	3199	0.0443
Polynomial SVR (4)	1.2183	0.1115	2985	0.0411
Polynomial SVR (6)	1.2194	0.0983	2987	0.0412
Gaussian SVR	1.2449	0.0064	3114	0.0431

Table 26: Results for R2-500GB, only ncores

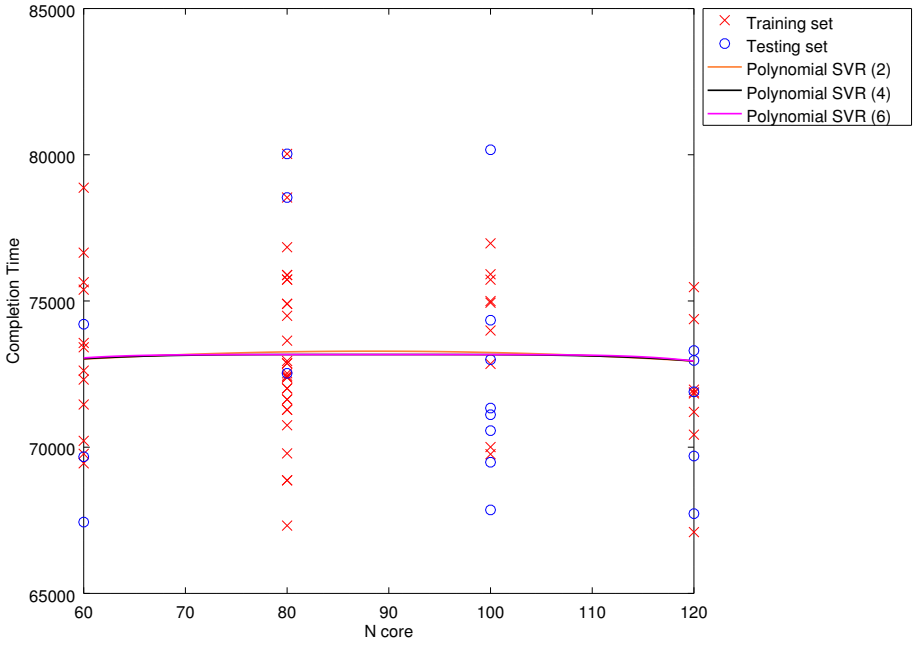

Figure 26: Completion time vs ncores for query R2 with datasize 500GB

2.2.3 Query R2 — Datasize 750GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.9999	-0.0420	2058	0.0263
Linear SVR	1.0055	0.0415	1982	0.0252
Polynomial SVR (2)	1.0883	0.0125	2252	0.0286
Polynomial SVR (3)	1.0177	0.0581	1973	0.0251
Polynomial SVR (4)	1.0924	0.0151	2266	0.0288
Polynomial SVR (6)	1.1161	0.0179	2327	0.0295
Gaussian SVR	1.0195	0.0150	2086	0.0266

Table 27: Results for R2-750GB, only ncores

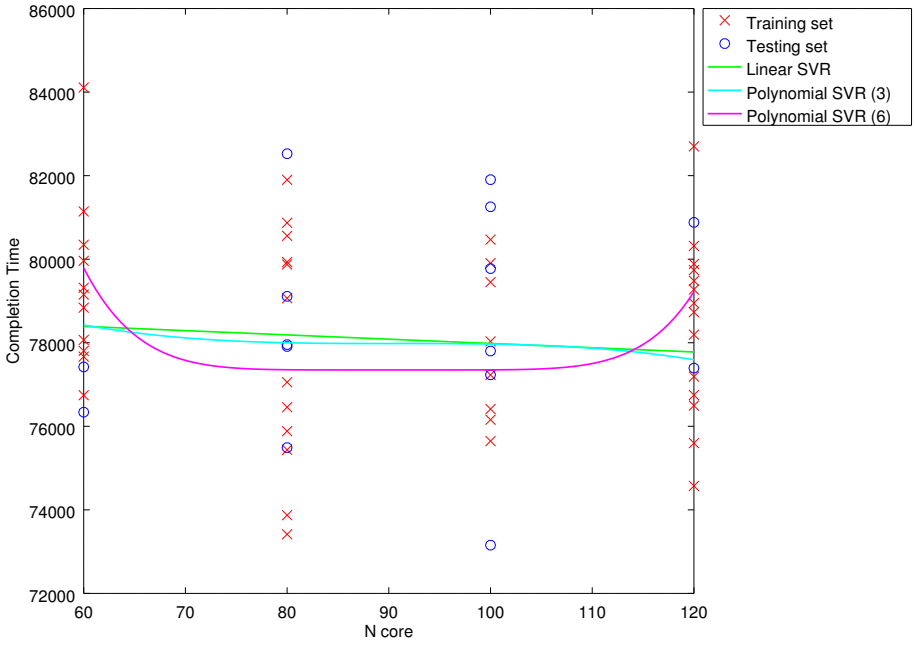

Figure 27: Completion time vs ncores for query R2 with datasize 750GB

2.2.4 Query R2 — Datasize 1000GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.5107	0.5100	255920	0.3049
Linear SVR	0.6277	0.6461	290435	0.3504
Polynomial SVR (2)	0.7885	0.0508	318924	0.4104
Polynomial SVR (3)	0.5858	0.4917	221149	0.2926
Polynomial SVR (4)	0.7500	0.0548	325104	0.3925
Polynomial SVR (6)	0.7566	0.0616	329967	0.3710
Gaussian SVR	0.2020	0.9521	87300	0.0985

Table 28: Results for R2-1000GB, only ncores

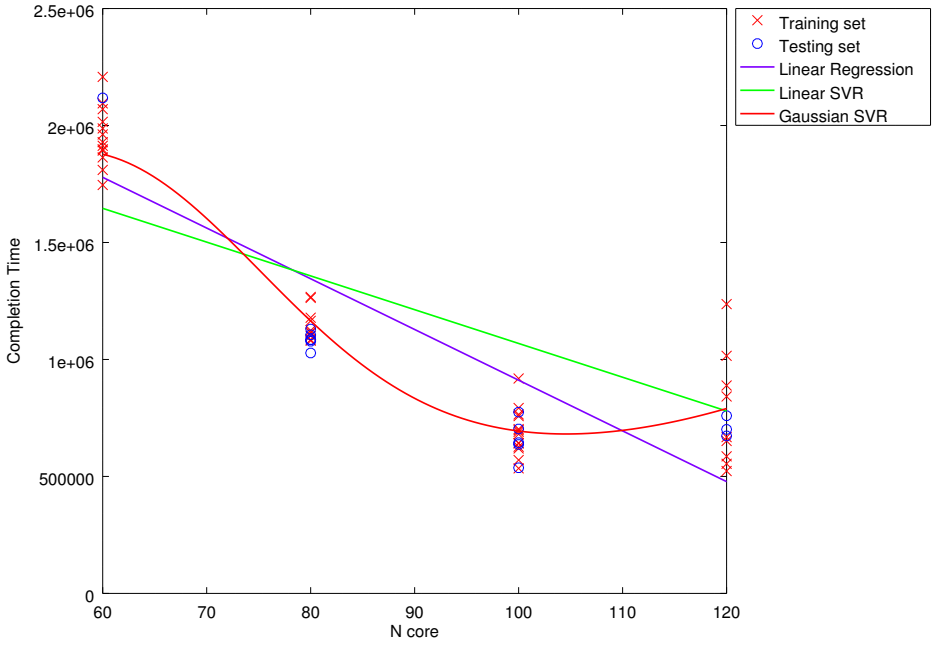

Figure 28: Completion time vs ncores for query R2 with datasize 1000GB

2.3 Query R3

2.3.1 Query R3 — Datasize 250GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.7631	0.4438	39680	0.1810
Linear SVR	0.7755	0.6918	39756	0.1747
Polynomial SVR (2)	1.2251	0.0307	67244	0.3277
Polynomial SVR (3)	0.7870	0.6821	40268	0.1779
Polynomial SVR (4)	1.1975	0.0394	65156	0.3207
Polynomial SVR (6)	1.1661	0.0508	63994	0.3249
Gaussian SVR	0.7771	0.6830	41435	0.1982

Table 29: Results for R3-250GB, only ncores

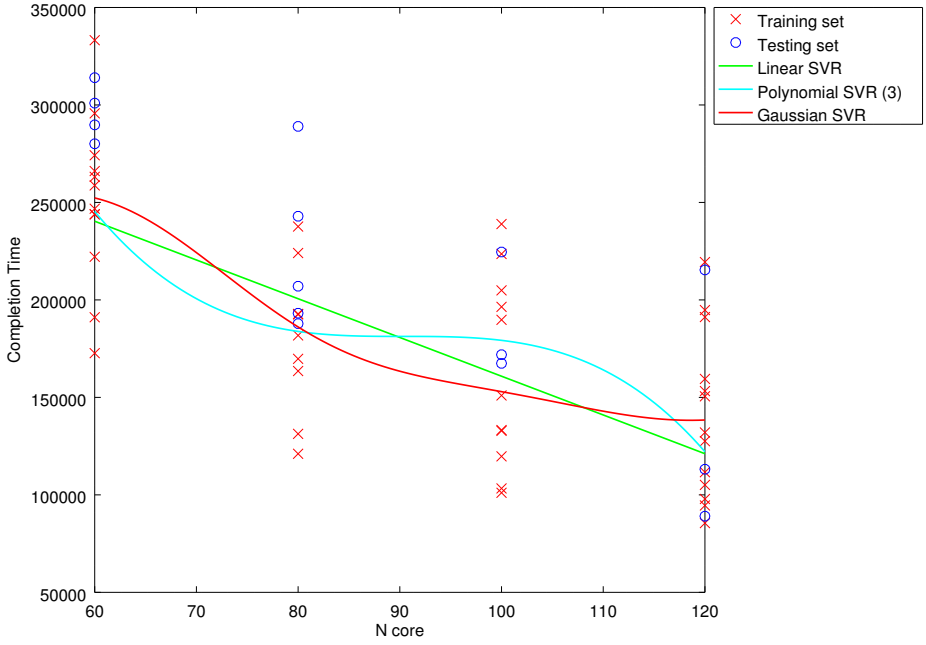

Figure 29: Completion time vs ncores for query R3 with datasize 250GB

2.3.2 Query R3 — Datasize 500GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.4959	0.7518	114279	0.2329
Linear SVR	0.5009	0.7582	114192	0.2258
Polynomial SVR (2)	1.0225	0.0258	185043	0.3254
Polynomial SVR (3)	0.5107	0.7434	115022	0.2303
Polynomial SVR (4)	1.0233	0.0032	189124	0.3289
Polynomial SVR (6)	1.0081	0.0024	186294	0.3216
Gaussian SVR	0.1438	0.9806	27885	0.0616

Table 30: Results for R3-500GB, only ncores

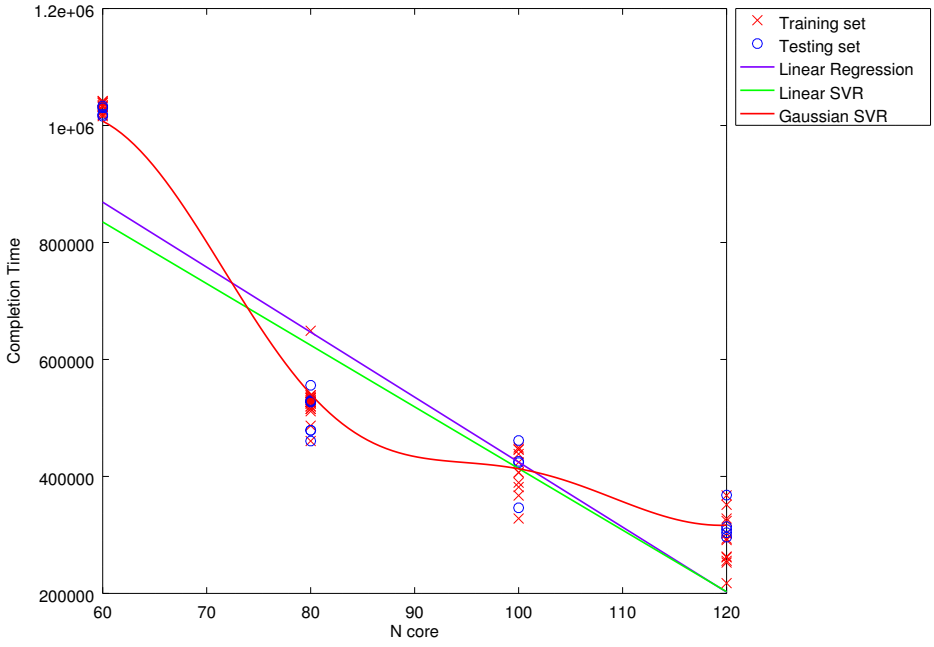

Figure 30: Completion time vs ncores for query R3 with datasize 500GB

2.3.3 Query R3 — Datasize 750GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.3793	0.8368	59263	0.0822
Linear SVR	0.3795	0.8464	58792	0.0810
Polynomial SVR (2)	1.0110	0.0321	138792	0.1985
Polynomial SVR (3)	0.5158	0.7551	59479	0.0780
Polynomial SVR (4)	1.0187	0.0211	139861	0.2001
Polynomial SVR (6)	1.0401	0.0000	151415	0.2135
Gaussian SVR	0.1659	0.9863	19628	0.0297

Table 31: Results for R3-750GB, only ncores

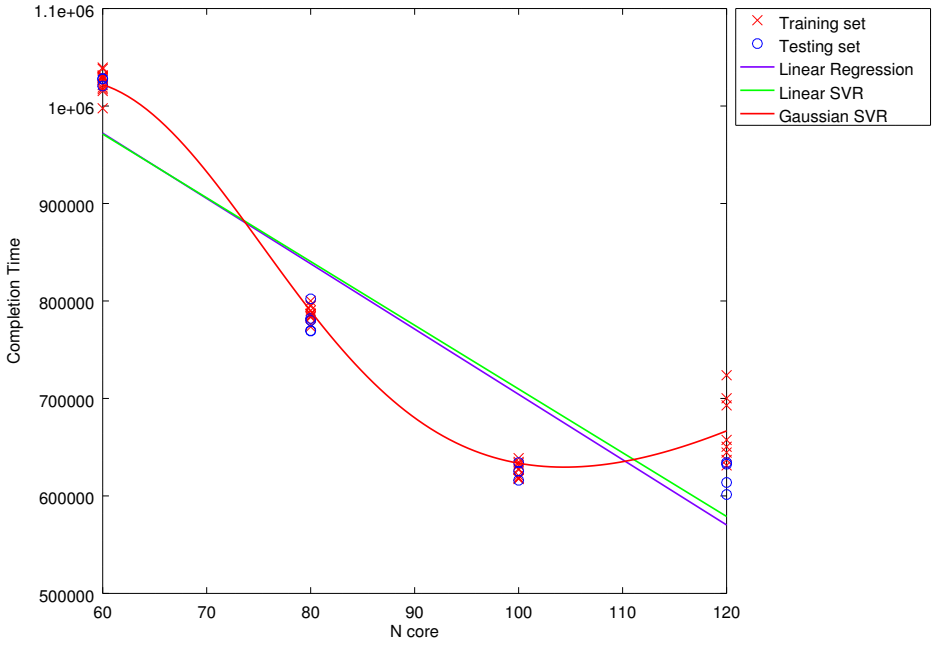

Figure 31: Completion time vs ncores for query R3 with datasize 750GB

2.3.4 Query R3 — Datasize 1000GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.4163	0.8159	86213	0.0920
Linear SVR	0.4160	0.8199	86180	0.0920
Polynomial SVR (2)	0.9122	0.1475	172594	0.1799
Polynomial SVR (3)	0.5750	0.6735	103081	0.1159
Polynomial SVR (4)	0.9131	0.1417	172440	0.1782
Polynomial SVR (6)	0.9813	0.1347	185181	0.2003
Gaussian SVR	0.0971	0.9938	16264	0.0168

Table 32: Results for R3-1000GB, only ncores

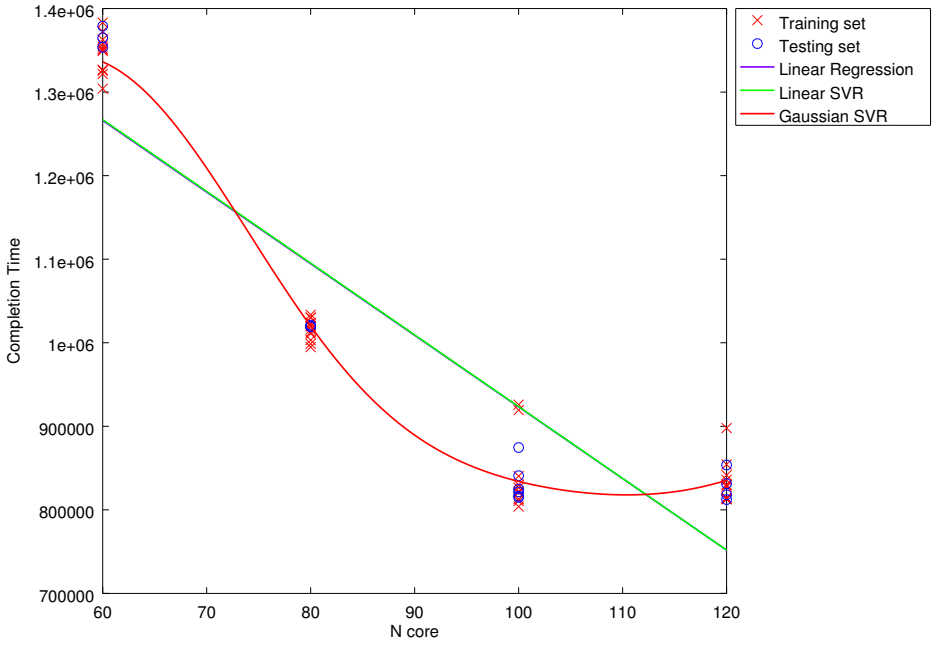

Figure 32: Completion time vs ncores for query R3 with datasize 1000GB

2.4 Query R4

2.4.1 Query R4 — Datasize 250GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.7819	0.3094	32190	0.2305
Linear SVR	0.7724	0.3411	31610	0.2369
Polynomial SVR (2)	1.1072	0.2345	43972	0.3869
Polynomial SVR (3)	0.8101	0.2724	29682	0.2136
Polynomial SVR (4)	1.0995	0.1685	42974	0.3796
Polynomial SVR (6)	1.0622	0.1020	42673	0.3596
Gaussian SVR	0.8106	0.2932	34104	0.2553

Table 33: Results for R4-250GB, only ncores

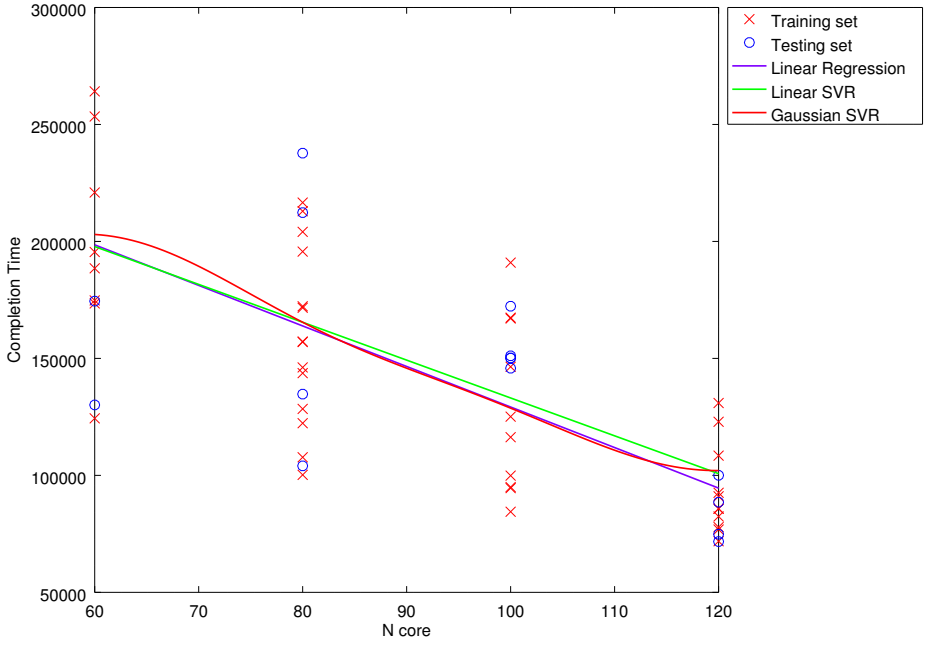

Figure 33: Completion time vs ncores for query R4 with datasize 250GB

2.4.2 Query R4 — Datasize 500GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.5643	0.7762	91169	0.1914
Linear SVR	0.6193	0.8227	93111	0.1792
Polynomial SVR (2)	1.3673	0.0397	168421	0.2839
Polynomial SVR (3)	0.5852	0.8088	97055	0.2142
Polynomial SVR (4)	1.3574	0.0034	166877	0.2794
Polynomial SVR (6)	1.3285	0.0062	162718	0.2726
Gaussian SVR	0.0984	0.9942	15104	0.0406

Table 34: Results for R4-500GB, only ncores

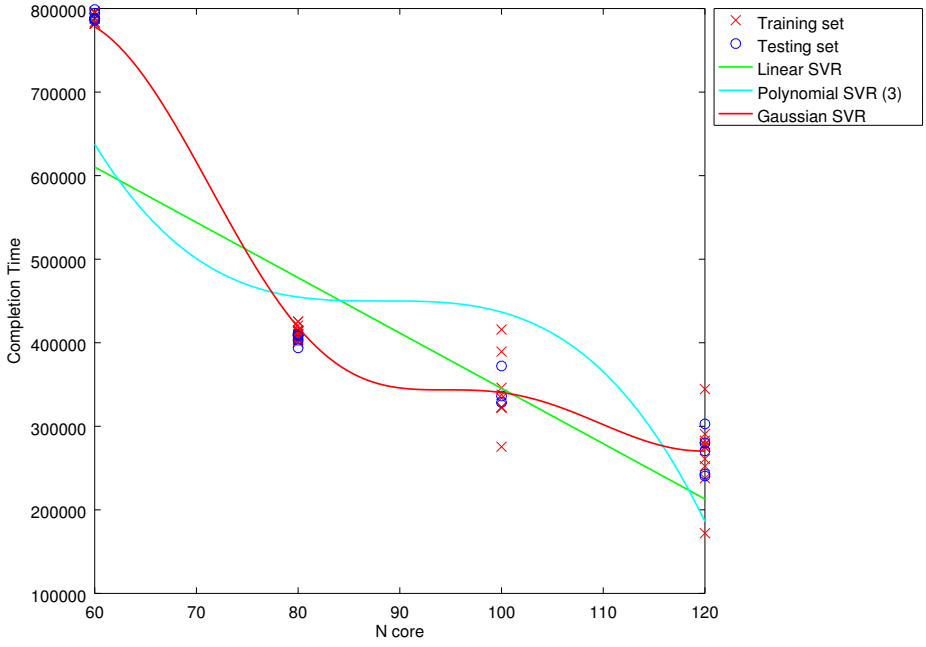

Figure 34: Completion time vs ncores for query R4 with datasize 500GB

2.4.3 Query R4 — Datasize 750GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.4434	0.6809	53831	0.0977
Linear SVR	0.3957	0.7879	48222	0.0866
Polynomial SVR (2)	0.7974	0.1681	84346	0.1474
Polynomial SVR (3)	0.5251	0.5563	59852	0.1091
Polynomial SVR (4)	0.7986	0.1533	84417	0.1475
Polynomial SVR (6)	0.8038	0.1311	84955	0.1485
Gaussian SVR	0.1123	0.9797	10796	0.0194

Table 35: Results for R4-750GB, only ncores

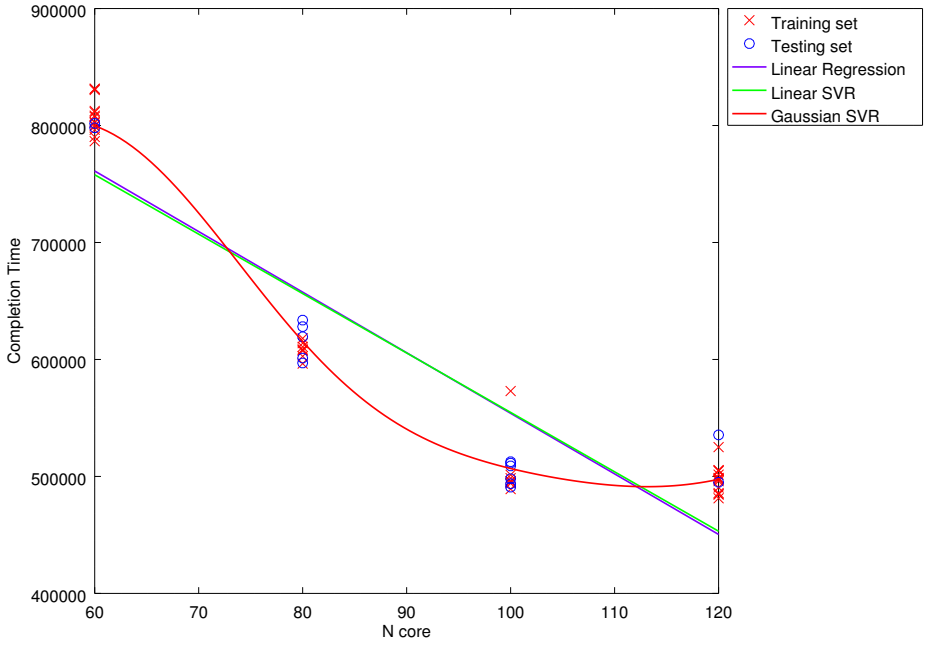

Figure 35: Completion time vs ncores for query R4 with datasize 750GB

2.4.4 Query R4 — Datasize 1000GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.5311	0.5655	342449	0.2699
Linear SVR	0.5613	0.5742	416843	0.3029
Polynomial SVR (2)	0.9229	0.0086	713254	0.5136
Polynomial SVR (3)	0.6772	0.3007	445678	0.3277
Polynomial SVR (4)	0.9151	0.0080	707727	0.5087
Polynomial SVR (6)	0.9127	0.0059	705796	0.5076
Gaussian SVR	0.2633	0.8936	147579	0.0877

Table 36: Results for R4-1000GB, only ncores

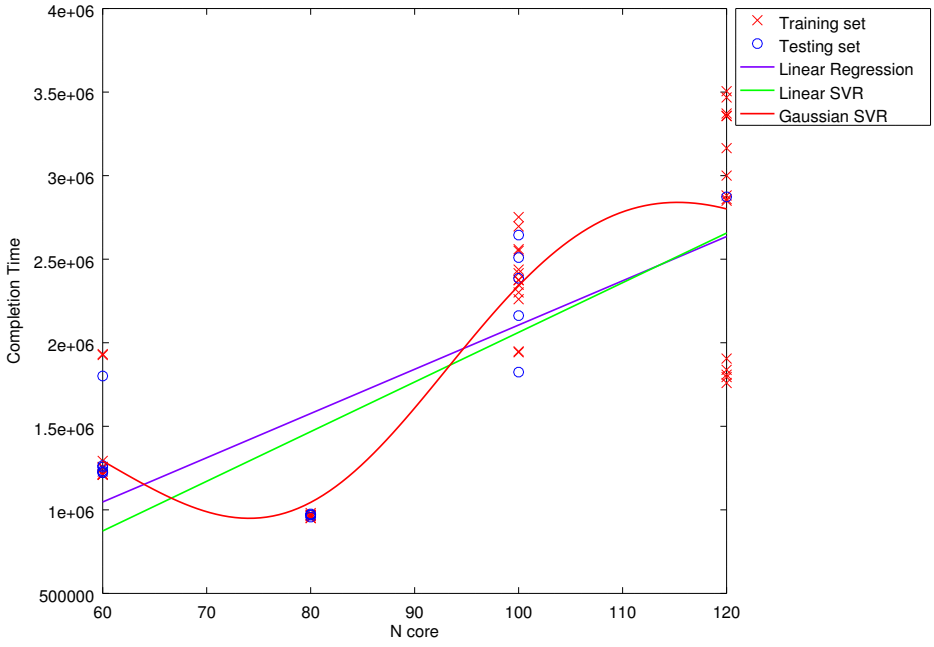

Figure 36: Completion time vs ncores for query R4 with datasize 1000GB

2.5 Query R5

2.5.1 Query R5 — Datasize 250GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.8259	-0.3016	536	0.0214
Linear SVR	0.7658	0.0010	512	0.0203
Polynomial SVR (2)	0.7559	0.0294	503	0.0200
Polynomial SVR (3)	0.7594	0.0060	504	0.0200
Polynomial SVR (4)	0.7642	0.0270	507	0.0201
Polynomial SVR (6)	0.7630	0.0253	508	0.0201
Gaussian SVR	0.8189	0.0013	486	0.0192

Table 37: Results for R5-250GB, only ncores

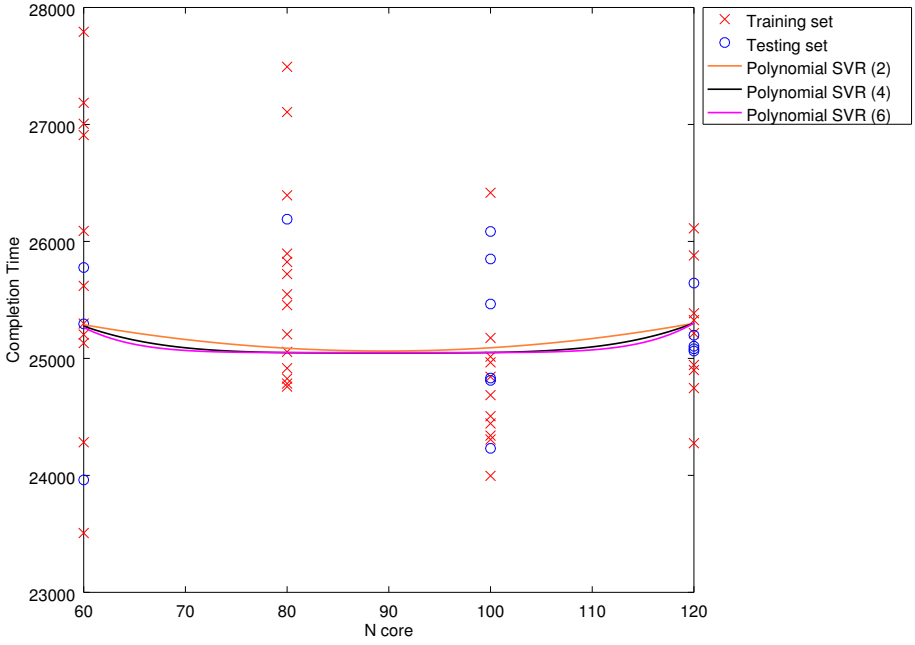

Figure 37: Completion time vs ncores for query R5 with datasize 250GB

2.5.2 Query R5 — Datasize 500GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	1.1147	-0.0213	1183	0.0475
Linear SVR	1.0929	0.1679	1192	0.0484
Polynomial SVR (2)	1.1591	0.0649	1276	0.0515
Polynomial SVR (3)	1.0733	0.1551	1170	0.0476
Polynomial SVR (4)	1.1591	0.0795	1276	0.0515
Polynomial SVR (6)	1.1530	0.0940	1265	0.0510
Gaussian SVR	1.1295	0.0016	1273	0.0518

Table 38: Results for R5-500GB, only ncores

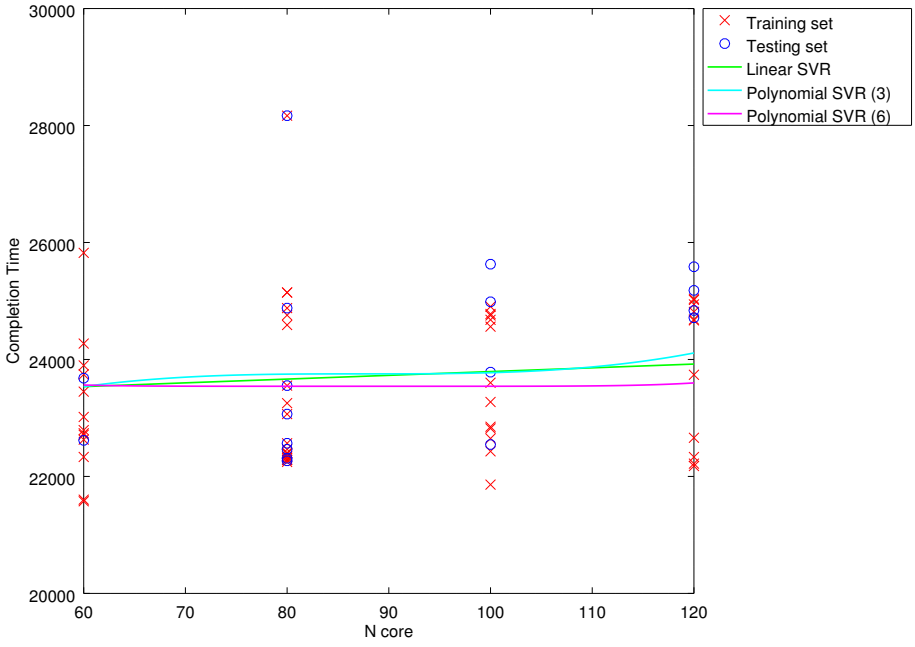

Figure 38: Completion time vs ncores for query R5 with datasize 500GB

2.5.3 Query R5 — Datasize 750GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.8928	-0.0234	1062	0.0443
Linear SVR	1.0199	-0.0000	983	0.0425
Polynomial SVR (2)	1.0199	-0.0000	983	0.0425
Polynomial SVR (3)	1.0199	-0.0000	983	0.0425
Polynomial SVR (4)	1.0199	-0.0000	983	0.0425
Polynomial SVR (6)	1.0199	-0.0000	983	0.0425
Gaussian SVR	1.0199	-0.0000	983	0.0425

Table 39: Results for R5-750GB, only ncores

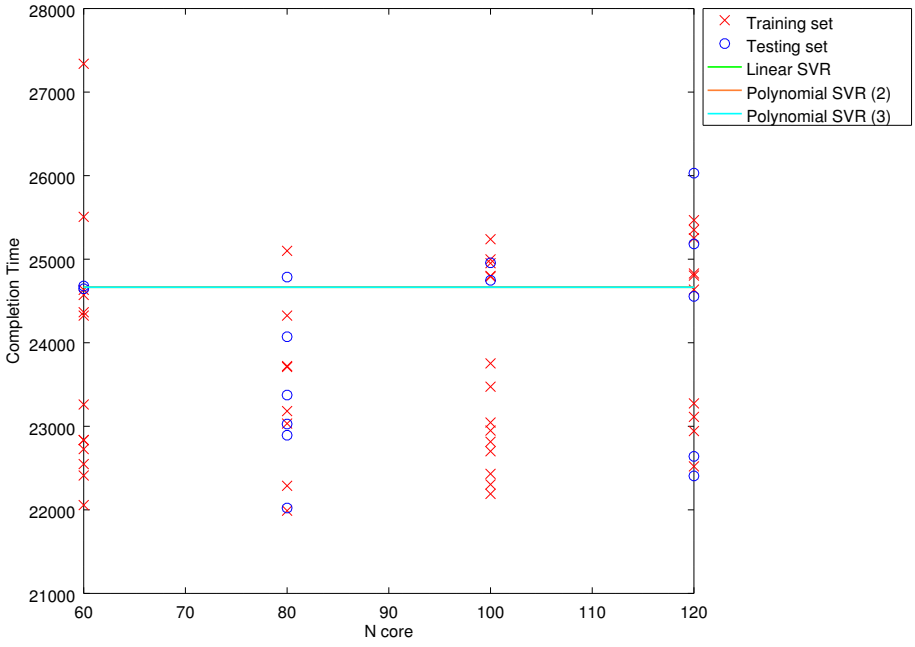

Figure 39: Completion time vs ncores for query R5 with datasize 750GB

2.5.4 Query R5 — Datasize 1000GB

2.5.5 Query R5 — Datasize 1000GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.8628	0.1111	3395	0.0850
Linear SVR	0.8578	0.1324	3181	0.0786
Polynomial SVR (2)	0.8655	0.1903	2222	0.0509
Polynomial SVR (3)	0.8833	0.1126	3533	0.0887
Polynomial SVR (4)	0.8454	0.2116	2146	0.0492
Polynomial SVR (6)	0.8380	0.2376	2150	0.0493
Gaussian SVR	0.8023	0.3023	2506	0.0595

Table 40: Results for R5-1000GB, only ncores

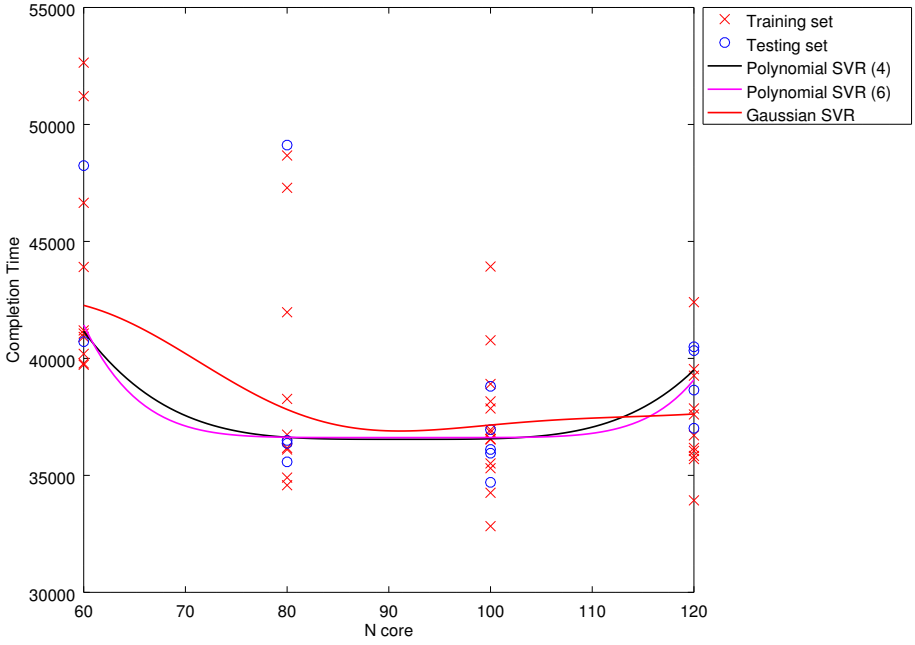

Figure 40: Completion time vs ncores for query R5 with datasize 1000GB

3 Fixed Datasize, $ncores^{-1}$, all the features

3.1 Query R1

3.1.1 Query R1 — Datasize 250GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.3569	0.9437	2586	0.0332
Linear SVR	0.3897	0.9727	2836	0.0362
Polynomial SVR (2)	0.7267	0.7759	10787	0.2045
Polynomial SVR (3)	3.6861	0.7898	20915	0.2254
Polynomial SVR (4)	2.1868	0.2291	18693	0.2484
Polynomial SVR (6)	3.1348	0.3153	25360	0.3269
Gaussian SVR	1.2154	0.4402	7712	0.0841

Table 41: Results for R1-250GB with the nonlinear 1/ncores feature

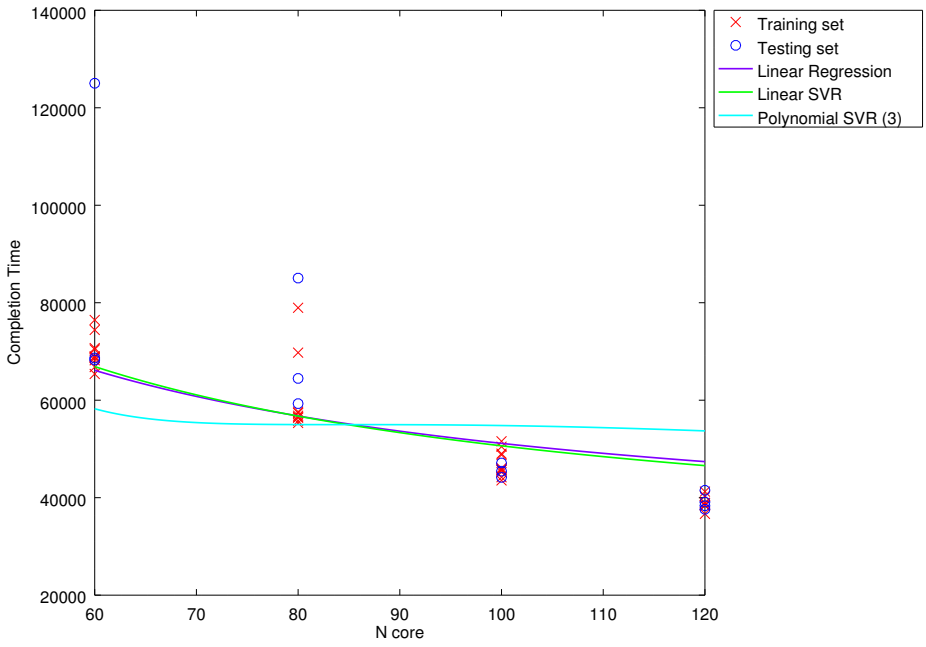

Figure 41: Completion time vs ncores for query R1 with datasize 250GB

3.1.2 Query R1 — Datasize 500GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.0878	0.9957	5544	0.0283
Linear SVR	0.0994	0.9979	6748	0.0357
Polynomial SVR (2)	0.5341	0.9077	32780	0.1738
Polynomial SVR (3)	0.1464	0.9900	8855	0.0540
Polynomial SVR (4)	0.6090	0.8974	38612	0.2160
Polynomial SVR (6)	1.0016	0.8146	55698	0.2804
Gaussian SVR	0.3567	0.9755	18202	0.0825

Table 42: Results for R1-500GB with the nonlinear 1/ncores feature

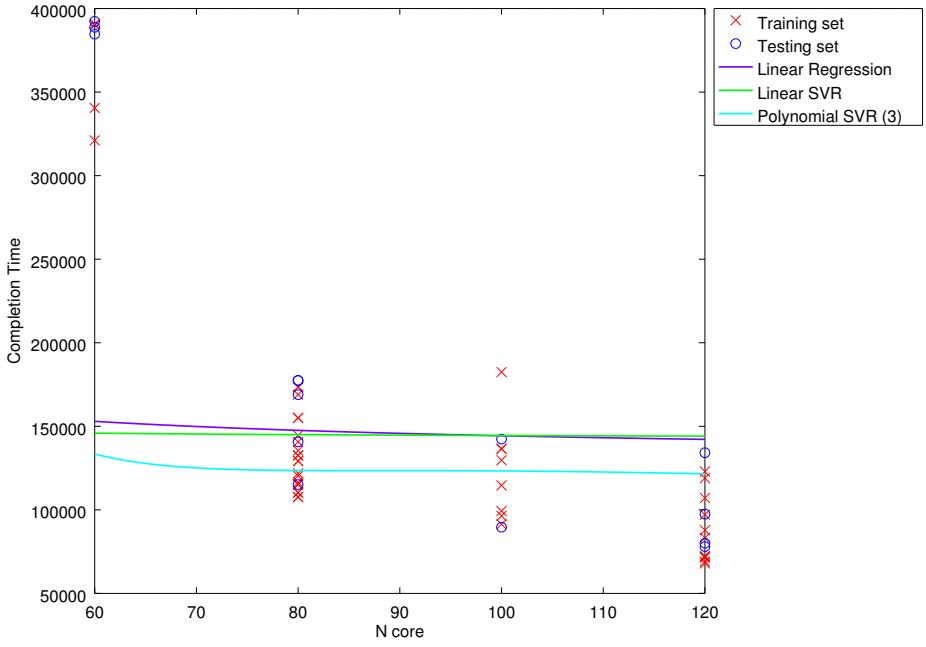

Figure 42: Completion time vs ncores for query R1 with datasize 500GB

3.1.3 Query R1 — Datasize 750GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.1409	0.9526	10604	0.0437
Linear SVR	0.1435	0.9626	10499	0.0427
Polynomial SVR (2)	0.6469	0.2361	42774	0.1805
Polynomial SVR (3)	0.4315	0.5604	27541	0.1053
Polynomial SVR (4)	0.8832	0.0817	53947	0.2200
Polynomial SVR (6)	0.5945	0.2311	42316	0.1704
Gaussian SVR	0.1752	0.9606	11544	0.0496

Table 43: Results for R1-750GB with the nonlinear 1/ncores feature

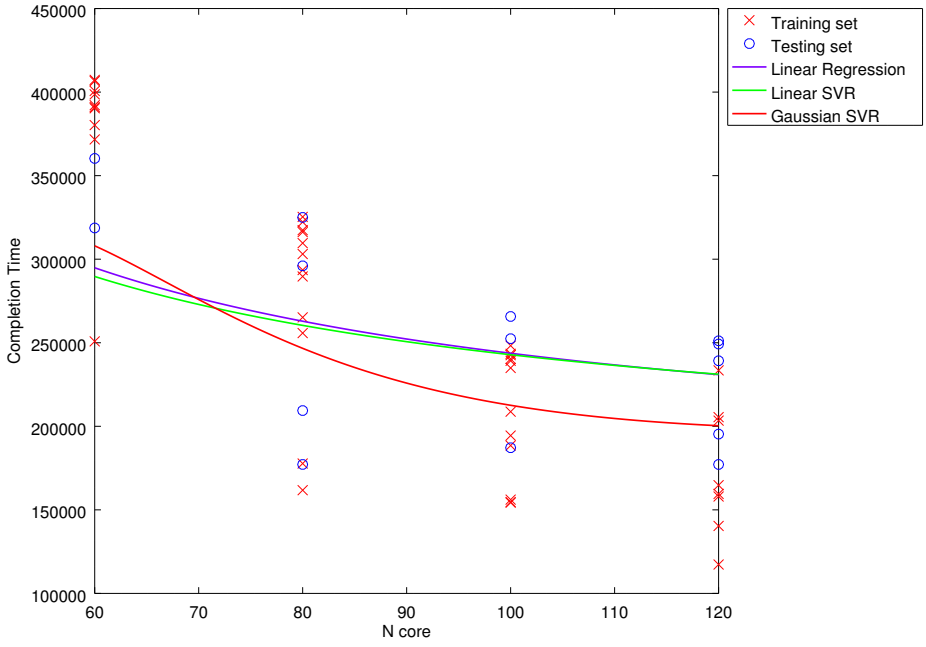

Figure 43: Completion time vs ncores for query R1 with datasize 750GB

3.1.4 Query R1 — Datasize 1000GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.1575	0.9544	10766	0.0283
Linear SVR	0.1449	0.9662	10234	0.0269
Polynomial SVR (2)	0.6723	0.3832	49056	0.1358
Polynomial SVR (3)	0.3719	0.7768	28996	0.0774
Polynomial SVR (4)	0.9733	0.0448	54529	0.1555
Polynomial SVR (6)	0.9608	0.0447	55540	0.1577
Gaussian SVR	0.3636	0.8120	15303	0.0446

Table 44: Results for R1-1000GB with the nonlinear 1/ncores feature

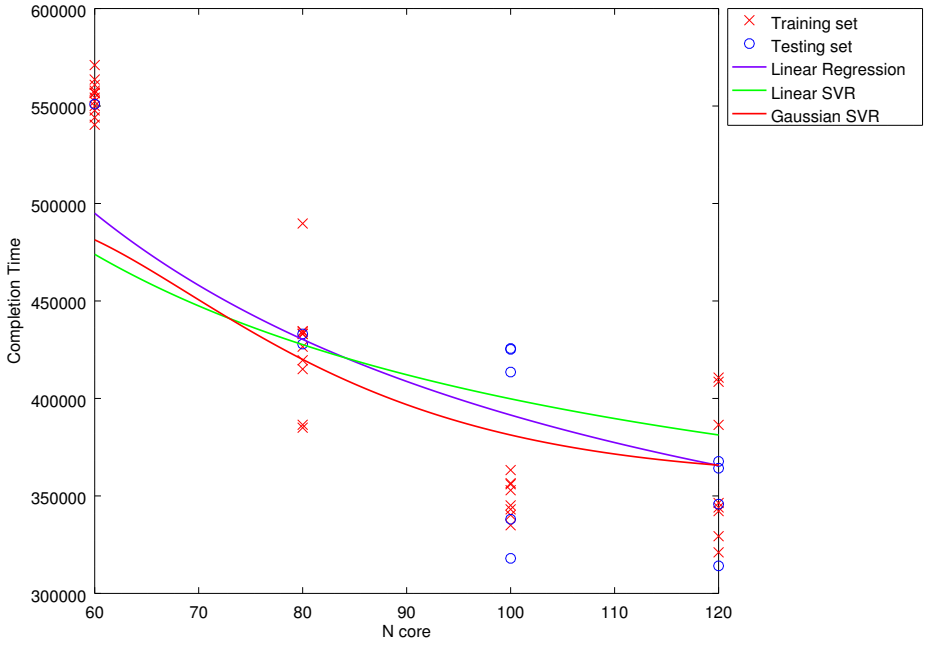

Figure 44: Completion time vs ncores for query R1 with datasize 1000GB

3.2 Query R2

3.2.1 Query R2 — Datasize 250GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.3116	0.8948	656	0.0079
Linear SVR	0.2890	0.9187	608	0.0073
Polynomial SVR (2)	1.0484	0.3489	2432	0.0302
Polynomial SVR (3)	0.5203	0.7803	1223	0.0153
Polynomial SVR (4)	0.9899	0.1627	2283	0.0284
Polynomial SVR (6)	1.0160	0.1256	2307	0.0287
Gaussian SVR	0.4203	0.8942	1005	0.0124

Table 45: Results for R2-250GB with the nonlinear 1/ncores feature

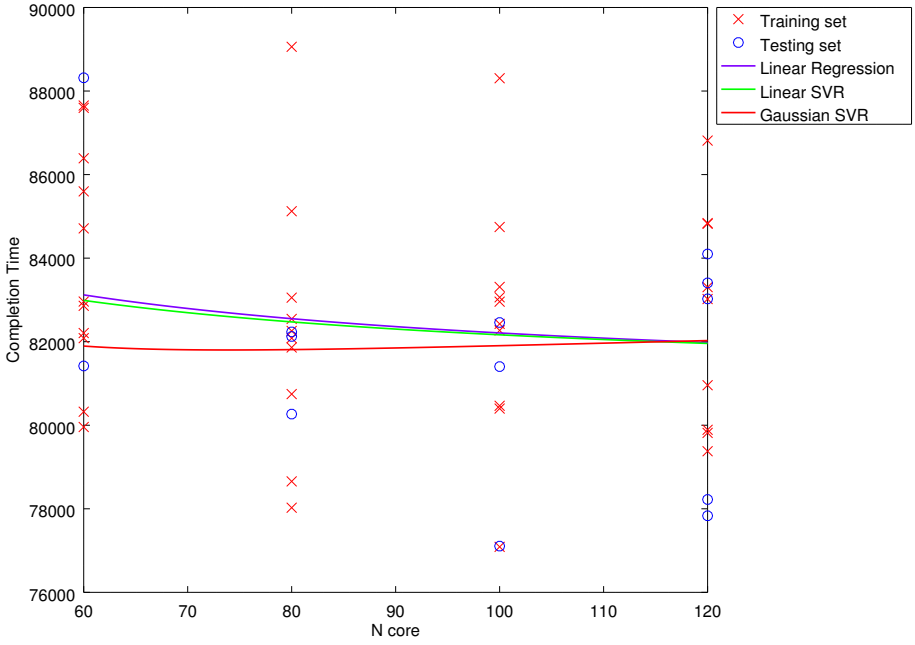

Figure 45: Completion time vs ncores for query R2 with datasize 250GB

3.2.2 Query R2 — Datasize 500GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.1459	0.9517	284	0.0040
Linear SVR	0.1396	0.9640	293	0.0041
Polynomial SVR (2)	0.6872	0.3168	1249	0.0177
Polynomial SVR (3)	0.3714	0.7477	824	0.0116
Polynomial SVR (4)	0.6409	0.5494	1218	0.0172
Polynomial SVR (6)	0.5920	0.5919	1350	0.0190
Gaussian SVR	0.2339	0.8938	536	0.0074

Table 46: Results for R2-500GB with the nonlinear 1/ncores feature

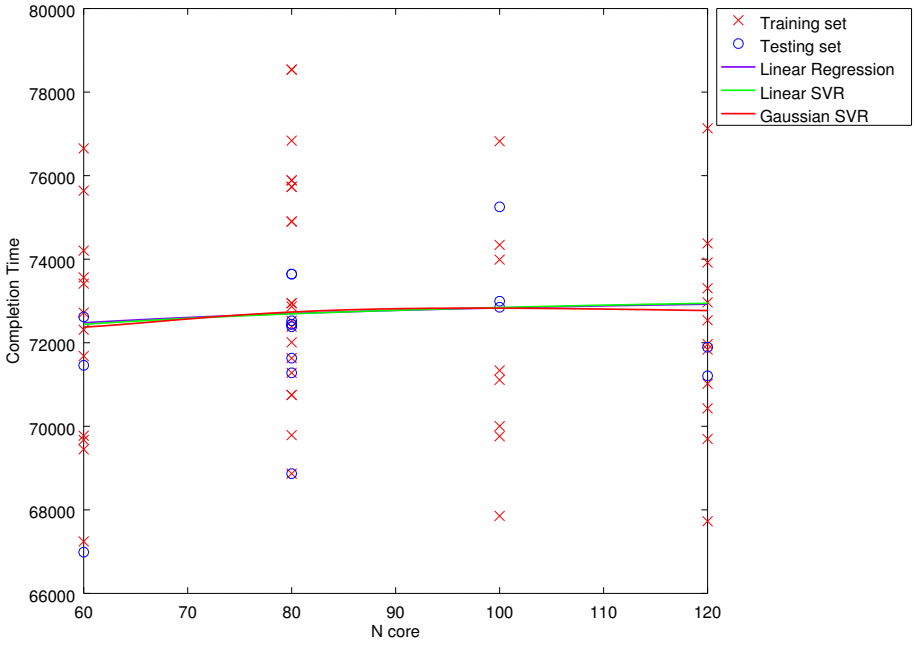

Figure 46: Completion time vs ncores for query R2 with datasize 500GB

3.2.3 Query R2 — Datasize 750GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.2407	0.9575	472	0.0060
Linear SVR	0.2404	0.9726	476	0.0060
Polynomial SVR (2)	1.2193	0.0000	2315	0.0292
Polynomial SVR (3)	0.6488	0.8600	1046	0.0133
Polynomial SVR (4)	1.2193	0.0000	2315	0.0292
Polynomial SVR (6)	1.2193	0.0000	2315	0.0292
Gaussian SVR	0.5945	0.8089	1108	0.0139

Table 47: Results for R2-750GB with the nonlinear 1/ncores feature

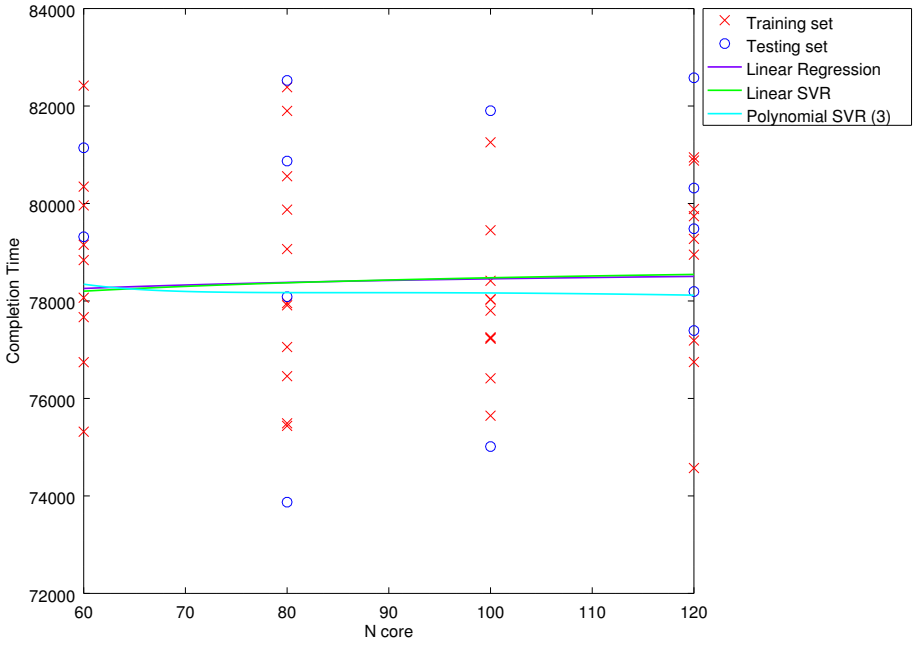

Figure 47: Completion time vs ncores for query R2 with datasize 750GB

3.2.4 Query R2 — Datasize 1000GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.0534	0.9971	21205	0.0221
Linear SVR	0.0843	0.9942	37707	0.0421
Polynomial SVR (2)	0.7009	0.5037	336756	0.4033
Polynomial SVR (3)	0.5103	0.8111	206774	0.2177
Polynomial SVR (4)	0.8812	0.5707	374286	0.4127
Polynomial SVR (6)	0.8290	0.5363	351110	0.4688
Gaussian SVR	0.2298	0.9590	93580	0.1148

Table 48: Results for R2-1000GB with the nonlinear 1/ncores feature

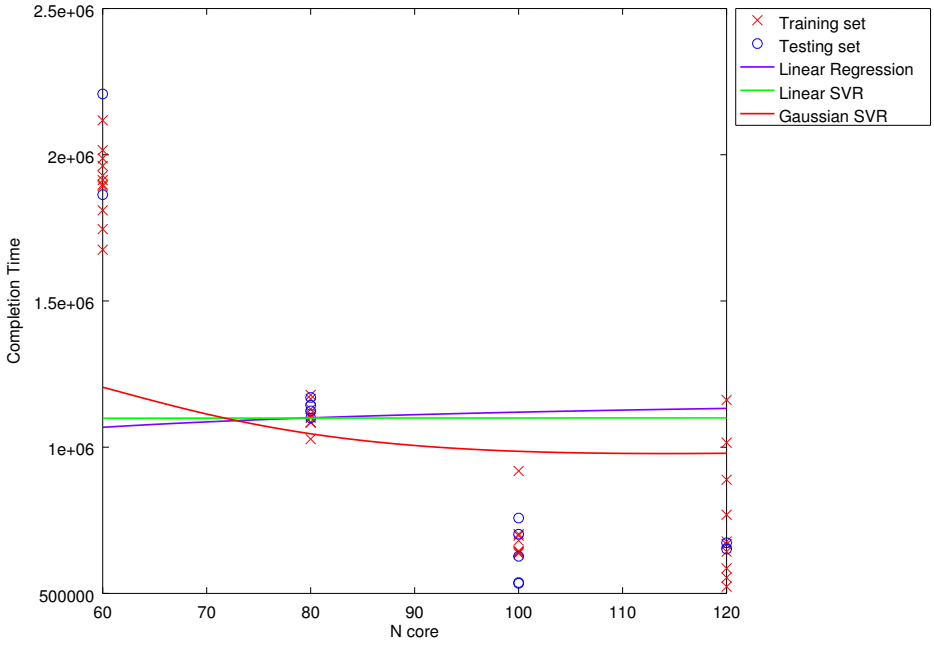

Figure 48: Completion time vs ncores for query R2 with datasize 1000GB

3.3 Query R3

3.3.1 Query R3 — Datasize 250GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.0780	0.9894	3915	0.0252
Linear SVR	0.1057	0.9816	5093	0.0299
Polynomial SVR (2)	0.7201	0.1170	41915	0.2811
Polynomial SVR (3)	0.2708	0.9232	15247	0.0945
Polynomial SVR (4)	0.6645	0.2415	36141	0.2452
Polynomial SVR (6)	0.6164	0.4095	34311	0.2194
Gaussian SVR	0.1788	0.9665	8642	0.0656

Table 49: Results for R3-250GB with the nonlinear 1/ncores feature

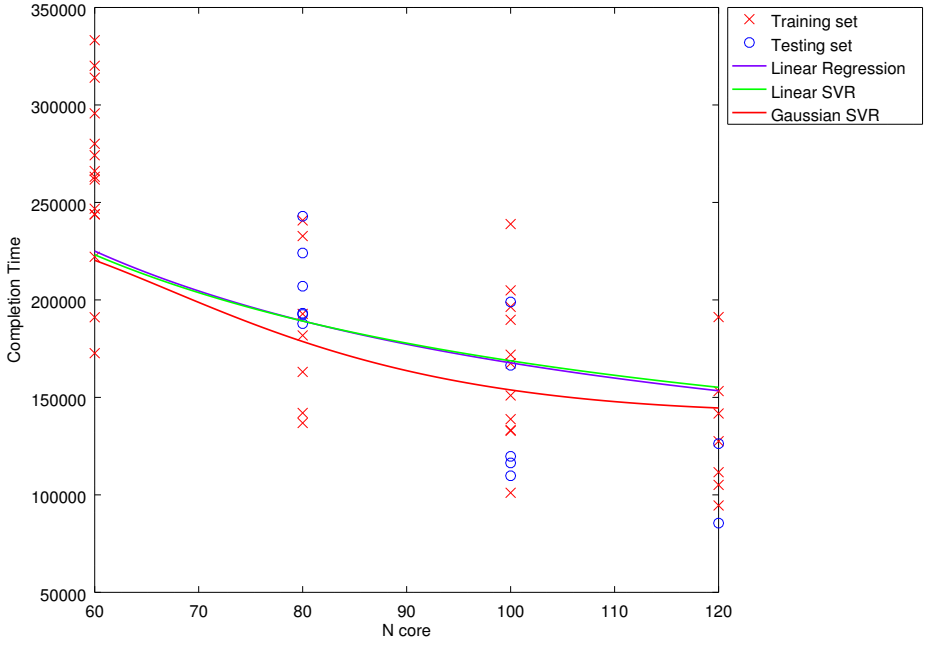

Figure 49: Completion time vs ncores for query R3 with datasize 250GB

3.3.2 Query R3 — Datasize 500GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.0539	0.9962	10845	0.0220
Linear SVR	0.0626	0.9949	12632	0.0269
Polynomial SVR (2)	0.2821	0.8979	59067	0.1231
Polynomial SVR (3)	0.1297	0.9782	32168	0.0654
Polynomial SVR (4)	0.1922	0.9545	40999	0.0889
Polynomial SVR (6)	0.2877	0.8937	54259	0.1205
Gaussian SVR	0.0703	0.9939	15349	0.0321

Table 50: Results for R3-500GB with the nonlinear 1/ncores feature

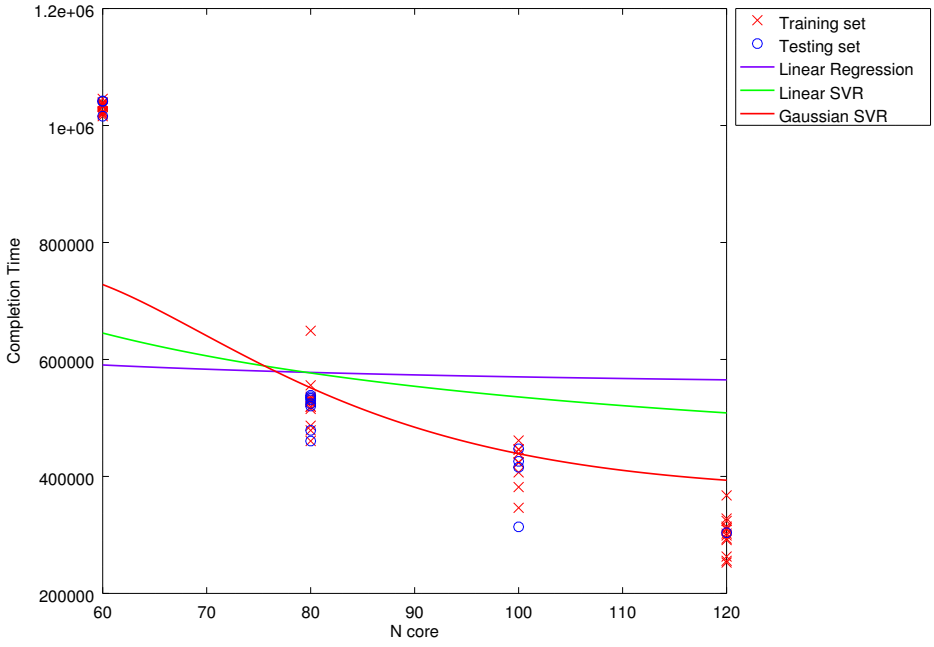

Figure 50: Completion time vs ncores for query R3 with datasize 500GB

3.3.3 Query R3 — Datasize 750GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.0209	0.9996	2658	0.0033
Linear SVR	0.0806	0.9961	11228	0.0149
Polynomial SVR (2)	1.0051	0.5158	140581	0.1721
Polynomial SVR (3)	0.4684	0.9239	65851	0.0849
Polynomial SVR (4)	0.7927	0.6494	111418	0.1502
Polynomial SVR (6)	0.6810	0.6845	95464	0.1247
Gaussian SVR	0.2061	0.9777	26875	0.0359

Table 51: Results for R3-750GB with the nonlinear 1/ncores feature

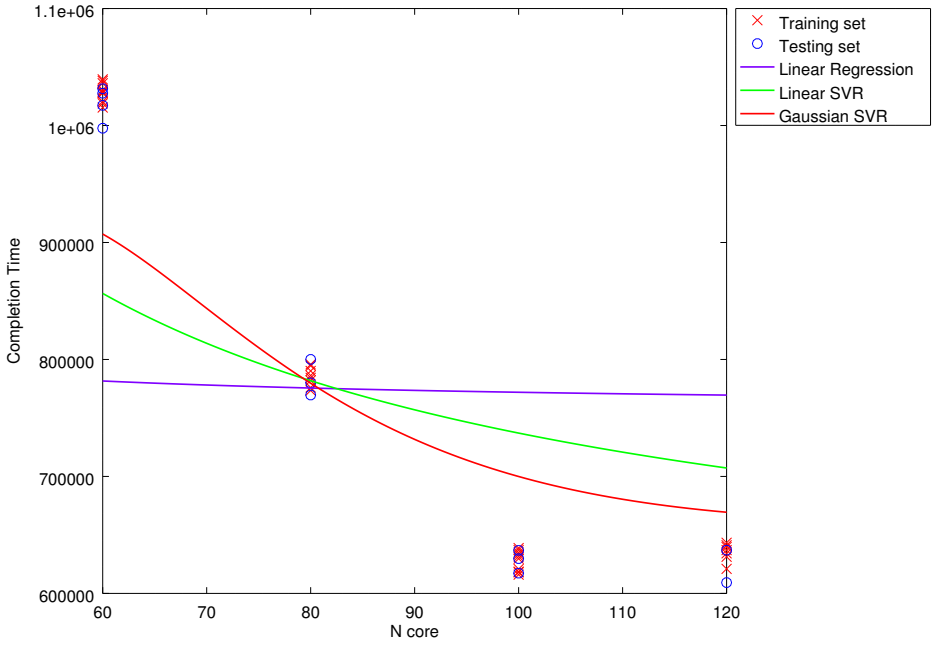

Figure 51: Completion time vs ncores for query R3 with datasize 750GB

3.3.4 Query R3 — Datasize 1000GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.0723	0.9956	10267	0.0094
Linear SVR	0.0959	0.9948	20283	0.0199
Polynomial SVR (2)	0.6954	0.6772	148312	0.1435
Polynomial SVR (3)	0.3570	0.9367	58398	0.0507
Polynomial SVR (4)	0.8566	0.5419	174113	0.1625
Polynomial SVR (6)	0.8709	0.4828	171350	0.1591
Gaussian SVR	0.2990	0.9557	49677	0.0438

Table 52: Results for R3-1000GB with the nonlinear 1/ncores feature

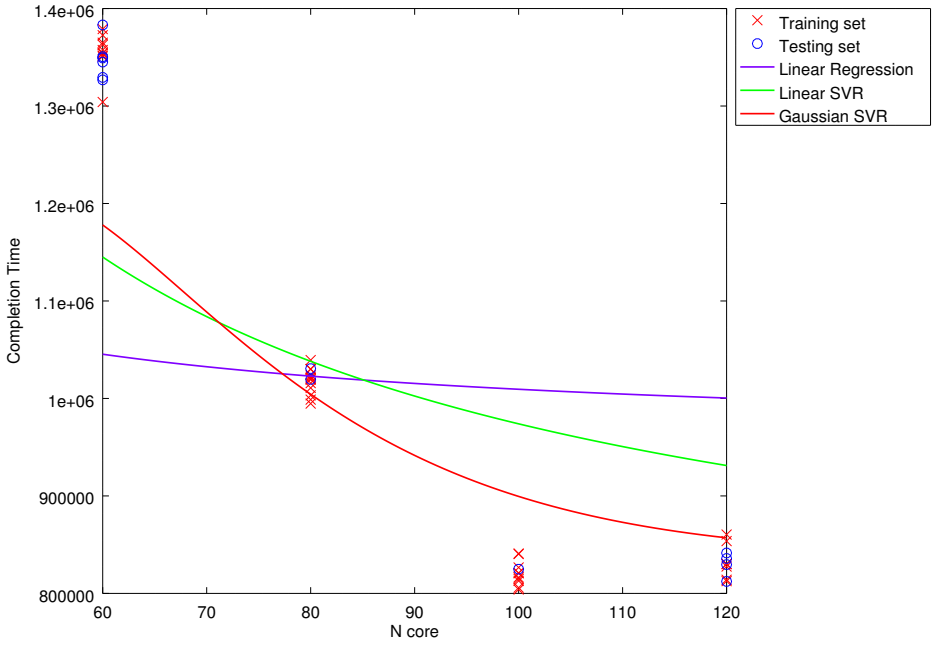

Figure 52: Completion time vs ncores for query R3 with datasize 1000GB

3.4 Query R4

3.4.1 Query R4 — Datasize 250GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.1407	0.9793	4859	0.0282
Linear SVR	0.1048	0.9893	4109	0.0239
Polynomial SVR (2)	0.8143	0.4103	33794	0.2429
Polynomial SVR (3)	0.4882	0.8326	18105	0.1119
Polynomial SVR (4)	0.8747	0.5418	37483	0.2594
Polynomial SVR (6)	0.8999	0.5869	38677	0.2634
Gaussian SVR	0.2798	0.9470	9282	0.0572

Table 53: Results for R4-250GB with the nonlinear 1/ncores feature

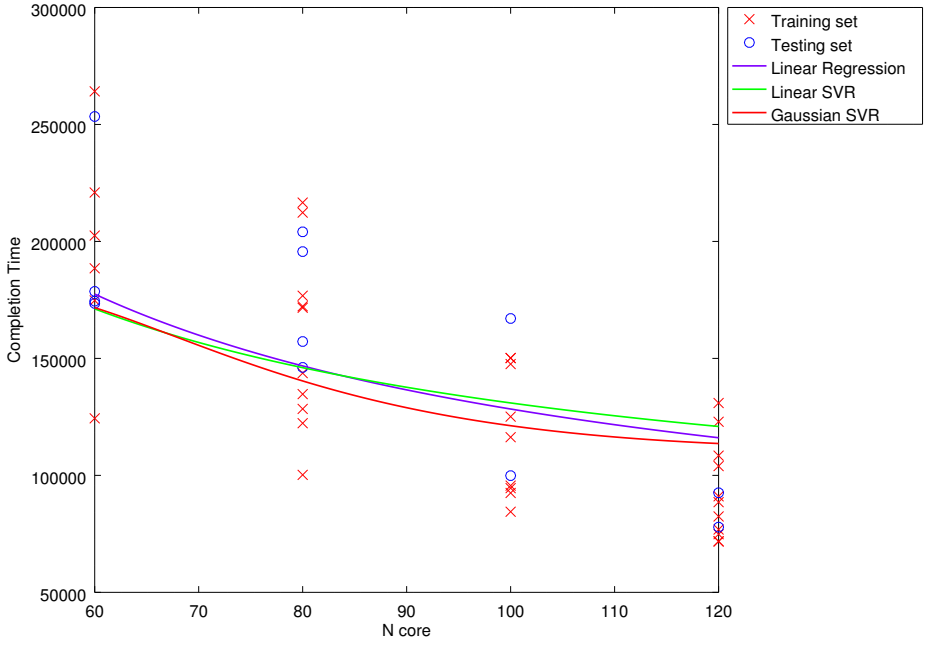

Figure 53: Completion time vs ncores for query R4 with datasize 250GB

3.4.2 Query R4 — Datasize 500GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.0501	0.9946	5178	0.0145
Linear SVR	0.0850	0.9872	14922	0.0397
Polynomial SVR (2)	0.2128	0.9400	29712	0.0914
Polynomial SVR (3)	0.1287	0.9766	19858	0.0507
Polynomial SVR (4)	0.2801	0.9035	42883	0.1152
Polynomial SVR (6)	0.2924	0.8767	40499	0.1112
Gaussian SVR	0.0791	0.9876	11411	0.0287

Table 54: Results for R4-500GB with the nonlinear 1/ncores feature

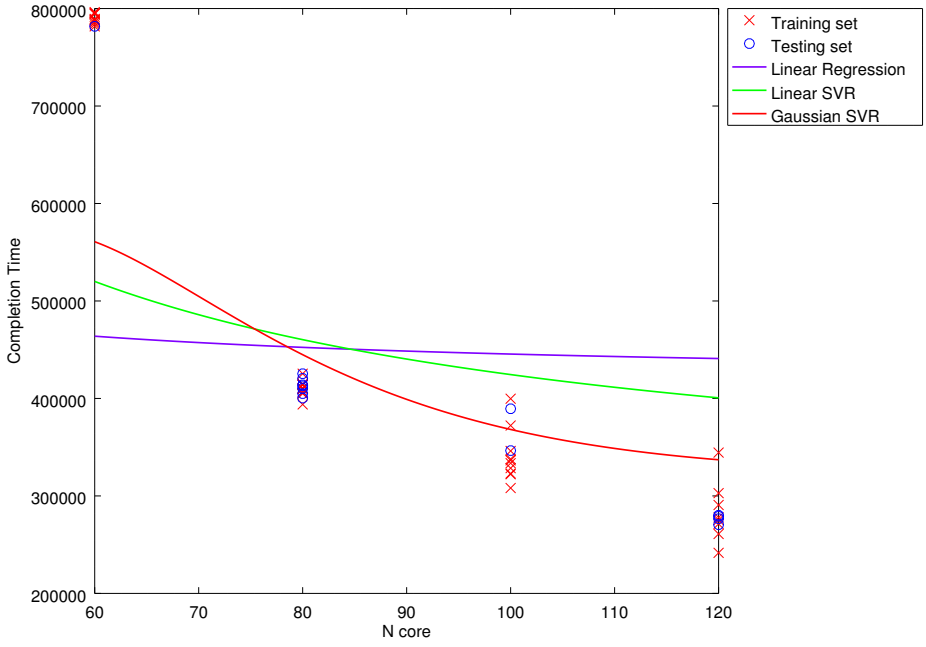

Figure 54: Completion time vs ncores for query R4 with datasize 500GB

3.4.3 Query R4 — Datasize 750GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.0216	0.9996	2333	0.0038
Linear SVR	0.0773	0.9968	8179	0.0146
Polynomial SVR (2)	0.6766	0.7235	82143	0.1361
Polynomial SVR (3)	0.3497	0.9401	37764	0.0591
Polynomial SVR (4)	0.7932	0.7572	91479	0.1430
Polynomial SVR (6)	0.9234	0.6666	101239	0.1550
Gaussian SVR	0.1553	0.9830	13013	0.0221

Table 55: Results for R4-750GB with the nonlinear 1/ncores feature

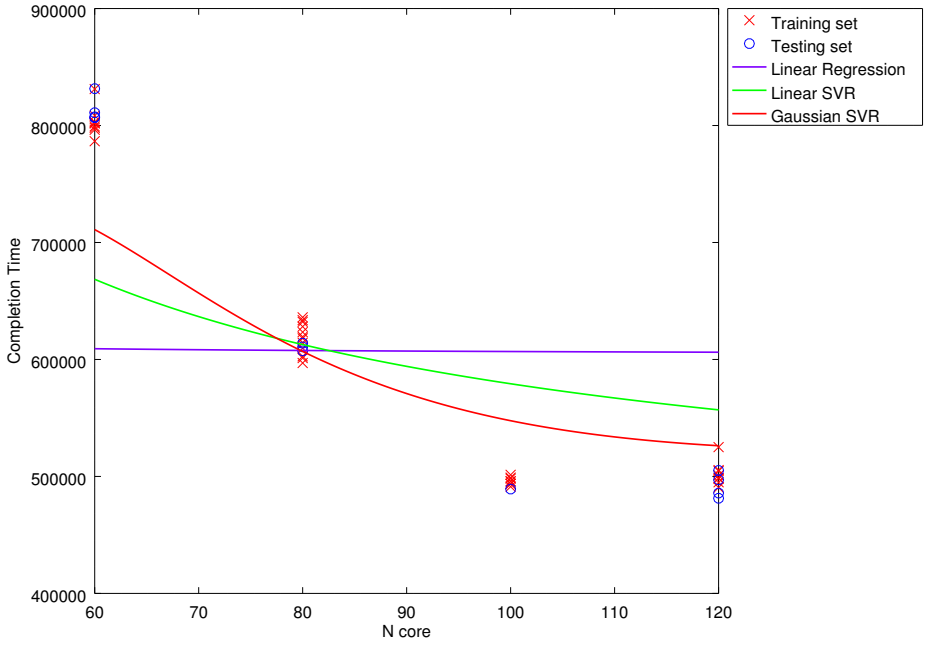

Figure 55: Completion time vs ncores for query R4 with datasize 750GB

3.4.4 Query R4 — Datasize 1000GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.0812	0.9944	45917	0.0244
Linear SVR	0.0785	0.9966	46627	0.0315
Polynomial SVR (2)	0.7046	0.6535	446048	0.2688
Polynomial SVR (3)	0.3197	0.9286	196869	0.1466
Polynomial SVR (4)	0.8131	0.7058	578071	0.4145
Polynomial SVR (6)	0.9574	0.5389	680038	0.4961
Gaussian SVR	0.2282	0.9582	129449	0.0703

Table 56: Results for R4-1000GB with the nonlinear 1/ncores feature

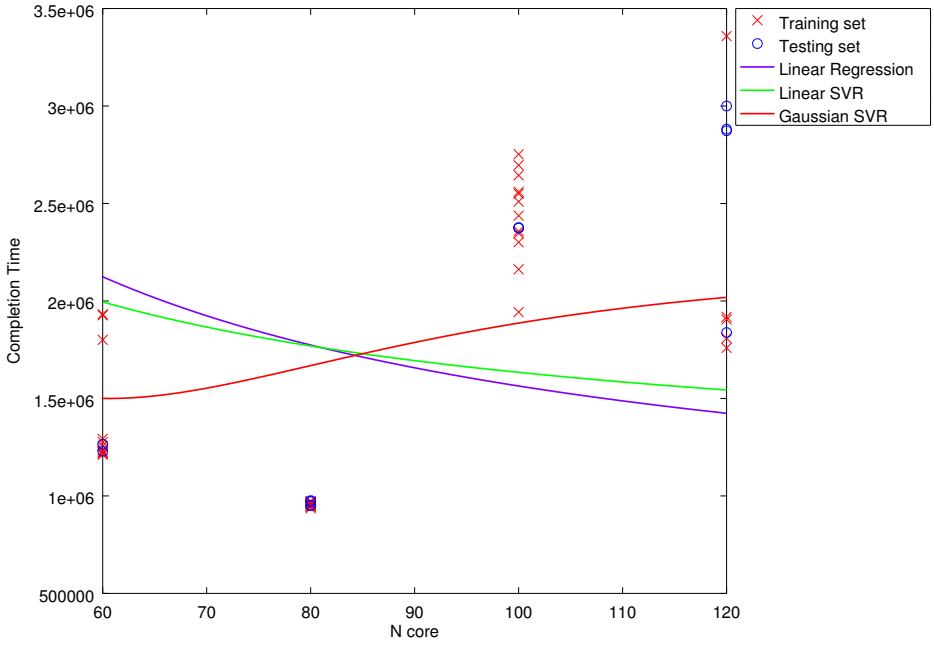

Figure 56: Completion time vs ncores for query R4 with datasize 1000GB

3.5 Query R5

3.5.1 Query R5 — Datasize 250GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	1.1046	-0.0599	720	0.0277
Linear SVR	1.0254	0.3409	642	0.0247
Polynomial SVR (2)	1.1448	0.0293	773	0.0301
Polynomial SVR (3)	1.0585	0.1048	706	0.0273
Polynomial SVR (4)	1.0900	0.0650	729	0.0283
Polynomial SVR (6)	1.0613	0.0496	679	0.0264
Gaussian SVR	0.9708	0.3643	606	0.0234

Table 57: Results for R5-250GB with the nonlinear 1/ncores feature

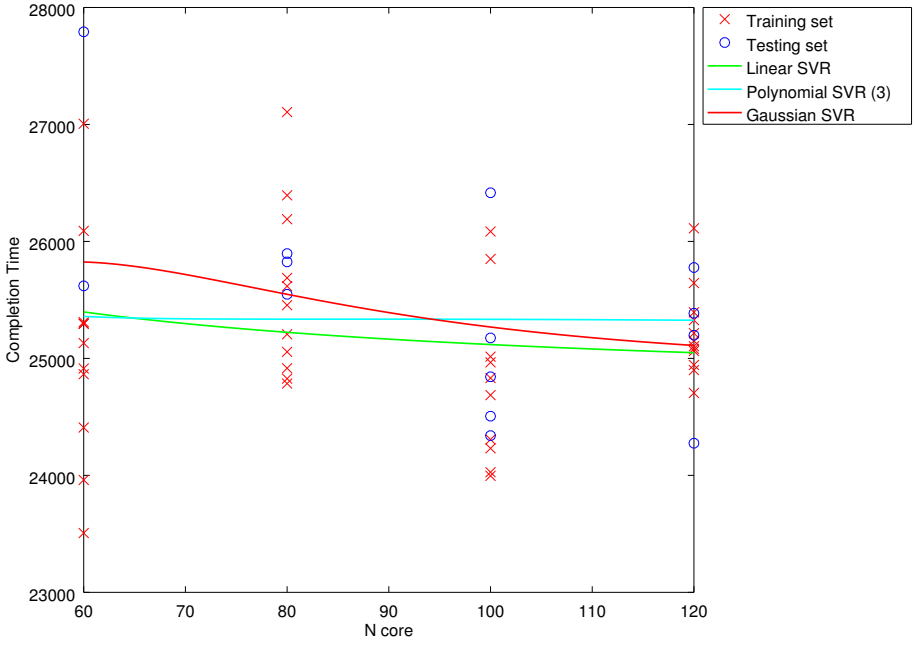

Figure 57: Completion time vs ncores for query R5 with datasize 250GB

3.5.2 Query R5 — Datasize 500GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.4987	0.5323	389	0.0164
Linear SVR	0.4932	0.5986	381	0.0161
Polynomial SVR (2)	0.7350	0.2975	705	0.0297
Polynomial SVR (3)	0.6196	0.4052	521	0.0220
Polynomial SVR (4)	0.7239	0.2502	636	0.0265
Polynomial SVR (6)	0.6725	0.1906	683	0.0289
Gaussian SVR	0.5501	0.5643	416	0.0175

Table 58: Results for R5-500GB with the nonlinear 1/ncores feature

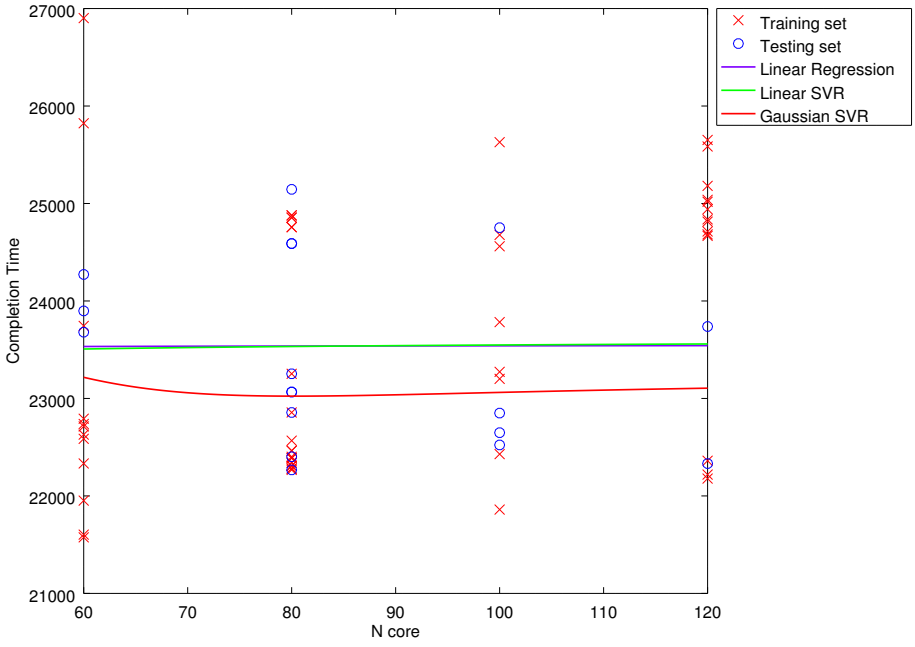

Figure 58: Completion time vs ncores for query R5 with datasize 500GB

3.5.3 Query R5 — Datasize 750GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.5758	0.6745	442	0.0177
Linear SVR	0.6121	0.7041	466	0.0186
Polynomial SVR (2)	2.5714	0.0675	1530	0.0618
Polynomial SVR (3)	0.8141	0.5165	830	0.0340
Polynomial SVR (4)	1.9622	0.0219	1418	0.0570
Polynomial SVR (6)	2.1235	0.0003	1574	0.0630
Gaussian SVR	0.6321	0.6870	486	0.0195

Table 59: Results for R5-750GB with the nonlinear 1/ncores feature

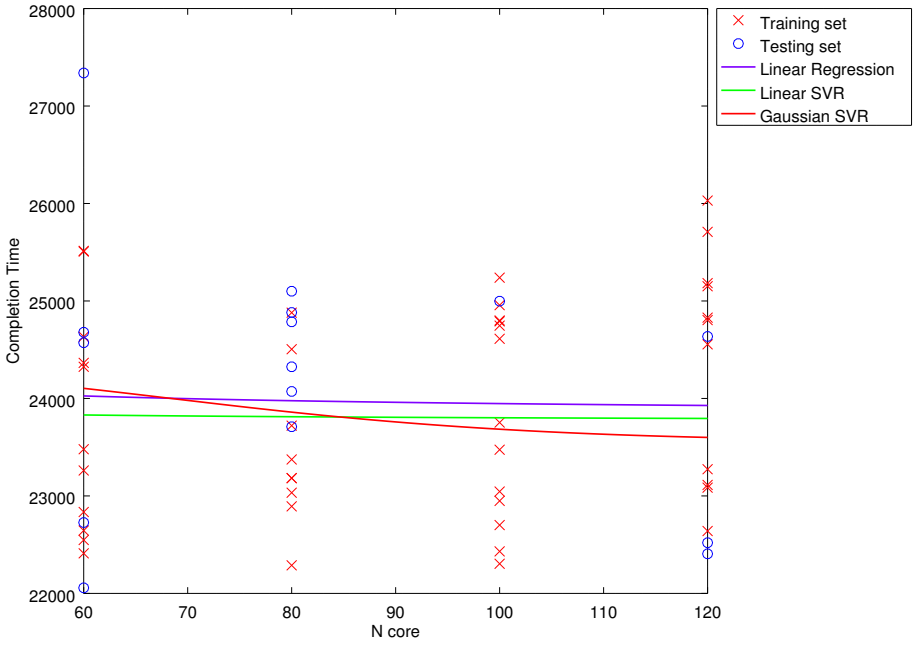

Figure 59: Completion time vs ncores for query R5 with datasize 750GB

3.5.4 Query R5 — Datasize 1000GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.5047	0.1390	1333	0.0341
Linear SVR	0.4982	0.3487	1237	0.0316
Polynomial SVR (2)	0.4124	0.5666	873	0.0219
Polynomial SVR (3)	0.6250	0.2996	1597	0.0404
Polynomial SVR (4)	0.5453	0.3911	1321	0.0332
Polynomial SVR (6)	0.5542	0.2966	1421	0.0358
Gaussian SVR	0.3568	0.5856	851	0.0217

Table 60: Results for R5-1000GB with the nonlinear 1/ncores feature

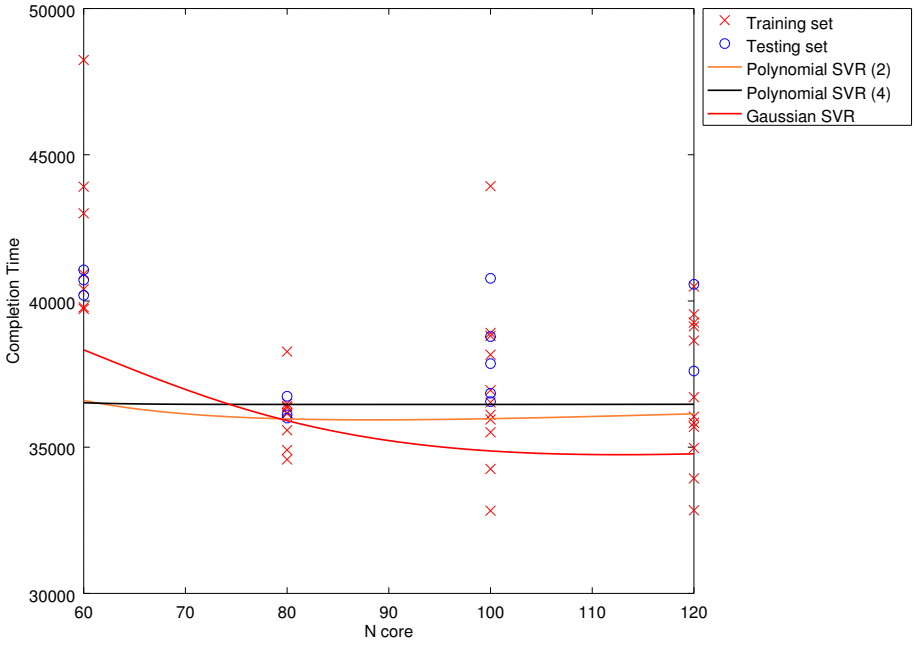

Figure 60: Completion time vs ncores for query R5 with datasize 1000GB

4 Fixed Datasize, $ncores^{-1}$, only $ncores$

4.1 Query R1

4.1.1 Query R1 — Datasize 250GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	1.2438	0.3259	12279	0.1330
Linear SVR	1.2946	0.3792	11957	0.1258
Polynomial SVR (2)	1.5679	0.0009	21296	0.3295
Polynomial SVR (3)	1.3800	0.2700	13385	0.1465
Polynomial SVR (4)	1.5428	0.0283	19555	0.2835
Polynomial SVR (6)	1.4977	0.0704	17963	0.2575
Gaussian SVR	1.2913	0.4092	11896	0.1246

Table 61: Results for R1-250GB with the nonlinear 1/ncores feature, only ncores

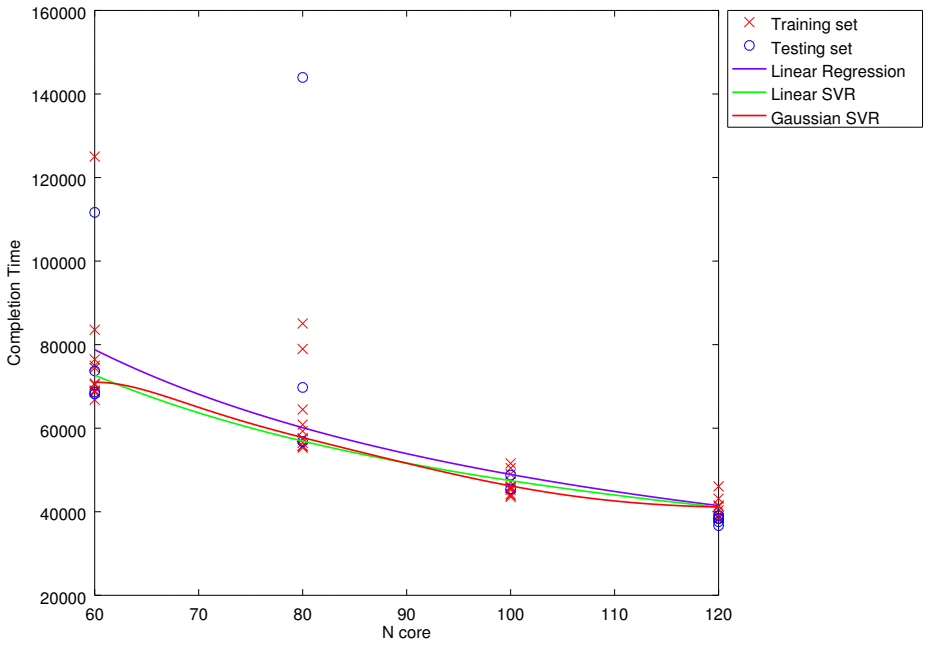

Figure 61: Completion time vs ncores for query R1 with datasize 250GB

4.1.2 Query R1 — Datasize 500GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.4321	0.8126	30237	0.2313
Linear SVR	0.4590	0.8228	29651	0.2051
Polynomial SVR (2)	0.7453	0.4527	56372	0.5019
Polynomial SVR (3)	0.3390	0.9148	25984	0.2401
Polynomial SVR (4)	0.5479	0.7103	41171	0.3775
Polynomial SVR (6)	0.4578	0.8009	34535	0.3294
Gaussian SVR	0.3841	0.9071	29393	0.2724

Table 62: Results for R1-500GB with the nonlinear 1/ncores feature, only ncores

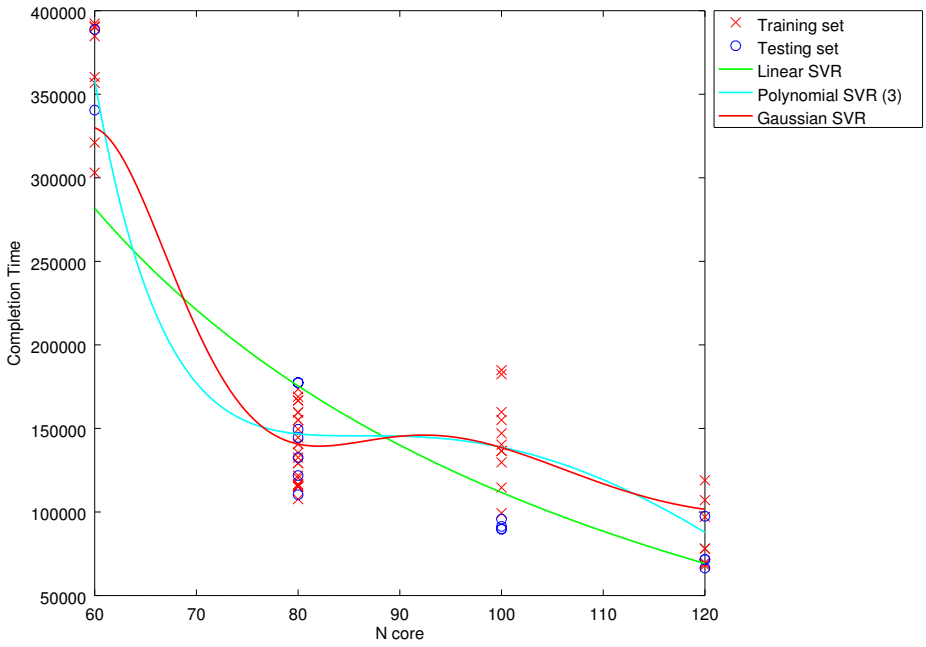

Figure 62: Completion time vs ncores for query R1 with datasize 500GB

4.1.3 Query R1 — Datasize 750GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.5526	0.8091	42884	0.2200
Linear SVR	0.5520	0.8666	39858	0.2325
Polynomial SVR (2)	0.9850	0.6193	74917	0.4174
Polynomial SVR (3)	0.5397	0.8813	38123	0.2276
Polynomial SVR (4)	0.8382	0.7906	57443	0.3611
Polynomial SVR (6)	0.7219	0.8510	49401	0.3107
Gaussian SVR	0.5938	0.8893	42964	0.2585

Table 63: Results for R1-750GB with the nonlinear 1/ncores feature, only ncores

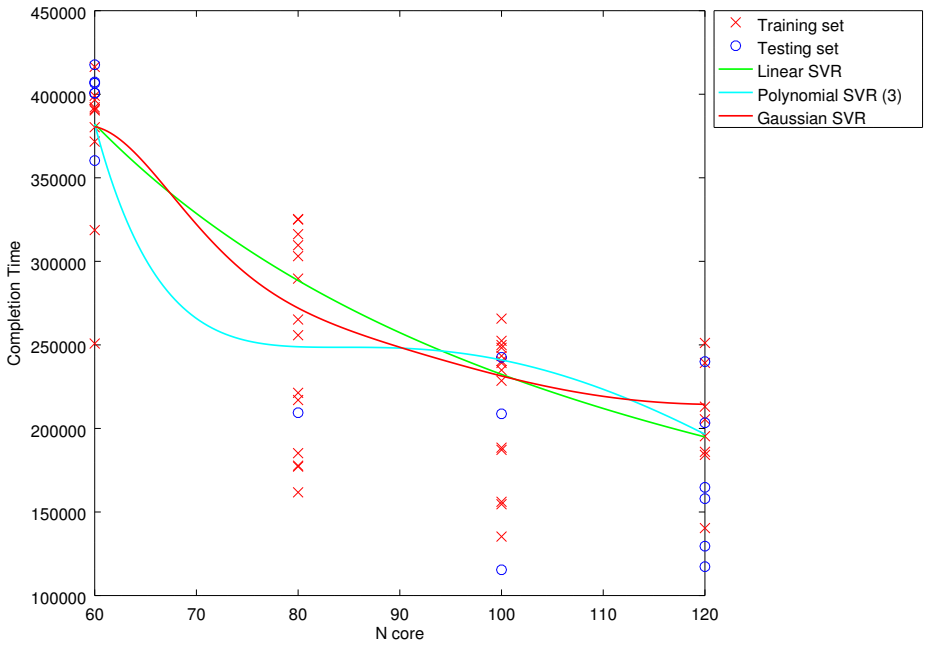

Figure 63: Completion time vs ncores for query R1 with datasize 750GB

4.1.4 Query R1 — Datasize 1000GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.3229	0.8946	22902	0.0553
Linear SVR	0.3305	0.8955	22857	0.0562
Polynomial SVR (2)	0.8438	0.4005	71595	0.1618
Polynomial SVR (3)	0.3683	0.8635	22807	0.0558
Polynomial SVR (4)	0.6342	0.6012	44393	0.1110
Polynomial SVR (6)	0.5642	0.6945	37885	0.0928
Gaussian SVR	0.3441	0.9190	26107	0.0602

Table 64: Results for R1-1000GB with the nonlinear 1/ncores feature, only ncores

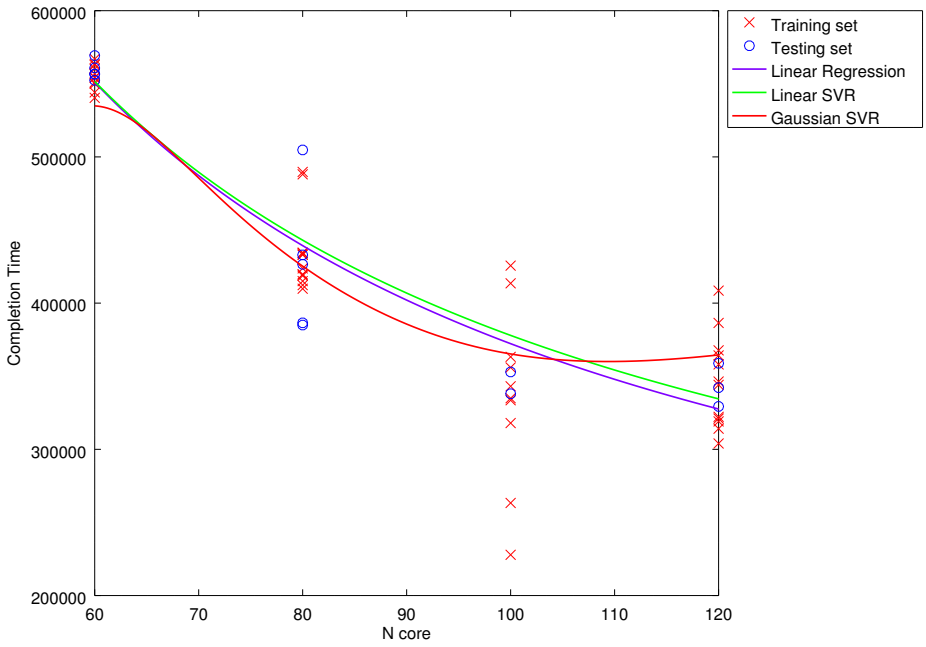

Figure 64: Completion time vs ncores for query R1 with datasize 1000GB

4.2 Query R2

4.2.1 Query R2 — Datasize 250GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.9065	-0.0568	2785	0.0339
Linear SVR	0.9191	0.0883	2817	0.0343
Polynomial SVR (2)	0.9837	0.0002	3054	0.0373
Polynomial SVR (3)	0.9247	0.0715	2809	0.0342
Polynomial SVR (4)	0.9515	0.0076	2885	0.0351
Polynomial SVR (6)	0.9422	0.0161	2813	0.0342
Gaussian SVR	0.9145	0.1079	2823	0.0344

Table 65: Results for R2-250GB with the nonlinear 1/ncores feature, only ncores

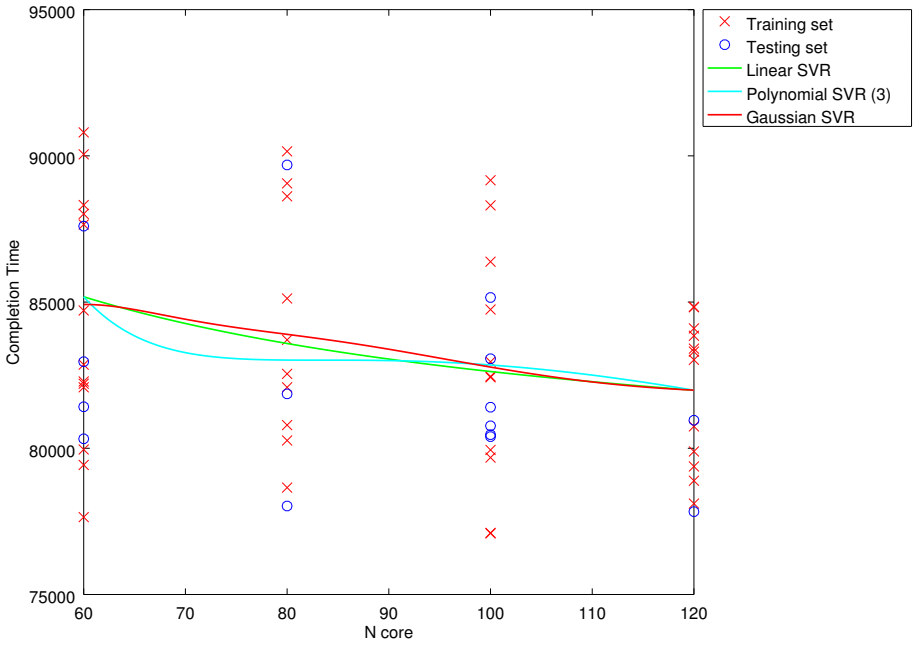

Figure 65: Completion time vs ncores for query R2 with datasize 250GB

4.2.2 Query R2 — Datasize 500GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	1.2224	-0.0227	2953	0.0406
Linear SVR	1.2263	0.0000	2982	0.0411
Polynomial SVR (2)	1.1928	0.1744	2978	0.0411
Polynomial SVR (3)	1.2252	0.0075	2985	0.0411
Polynomial SVR (4)	1.2031	0.1128	2964	0.0408
Polynomial SVR (6)	1.2064	0.0870	2956	0.0407
Gaussian SVR	1.2449	0.0043	3126	0.0432

Table 66: Results for R2-500GB with the nonlinear 1/ncores feature, only ncores

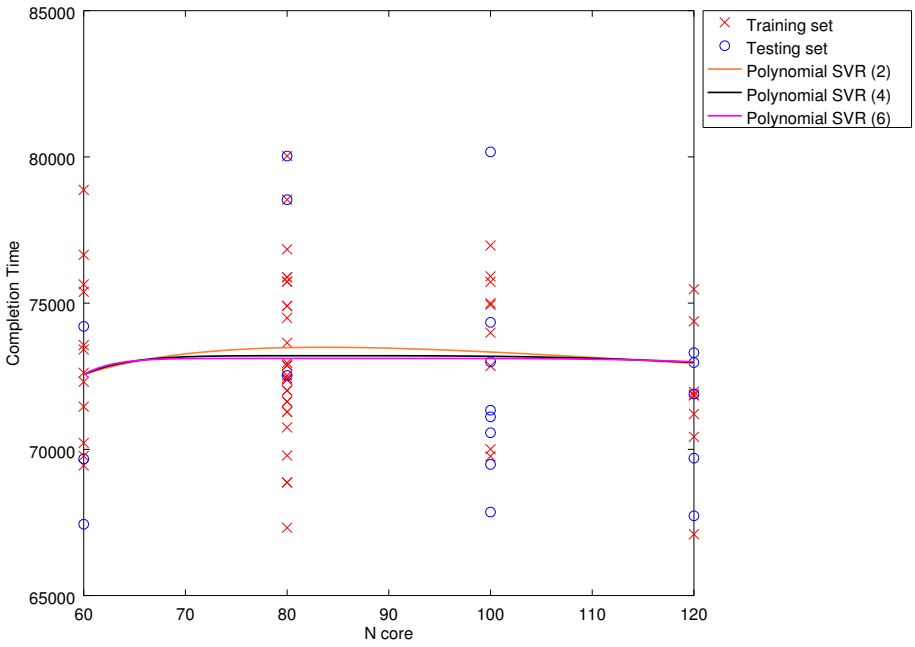

Figure 66: Completion time vs ncores for query R2 with datasize 500GB

4.2.3 Query R2 — Datasize 750GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	1.0119	-0.0671	2082	0.0266
Linear SVR	1.0201	0.0504	2009	0.0255
Polynomial SVR (2)	1.0331	0.0354	2078	0.0264
Polynomial SVR (3)	1.0269	0.0655	2050	0.0261
Polynomial SVR (4)	1.0306	0.0485	2052	0.0261
Polynomial SVR (6)	1.0210	0.0548	2045	0.0260
Gaussian SVR	1.0305	0.0202	2063	0.0262

Table 67: Results for R2-750GB with the nonlinear 1/ncores feature, only ncores

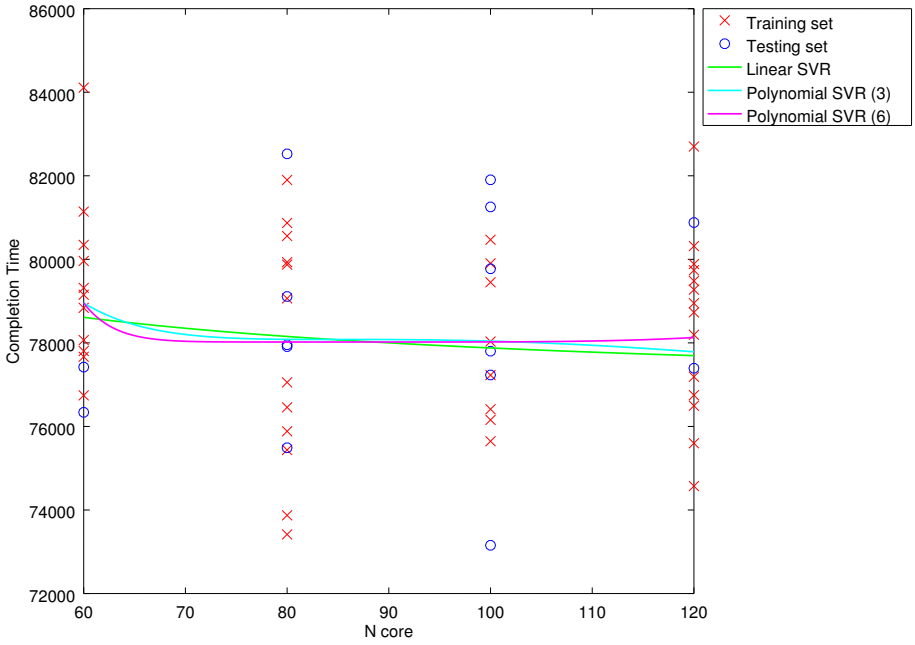

Figure 67: Completion time vs ncores for query R2 with datasize 750GB

4.2.4 Query R2 — Datasize 1000GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.3258	0.8006	157650	0.1926
Linear SVR	0.4842	0.8331	223250	0.2752
Polynomial SVR (2)	0.6883	0.1758	329819	0.3776
Polynomial SVR (3)	0.4559	0.7612	173610	0.2351
Polynomial SVR (4)	0.5423	0.4599	258020	0.2906
Polynomial SVR (6)	0.4626	0.6044	229115	0.2737
Gaussian SVR	0.1730	0.9691	69502	0.0818

Table 68: Results for R2-1000GB with the nonlinear 1/ncores feature, only ncores

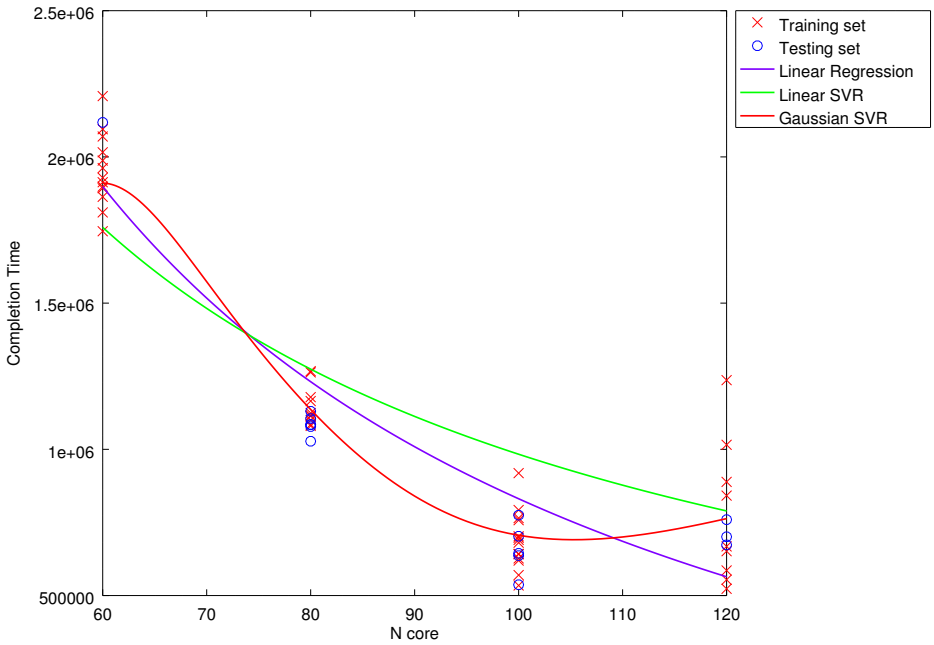

Figure 68: Completion time vs ncores for query R2 with datasize 1000GB

4.3 Query R3

4.3.1 Query R3 — Datasize 250GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.7627	0.4444	40061	0.1892
Linear SVR	0.7863	0.7014	40887	0.1865
Polynomial SVR (2)	1.0935	0.1924	61331	0.3228
Polynomial SVR (3)	0.8169	0.6371	43254	0.2055
Polynomial SVR (4)	0.9730	0.3470	54074	0.2851
Polynomial SVR (6)	1.0190	0.4264	57989	0.2907
Gaussian SVR	0.7857	0.6768	42184	0.2023

Table 69: Results for R3-250GB with the nonlinear 1/ncores feature, only ncores

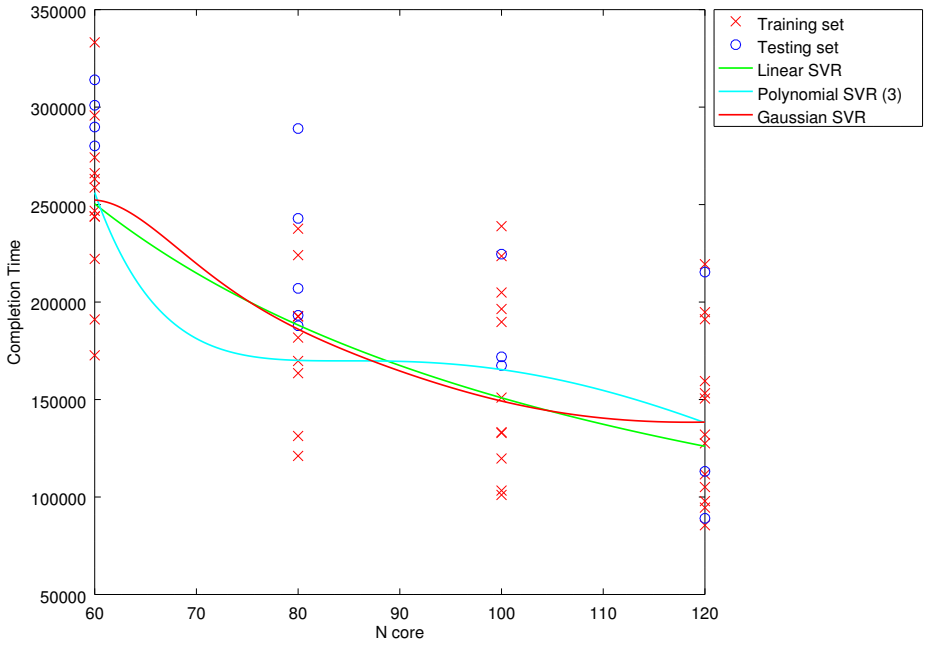

Figure 69: Completion time vs ncores for query R3 with datasize 250GB

4.3.2 Query R3 — Datasize 500GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.3162	0.8991	74328	0.1622
Linear SVR	0.3102	0.9030	71562	0.1475
Polynomial SVR (2)	0.8017	0.3717	170794	0.4134
Polynomial SVR (3)	0.1738	0.9698	33539	0.0823
Polynomial SVR (4)	0.6309	0.6249	122119	0.3179
Polynomial SVR (6)	0.5001	0.7625	102471	0.2574
Gaussian SVR	0.1390	0.9814	27308	0.0611

Table 70: Results for R3-500GB with the nonlinear 1/ncores feature, only ncores

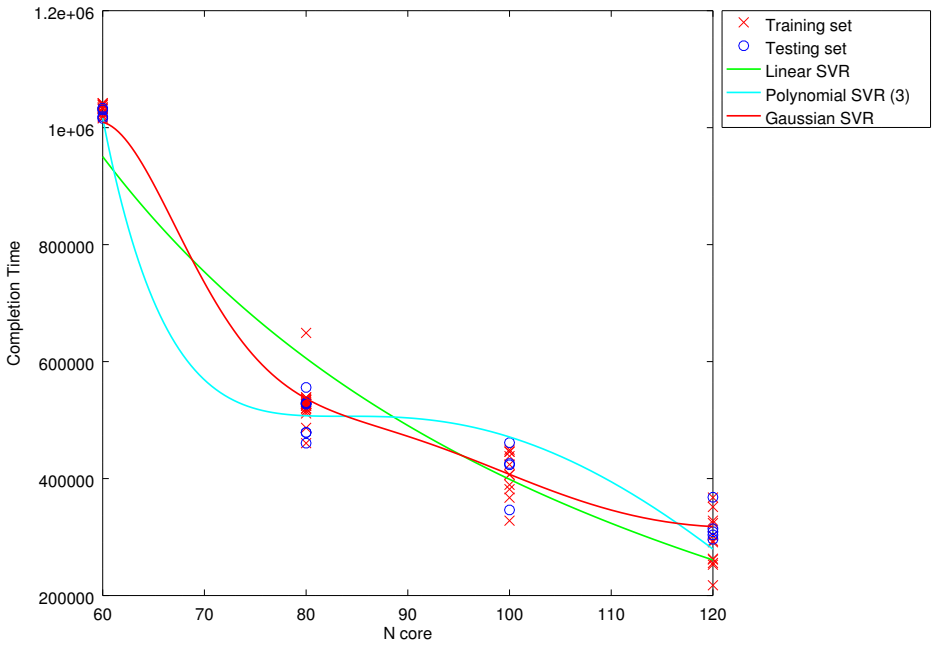

Figure 70: Completion time vs ncores for query R3 with datasize 500GB

4.3.3 Query R3 — Datasize 750GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.2049	0.9524	28016	0.0404
Linear SVR	0.2057	0.9575	28111	0.0401
Polynomial SVR (2)	0.8224	0.2464	128039	0.1753
Polynomial SVR (3)	0.3301	0.8941	46655	0.0635
Polynomial SVR (4)	0.6989	0.5009	93688	0.1338
Polynomial SVR (6)	0.5836	0.6402	78143	0.1099
Gaussian SVR	0.1629	0.9849	18453	0.0281

Table 71: Results for R3-750GB with the nonlinear 1/ncores feature, only ncores

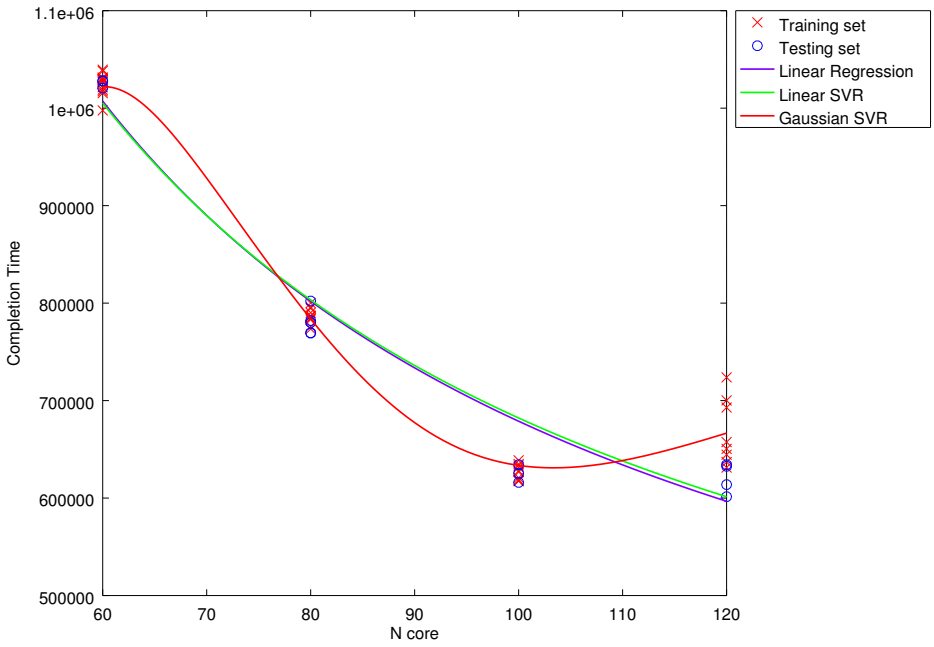

Figure 71: Completion time vs ncores for query R3 with datasize 750GB

4.3.4 Query R3 — Datasize 1000GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.2395	0.9391	47342	0.0519
Linear SVR	0.2429	0.9424	47915	0.0527
Polynomial SVR (2)	0.7601	0.4483	150763	0.1636
Polynomial SVR (3)	0.3880	0.8856	63063	0.0725
Polynomial SVR (4)	0.5732	0.6936	111389	0.1218
Polynomial SVR (6)	0.4757	0.7969	94384	0.1027
Gaussian SVR	0.0893	0.9941	15506	0.0164

Table 72: Results for R3-1000GB with the nonlinear 1/ncores feature, only ncores

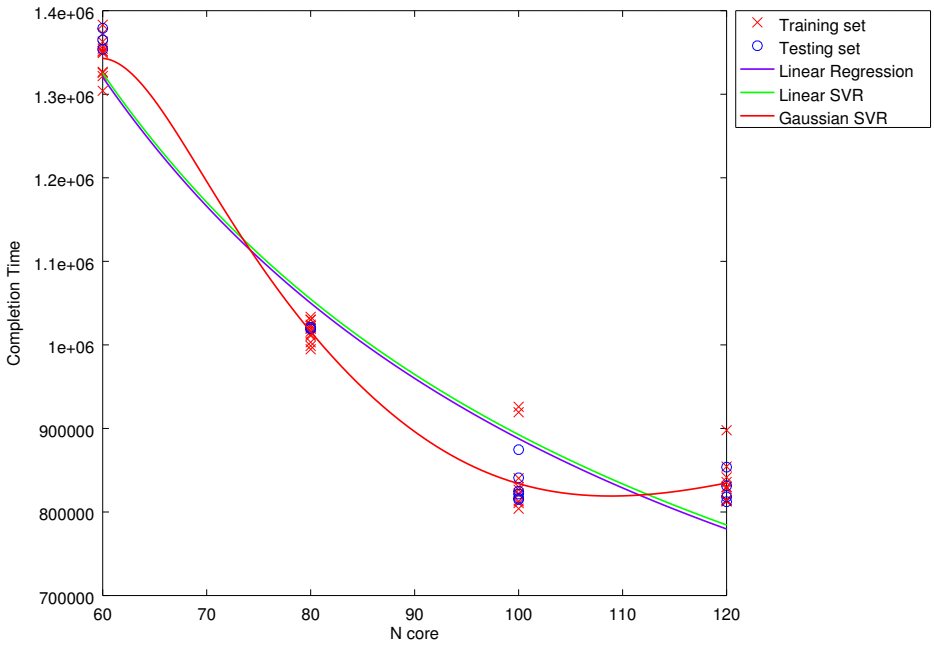

Figure 72: Completion time vs ncores for query R3 with datasize 1000GB

4.4 Query R4

4.4.1 Query R4 — Datasize 250GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.8686	0.1477	36477	0.2638
Linear SVR	0.8530	0.2189	36318	0.2718
Polynomial SVR (2)	1.0554	0.0583	42311	0.3487
Polynomial SVR (3)	0.9447	0.0901	37389	0.2771
Polynomial SVR (4)	1.0542	0.0017	43480	0.3472
Polynomial SVR (6)	1.0304	0.0023	42378	0.3295
Gaussian SVR	0.8335	0.2617	35373	0.2615

Table 73: Results for R4-250GB with the nonlinear 1/ncores feature, only ncores

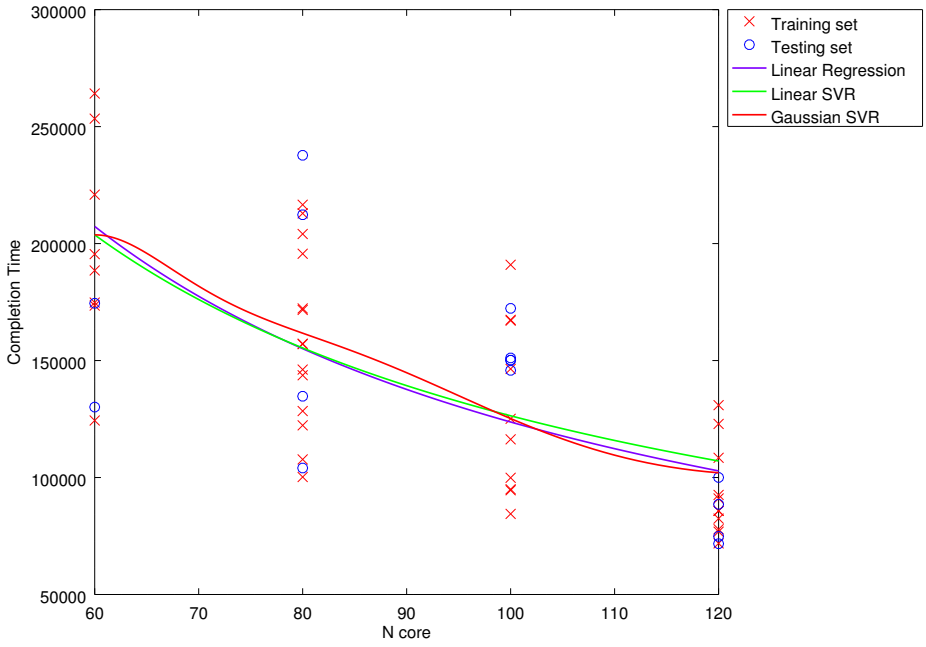

Figure 73: Completion time vs ncores for query R4 with datasize 250GB

4.4.2 Query R4 — Datasize 500GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.3338	0.9217	56936	0.1305
Linear SVR	0.3256	0.9357	55121	0.1264
Polynomial SVR (2)	0.9065	0.4895	144214	0.3545
Polynomial SVR (3)	0.2259	0.9863	34171	0.0723
Polynomial SVR (4)	0.6572	0.7210	94909	0.2849
Polynomial SVR (6)	0.5188	0.8340	78746	0.2291
Gaussian SVR	0.0987	0.9941	15135	0.0407

Table 74: Results for R4-500GB with the nonlinear 1/ncores feature, only ncores

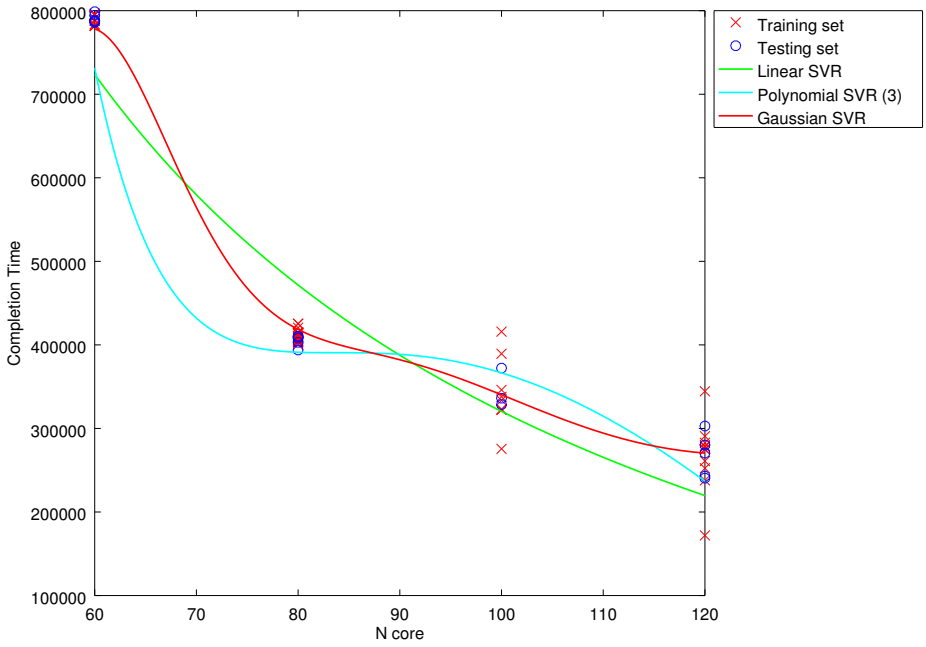

Figure 74: Completion time vs ncores for query R4 with datasize 500GB

4.4.3 Query R4 — Datasize 750GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.2516	0.8972	26818	0.0505
Linear SVR	0.2485	0.9195	27232	0.0506
Polynomial SVR (2)	0.6507	0.3168	79129	0.1386
Polynomial SVR (3)	0.4005	0.7802	41995	0.0791
Polynomial SVR (4)	0.5367	0.5595	58504	0.1007
Polynomial SVR (6)	0.4790	0.6565	54216	0.0925
Gaussian SVR	0.1194	0.9780	12187	0.0212

Table 75: Results for R4-750GB with the nonlinear 1/ncores feature, only ncores

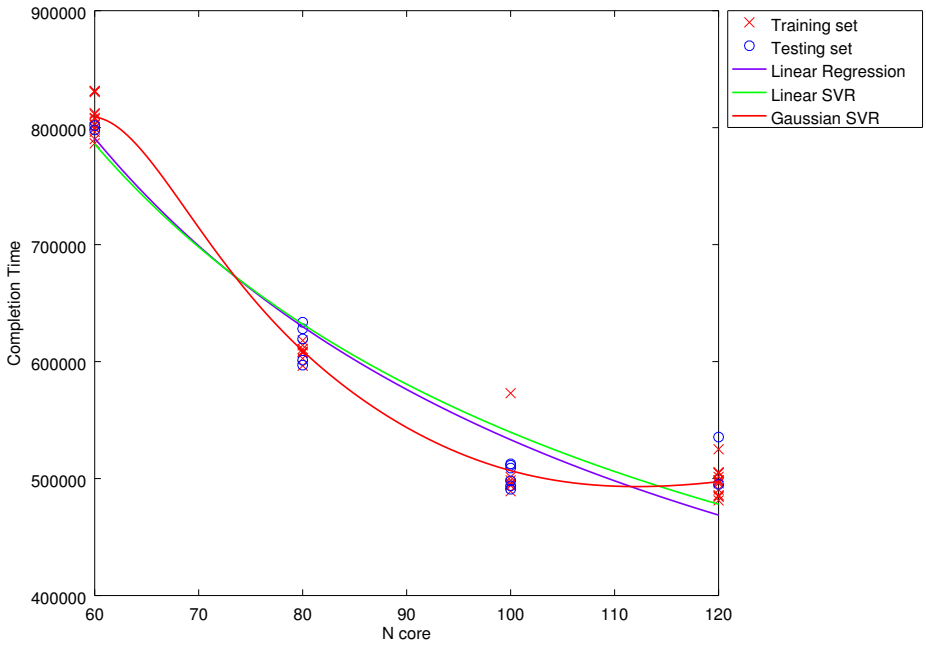

Figure 75: Completion time vs ncores for query R4 with datasize 750GB

4.4.4 Query R4 — Datasize 1000GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.6309	0.3868	402429	0.3275
Linear SVR	0.6552	0.4433	491316	0.3635
Polynomial SVR (2)	0.8742	0.0220	656146	0.4904
Polynomial SVR (3)	0.7401	0.2208	516331	0.3896
Polynomial SVR (4)	0.9535	0.0796	721224	0.5477
Polynomial SVR (6)	0.9594	0.1062	724319	0.5523
Gaussian SVR	0.2640	0.8931	147333	0.0879

Table 76: Results for R4-1000GB with the nonlinear 1/ncores feature, only ncores

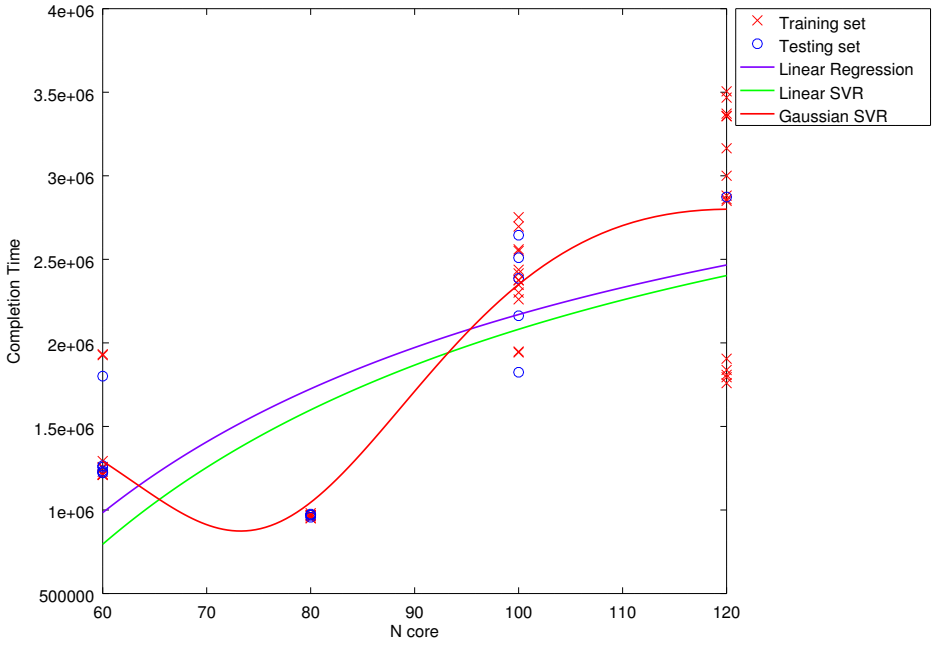

Figure 76: Completion time vs ncores for query R4 with datasize 1000GB

4.5 Query R5

4.5.1 Query R5 — Datasize 250GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.8453	-0.3635	534	0.0213
Linear SVR	0.7863	0.0052	514	0.0204
Polynomial SVR (2)	0.7718	0.0620	498	0.0198
Polynomial SVR (3)	0.7844	0.0178	515	0.0204
Polynomial SVR (4)	0.8009	0.0468	502	0.0199
Polynomial SVR (6)	0.8011	0.0419	497	0.0197
Gaussian SVR	0.8189	0.0012	486	0.0192

Table 77: Results for R5-250GB with the nonlinear 1/ncores feature, only ncores

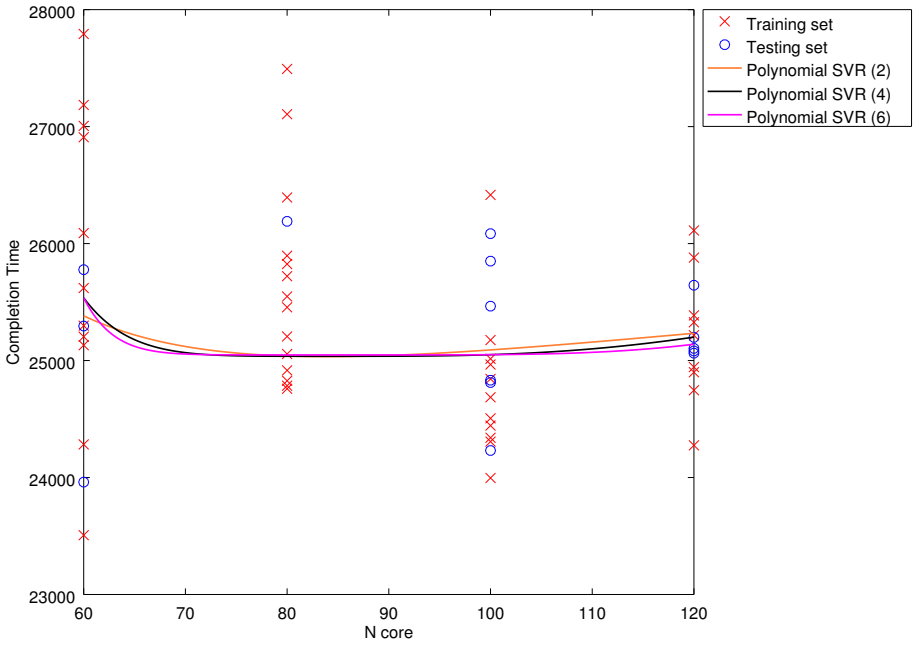

Figure 77: Completion time vs ncores for query R5 with datasize 250GB

4.5.2 Query R5 — Datasize 500GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	1.1189	-0.0289	1200	0.0483
Linear SVR	1.1011	0.1510	1214	0.0493
Polynomial SVR (2)	1.1335	0.0115	1295	0.0528
Polynomial SVR (3)	1.0972	0.1235	1221	0.0497
Polynomial SVR (4)	1.1256	0.0000	1275	0.0519
Polynomial SVR (6)	1.1101	0.0053	1273	0.0520
Gaussian SVR	1.1255	0.0005	1276	0.0520

Table 78: Results for R5-500GB with the nonlinear 1/ncores feature, only ncores

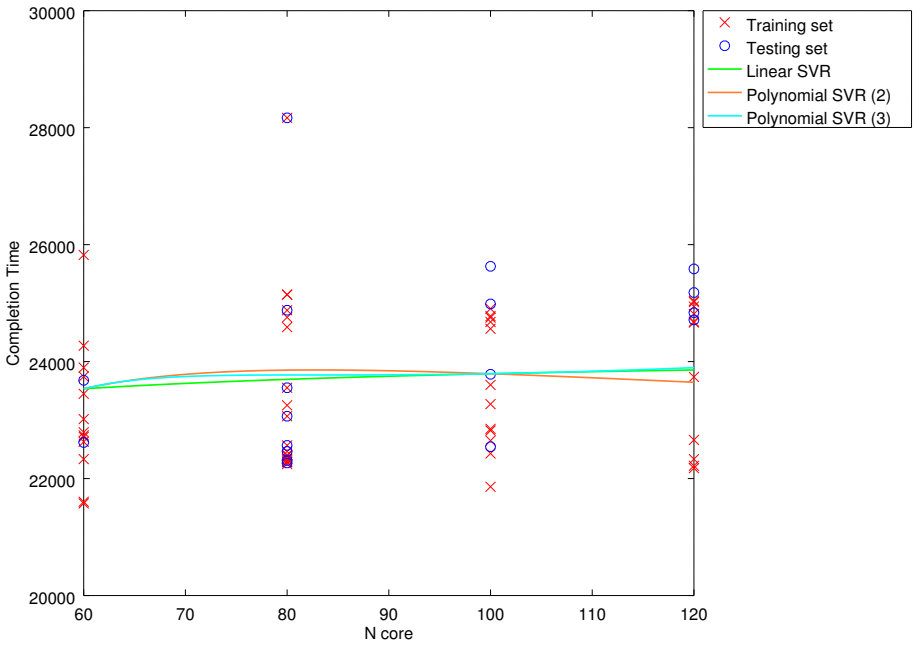

Figure 78: Completion time vs ncores for query R5 with datasize 500GB

4.5.3 Query R5 — Datasize 750GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.8980	-0.0354	1068	0.0445
Linear SVR	1.0199	-0.0000	983	0.0425
Polynomial SVR (2)	1.0199	-0.0000	983	0.0425
Polynomial SVR (3)	1.0199	-0.0000	983	0.0425
Polynomial SVR (4)	1.0199	-0.0000	983	0.0425
Polynomial SVR (6)	1.0199	-0.0000	983	0.0425
Gaussian SVR	1.0199	-0.0000	983	0.0425

Table 79: Results for R5-750GB with the nonlinear 1/ncores feature, only ncores

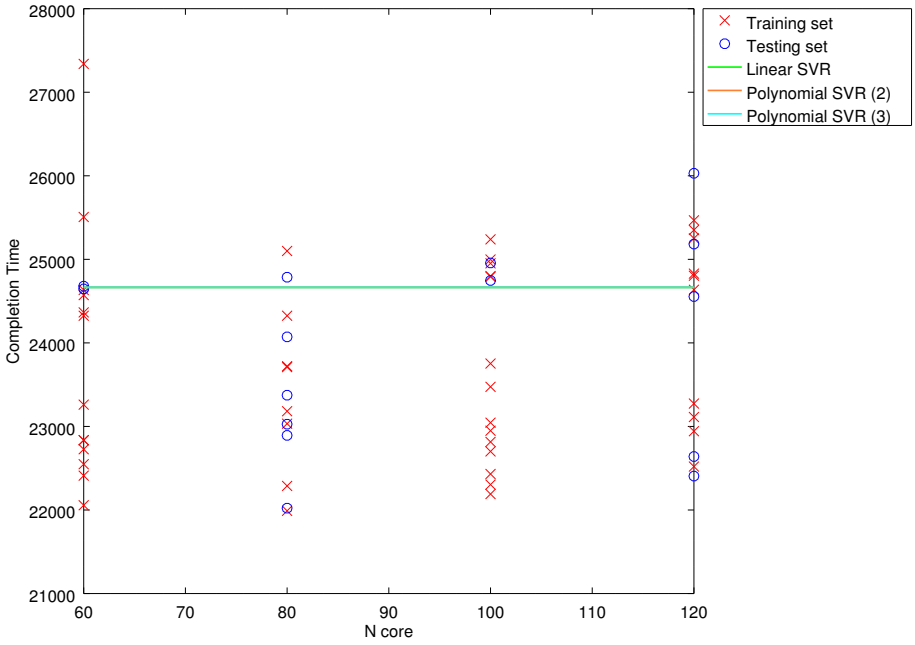

Figure 79: Completion time vs ncores for query R5 with datasize 750GB

4.5.4 Query R5 — Datasize 1000GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.9881	0.2649	3615	0.0851
Linear SVR	1.0305	0.2927	3493	0.0805
Polynomial SVR (2)	1.0317	0.2871	3100	0.0694
Polynomial SVR (3)	1.0104	0.3333	3346	0.0763
Polynomial SVR (4)	1.0033	0.3402	2998	0.0667
Polynomial SVR (6)	1.0049	0.3490	3055	0.0682
Gaussian SVR	0.9937	0.3578	3236	0.0737

Table 80: Results for R5-1000GB with the nonlinear 1/ncores feature, only ncores

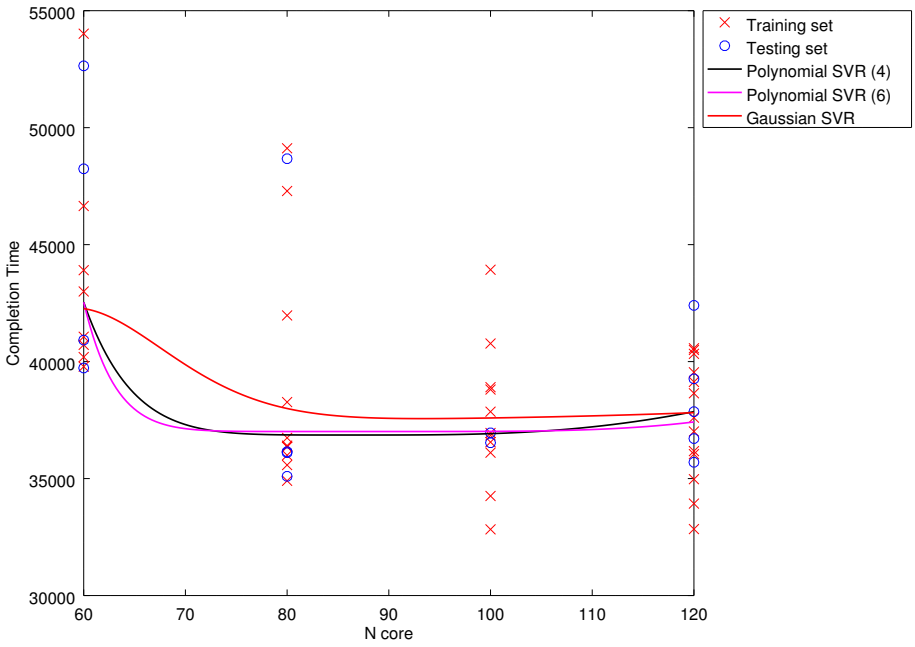

Figure 80: Completion time vs ncores for query R5 with datasize 1000GB

5 Query comparison

5.1 R5 → R2 (Fast - Fast)

5.1.1 R5 → R2 — Datasize 250GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.2664	-3.8012	7131	0.0868
Linear SVR	1.4192	0.4500	41150	0.4928
Polynomial SVR (2)	2.0017	0.3318	57963	0.6936
Polynomial SVR (3)	1.7351	0.3201	50336	0.6032
Polynomial SVR (4)	2.0052	0.2591	58051	0.6946
Polynomial SVR (6)	2.0177	0.3148	58314	0.6974
Gaussian SVR	1.9531	0.0953	56646	0.6783

Table 81: Results for R5 → R2-250

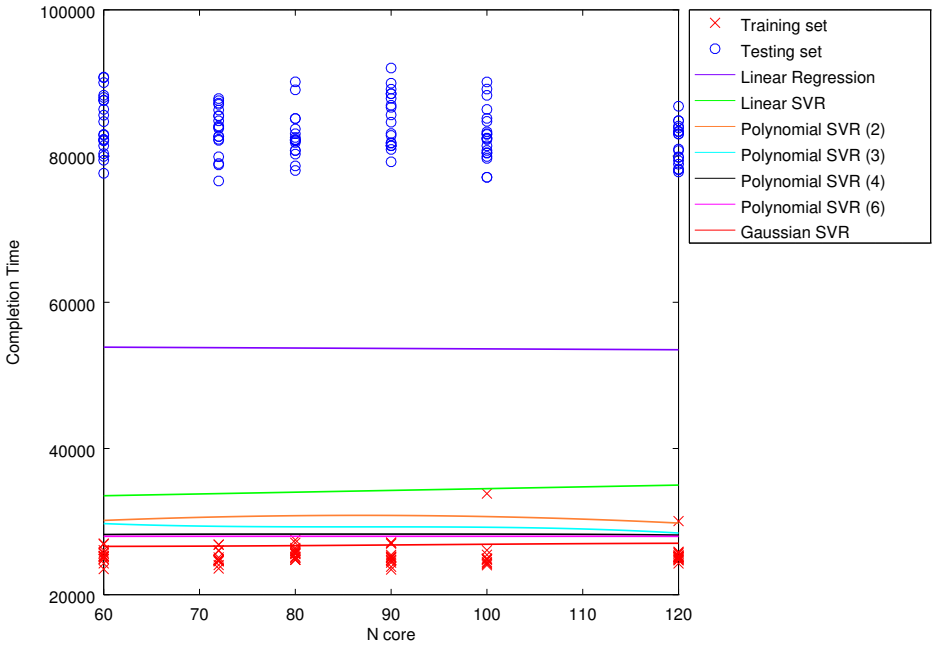

Figure 81: Completion time vs ncores for: R5 → R2 with datasize 250GB

5.1.2 R5 → R2 — Datasize 500GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.0927	0.4825	1627	0.0219
Linear SVR	0.8177	0.8804	20156	0.2760
Polynomial SVR (2)	1.9565	0.8888	47958	0.6544
Polynomial SVR (3)	1.3110	0.8506	32346	0.4437
Polynomial SVR (4)	1.9813	0.8130	48399	0.6598
Polynomial SVR (6)	1.9723	0.7576	48394	0.6606
Gaussian SVR	1.8458	0.4395	45448	0.6213

Table 82: Results for R5 → R2-500

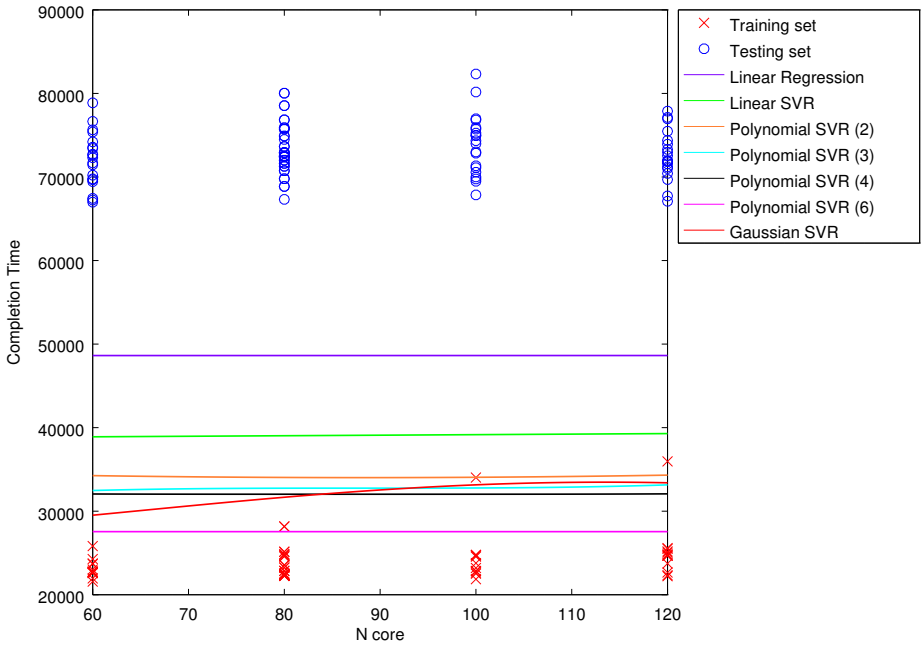

Figure 82: Completion time vs ncores for: R5 → R2 with datasize 500GB

5.1.3 R5 → R2 — Datasize 750GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.7284	-36.9573	19886	0.2523
Linear SVR	1.2200	0.7629	33371	0.4226
Polynomial SVR (2)	1.9610	0.7081	53638	0.6792
Polynomial SVR (3)	1.7155	0.7155	46976	0.5953
Polynomial SVR (4)	1.9544	0.6962	53405	0.6760
Polynomial SVR (6)	1.9545	0.6360	53347	0.6749
Gaussian SVR	1.8982	0.4132	51944	0.6579

Table 83: Results for R5 → R2-750

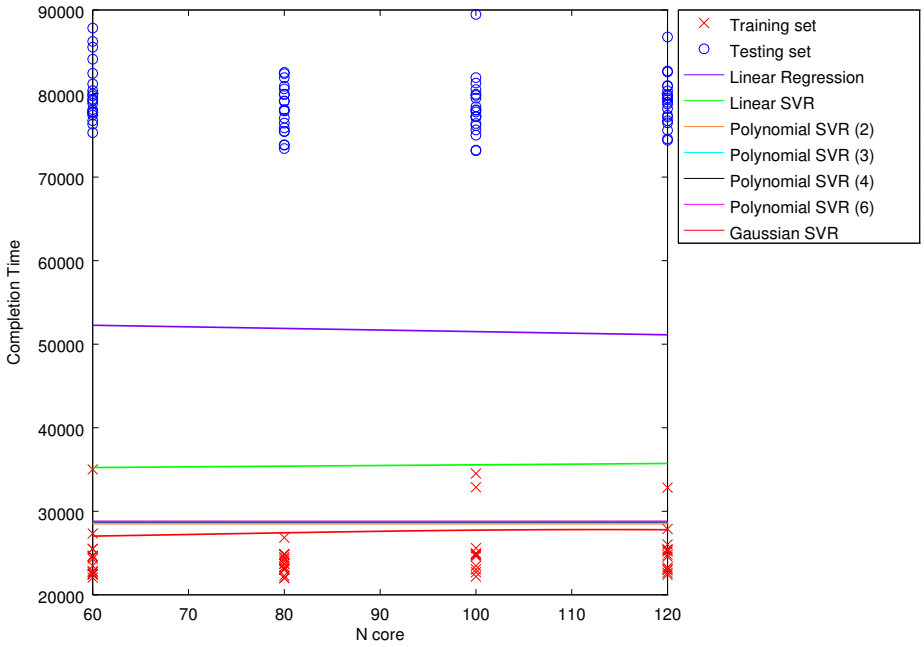

Figure 83: Completion time vs ncores for: R5 → R2 with datasize 750GB

5.1.4 R5 → R2 — Datasize 1000GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.2791	0.8182	124323	0.1468
Linear SVR	1.9131	-0.0000	865805	0.9413
Polynomial SVR (2)	1.9131	-0.0000	865805	0.9413
Polynomial SVR (3)	1.9131	-0.0000	865805	0.9413
Polynomial SVR (4)	1.9131	-0.0000	865805	0.9413
Polynomial SVR (6)	1.9131	-0.0000	865805	0.9413
Gaussian SVR	1.9131	-0.0000	865805	0.9413

Table 84: Results for R5 → R2-1000

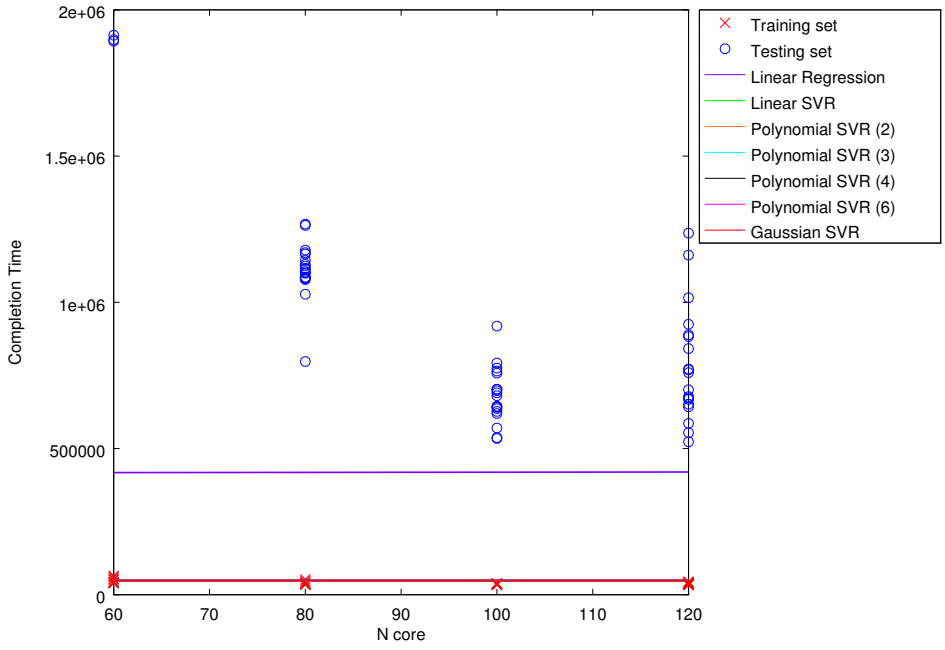

Figure 84: Completion time vs ncores for: R5 → R2 with datasize 1000GB

5.2 R2 → R5 (Fast - Fast)

5.2.1 R2 → R5 — Datasize 250GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.5076	-54.2010	14618	0.5690
Linear SVR	0.0742	0.5354	1906	0.0748
Polynomial SVR (2)	1.9582	0.5005	56793	2.2284
Polynomial SVR (3)	1.4334	0.4575	41627	1.6299
Polynomial SVR (4)	1.9578	0.4334	56818	2.2282
Polynomial SVR (6)	1.9529	0.3733	56692	2.2226
Gaussian SVR	1.8646	0.4311	54134	2.1221

Table 85: Results for R2 → R5-250

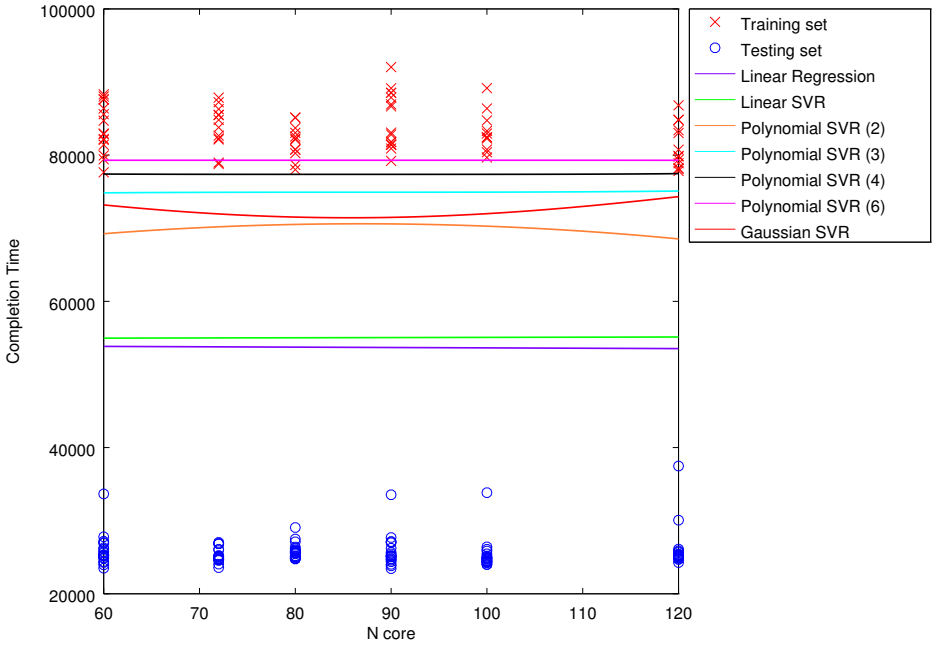

Figure 85: Completion time vs ncores for: R2 → R5 with datasize 250GB

5.2.2 R2 → R5 — Datasize 500GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.2910	-3.7899	6672	0.2718
Linear SVR	0.7805	0.3478	19144	0.8083
Polynomial SVR (2)	1.9764	0.3328	48648	2.0492
Polynomial SVR (3)	1.5362	0.3112	37827	1.5916
Polynomial SVR (4)	1.9546	0.2841	48117	2.0265
Polynomial SVR (6)	1.9390	0.2654	47733	2.0103
Gaussian SVR	1.9761	0.1983	48681	2.0483

Table 86: Results for R2 → R5-500

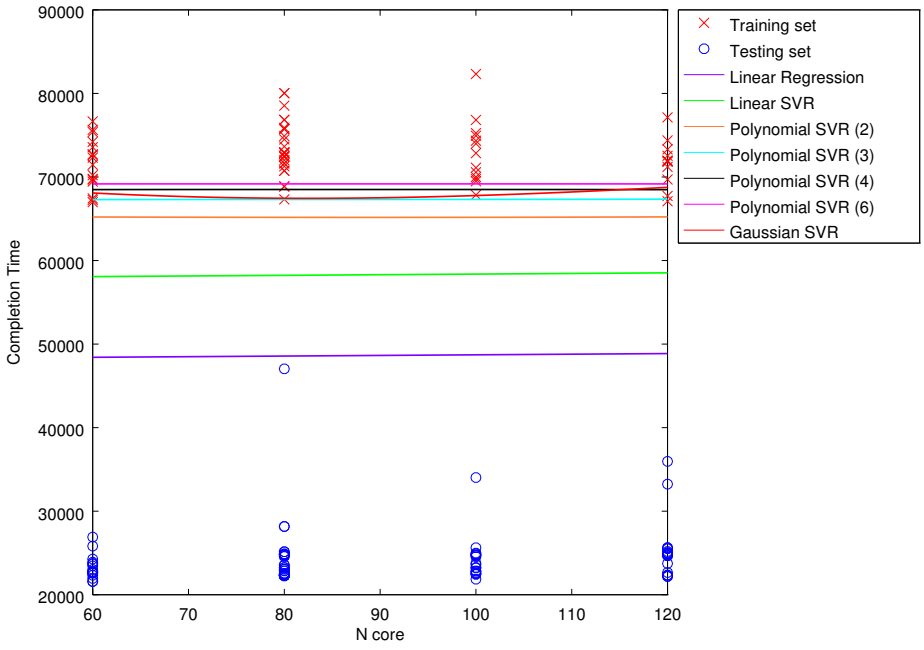

Figure 86: Completion time vs ncores for: R2 → R5 with datasize 500GB

5.2.3 R2 → R5 — Datasize 750GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.7050	-58.7255	19273	0.7906
Linear SVR	0.5814	0.1630	15781	0.6549
Polynomial SVR (2)	1.9918	0.3443	54498	2.2536
Polynomial SVR (3)	1.5014	0.3534	41121	1.6974
Polynomial SVR (4)	1.9754	0.2597	54059	2.2349
Polynomial SVR (6)	1.9710	0.2735	53943	2.2300
Gaussian SVR	2.0285	0.4331	55559	2.2948

Table 87: Results for R2 → R5-750

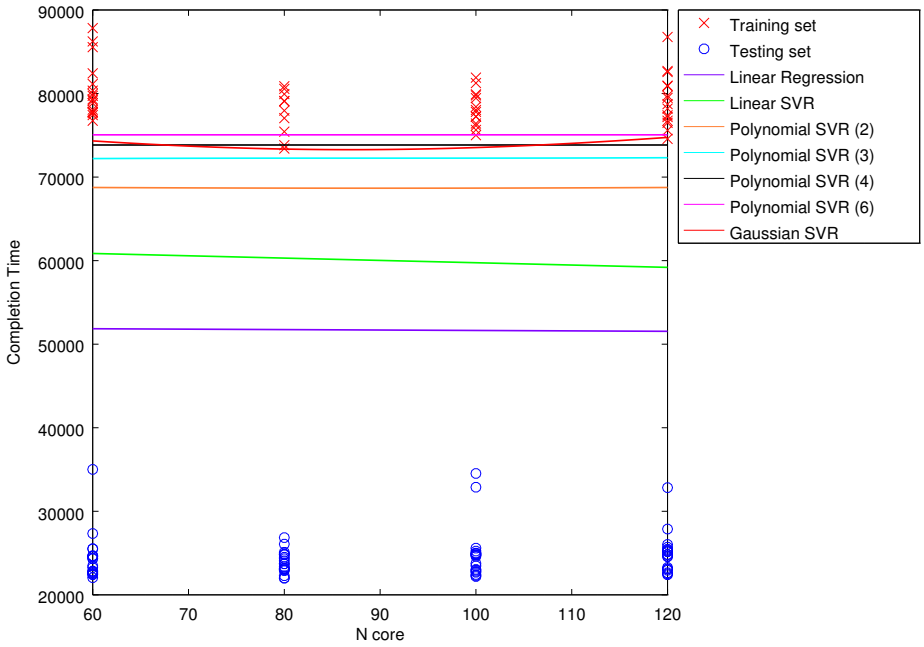

Figure 87: Completion time vs ncores for: R2 → R5 with datasize 750GB

5.2.4 R2 → R5 — Datasize 1000GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.2457	-258.0712	117597	2.9684
Linear SVR	0.0746	0.0791	31327	0.7786
Polynomial SVR (2)	1.5254	0.2338	734175	18.4731
Polynomial SVR (3)	0.8871	0.1001	426009	10.6683
Polynomial SVR (4)	1.4660	0.0268	705234	17.7261
Polynomial SVR (6)	1.4552	0.0225	700727	17.5862
Gaussian SVR	2.1204	0.0853	1020795	25.5869

Table 88: Results for R2 → R5-1000

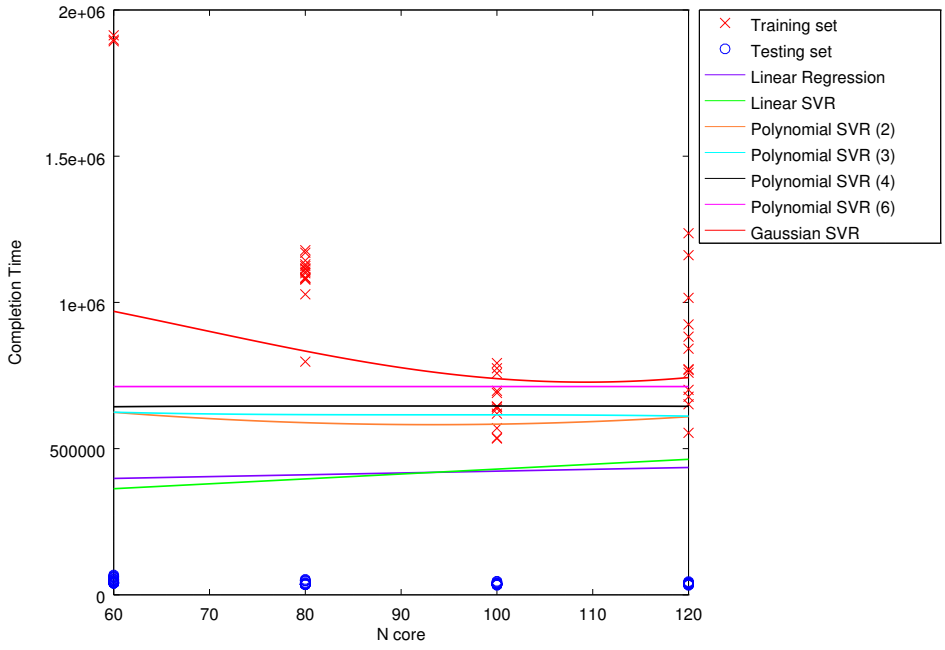

Figure 88: Completion time vs ncores for: R2 → R5 with datasize 1000GB

5.3 R3 → R4 (Slow - Slow)

5.3.1 R3 → R4 — Datasize 250GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.0999	0.9905	7743	0.0556
Linear SVR	0.1017	0.9936	7894	0.0559
Polynomial SVR (2)	1.2864	0.0000	109952	0.8321
Polynomial SVR (3)	0.5276	0.8127	40426	0.2156
Polynomial SVR (4)	0.9298	0.2785	77573	0.5374
Polynomial SVR (6)	0.8746	0.2982	69520	0.4722
Gaussian SVR	0.5125	0.8738	39988	0.2815

Table 89: Results for R3 → R4-250

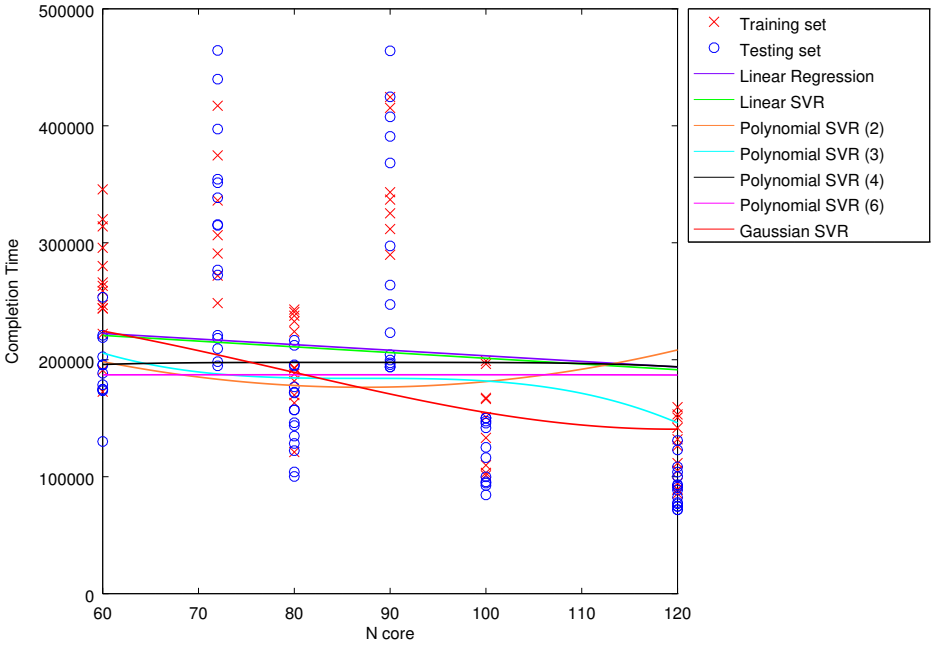

Figure 89: Completion time vs ncores for: R3 → R4 with datasize 250GB

5.3.2 R3 → R4 — Datasize 500GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.5575	0.7917	79305	0.1806
Linear SVR	0.6248	0.8895	87857	0.1990
Polynomial SVR (2)	1.4064	0.3856	175966	0.4039
Polynomial SVR (3)	0.9288	0.7713	129311	0.3215
Polynomial SVR (4)	1.3185	0.1435	177714	0.4267
Polynomial SVR (6)	1.4713	0.1476	177441	0.4040
Gaussian SVR	1.0797	0.3126	134221	0.2887

Table 90: Results for R3 → R4-500

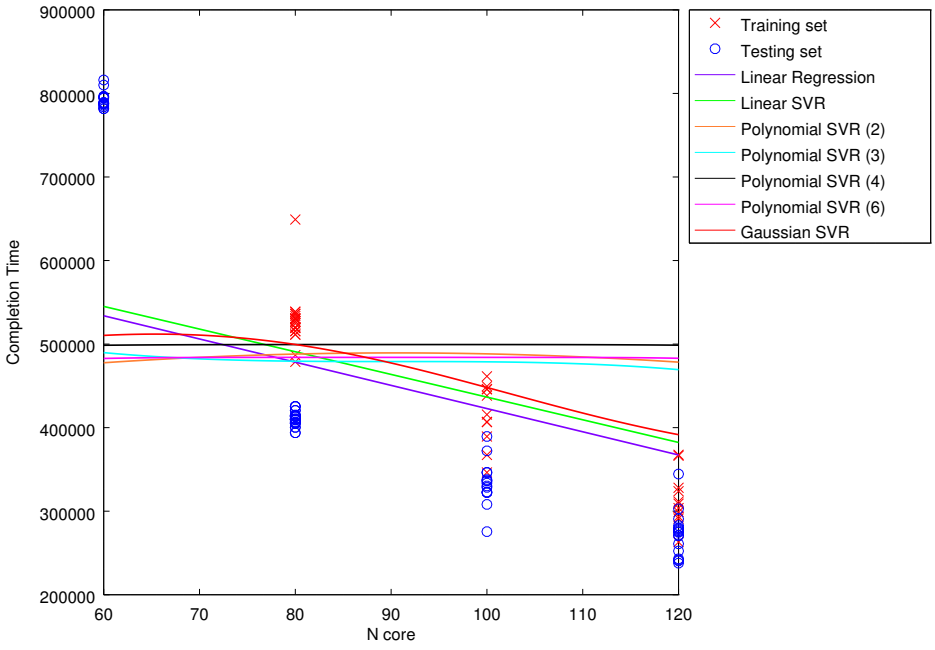

Figure 90: Completion time vs ncores for: R3 → R4 with datasize 500GB

5.3.3 R3 → R4 — Datasize 750GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.0933	0.9899	12026	0.0195
Linear SVR	0.2629	0.9912	33564	0.0561
Polynomial SVR (2)	1.8337	0.7689	217984	0.3828
Polynomial SVR (3)	0.7352	0.7039	87799	0.1548
Polynomial SVR (4)	1.8539	0.4641	210264	0.3621
Polynomial SVR (6)	1.1873	0.4401	148303	0.2645
Gaussian SVR	1.0774	0.8421	126509	0.2349

Table 91: Results for R3 → R4-750

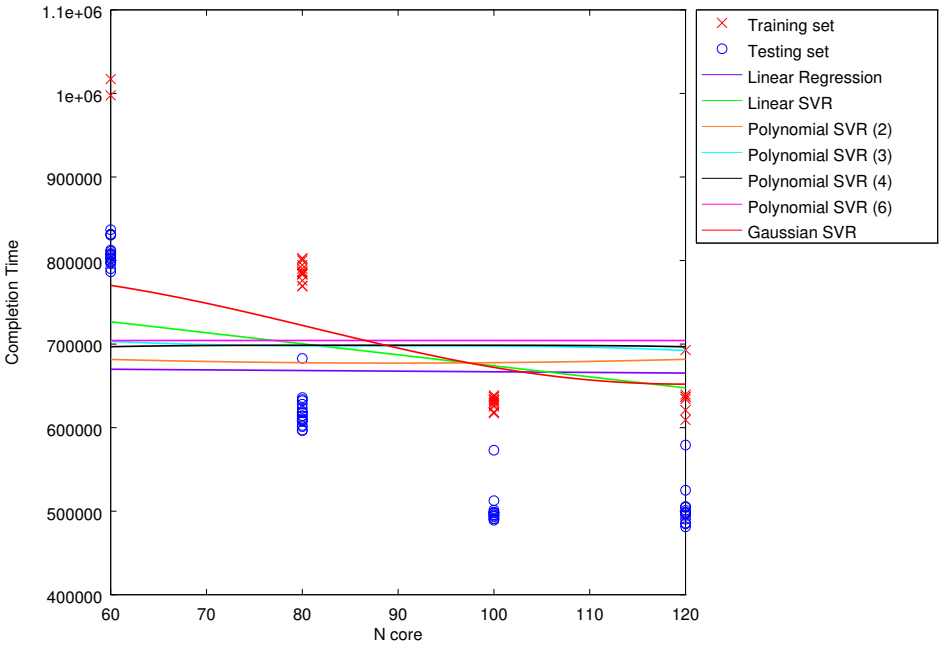

Figure 91: Completion time vs ncores for: R3 → R4 with datasize 750GB

5.3.4 R3 → R4 — Datasize 1000GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	1.7878	-1.2231	784437	0.3797
Linear SVR	1.6361	0.0281	660983	0.2902
Polynomial SVR (2)	1.6549	0.0157	689481	0.3172
Polynomial SVR (3)	1.6757	0.0650	705047	0.3218
Polynomial SVR (4)	1.9237	0.1209	764929	0.3431
Polynomial SVR (6)	1.8627	0.0254	735220	0.3294
Gaussian SVR	1.5644	0.0772	640303	0.2875

Table 92: Results for R3 → R4-1000

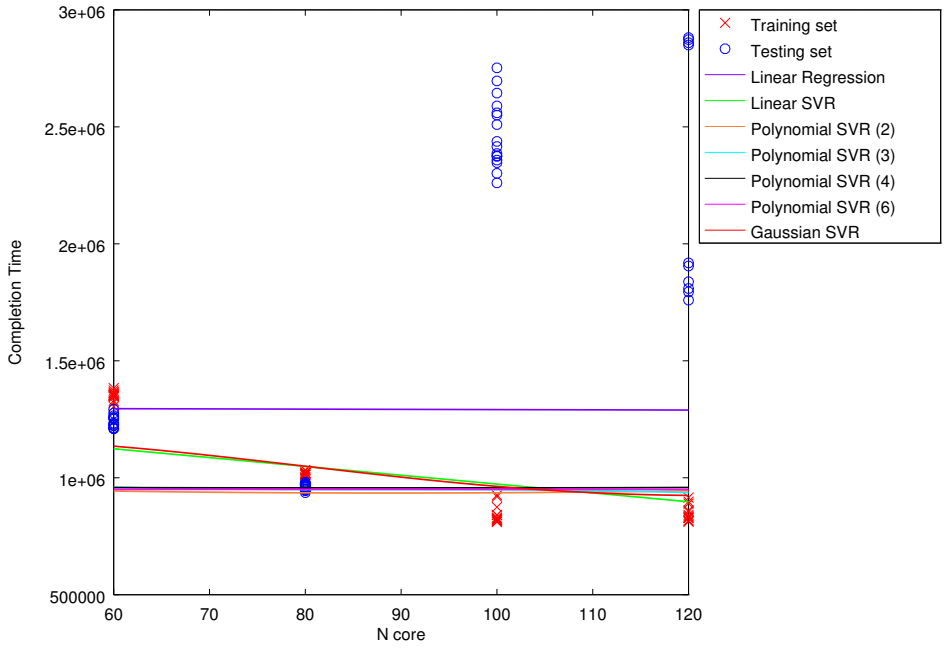

Figure 92: Completion time vs ncores for: R3 → R4 with datasize 1000GB

5.4 R4 → R3 (Slow - Slow)

5.4.1 R4 → R3 — Datasize 250GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.1117	0.9859	6402	0.0272
Linear SVR	0.1241	0.9882	8184	0.0379
Polynomial SVR (2)	1.0187	0.0102	84602	0.4395
Polynomial SVR (3)	0.7609	0.7115	53158	0.2639
Polynomial SVR (4)	1.0989	0.0178	85066	0.3771
Polynomial SVR (6)	1.0632	0.0244	82125	0.3725
Gaussian SVR	0.5420	0.8101	41963	0.2180

Table 93: Results for R4 → R3-250

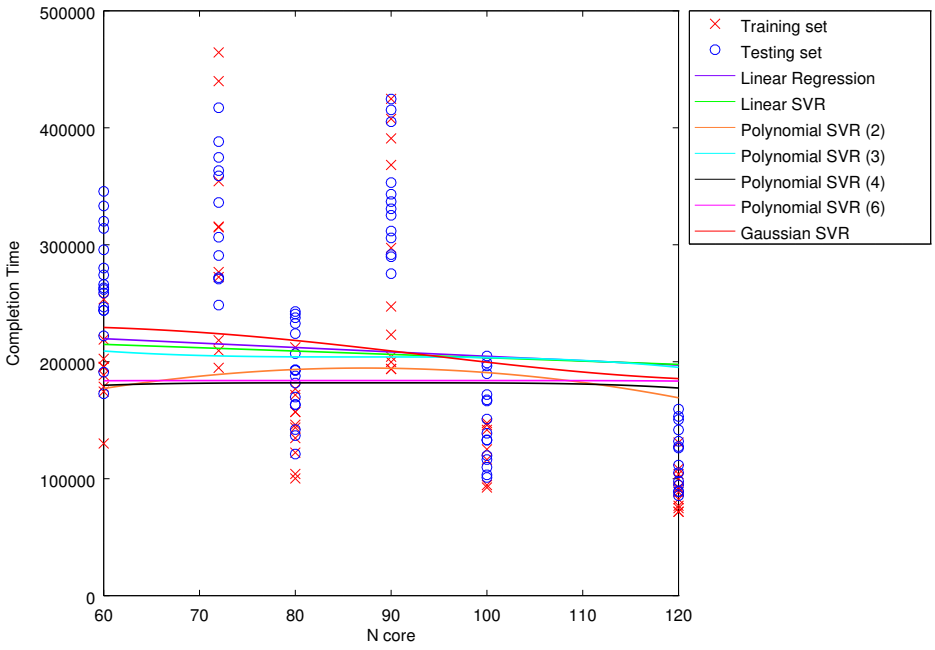

Figure 93: Completion time vs ncores for: R4 → R3 with datasize 250GB

5.4.2 R4 → R3 — Datasize 500GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.1181	0.9592	13349	0.0297
Linear SVR	0.1281	0.9719	13961	0.0304
Polynomial SVR (2)	1.0739	0.0989	142216	0.3246
Polynomial SVR (3)	0.5917	0.7504	77226	0.1606
Polynomial SVR (4)	1.0211	0.2288	134494	0.3151
Polynomial SVR (6)	0.9462	0.0853	117567	0.2718
Gaussian SVR	0.3507	0.7951	46118	0.1032

Table 94: Results for R4 → R3-500

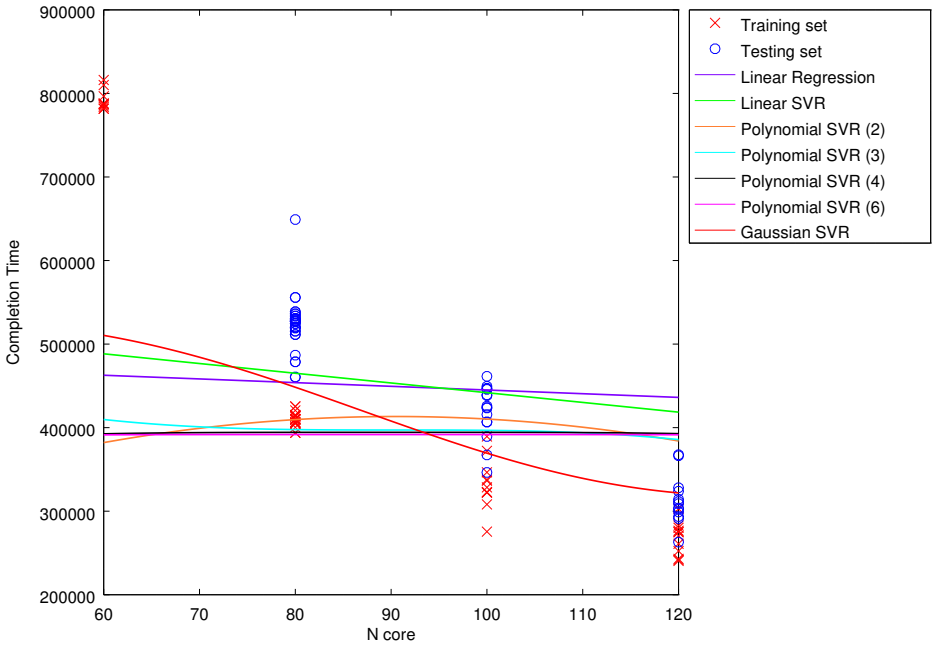

Figure 94: Completion time vs ncores for: R4 → R3 with datasize 500GB

5.4.3 R4 → R3 — Datasize 750GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.1065	0.9853	13983	0.0192
Linear SVR	0.1572	0.9937	18494	0.0250
Polynomial SVR (2)	1.7890	0.8077	185642	0.2313
Polynomial SVR (3)	1.0593	0.8125	126124	0.1704
Polynomial SVR (4)	2.8642	0.6965	233765	0.2775
Polynomial SVR (6)	2.9392	0.5268	213611	0.2468
Gaussian SVR	0.8538	0.7386	85514	0.1055

Table 95: Results for R4 → R3-750

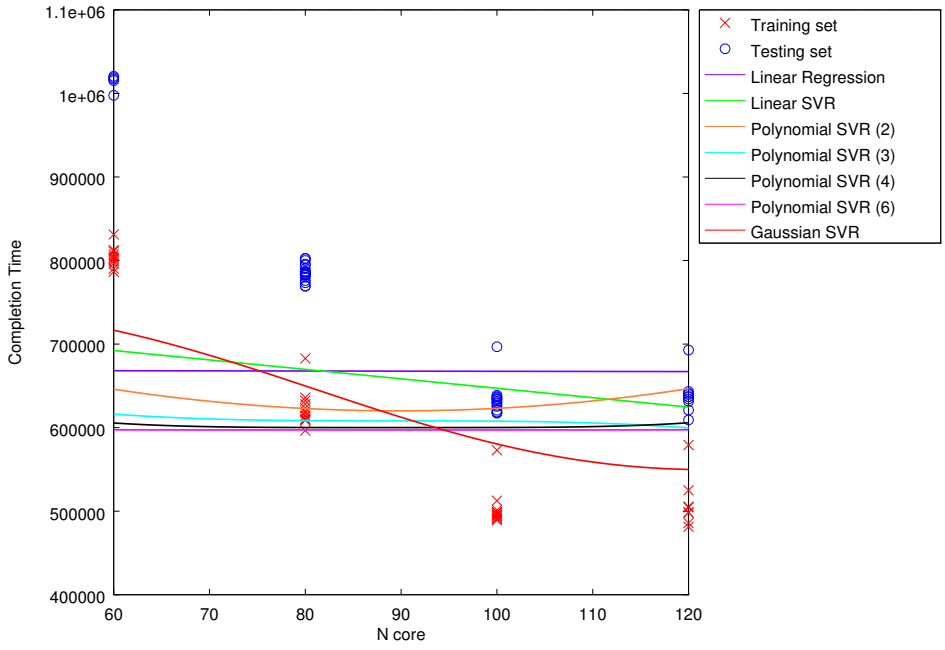

Figure 95: Completion time vs ncores for: R4 → R3 with datasize 750GB

5.4.4 R4 → R3 — Datasize 1000GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.2979	0.2999	163252	0.1736
Linear SVR	0.2150	0.9597	109503	0.1219
Polynomial SVR (2)	0.7033	0.6852	382105	0.4195
Polynomial SVR (3)	0.5805	0.7986	324151	0.3440
Polynomial SVR (4)	0.3638	0.7740	189564	0.2076
Polynomial SVR (6)	0.4766	0.6324	252222	0.2741
Gaussian SVR	1.0142	0.5173	574326	0.6115

Table 96: Results for R4 → R3-1000

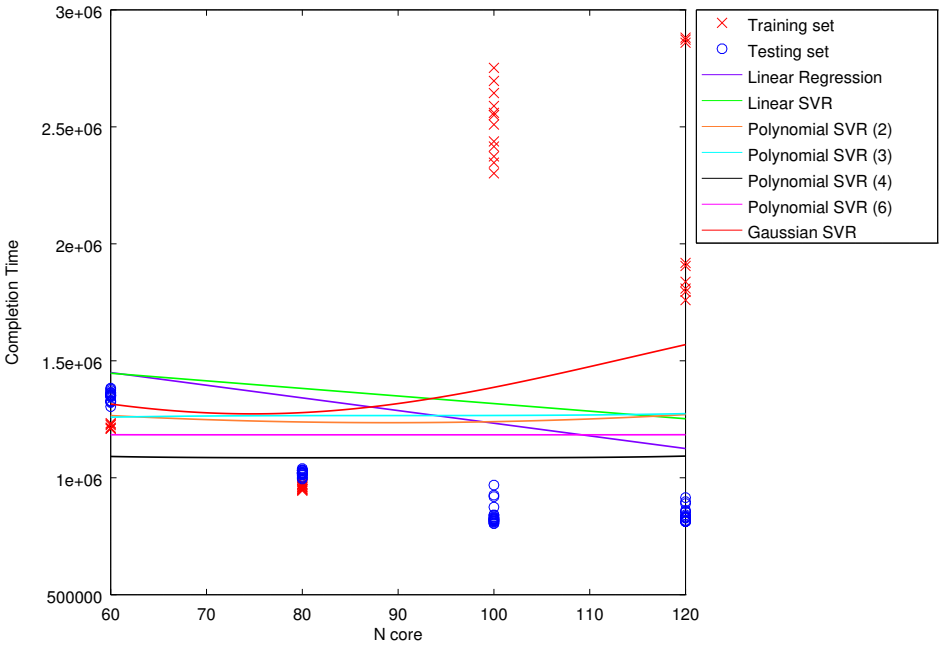

Figure 96: Completion time vs ncores for: R4 → R3 with datasize 1000GB

5.5 R1, R2, R4 → R3 (Fast and Slow - Average)

5.5.1 R1,R2,R4 → R3 — Datasize 250GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.1484	0.9772	9501	0.0466
Linear SVR	0.1560	0.9845	10574	0.0553
Polynomial SVR (2)	0.8548	0.7060	57023	0.2555
Polynomial SVR (3)	0.4469	0.8616	28093	0.1225
Polynomial SVR (4)	0.7404	0.7503	42222	0.1783
Polynomial SVR (6)	1.1648	0.5504	61124	0.2455
Gaussian SVR	0.3926	0.9405	24779	0.1084

Table 97: Results for R1,R2,R4 → R3-250

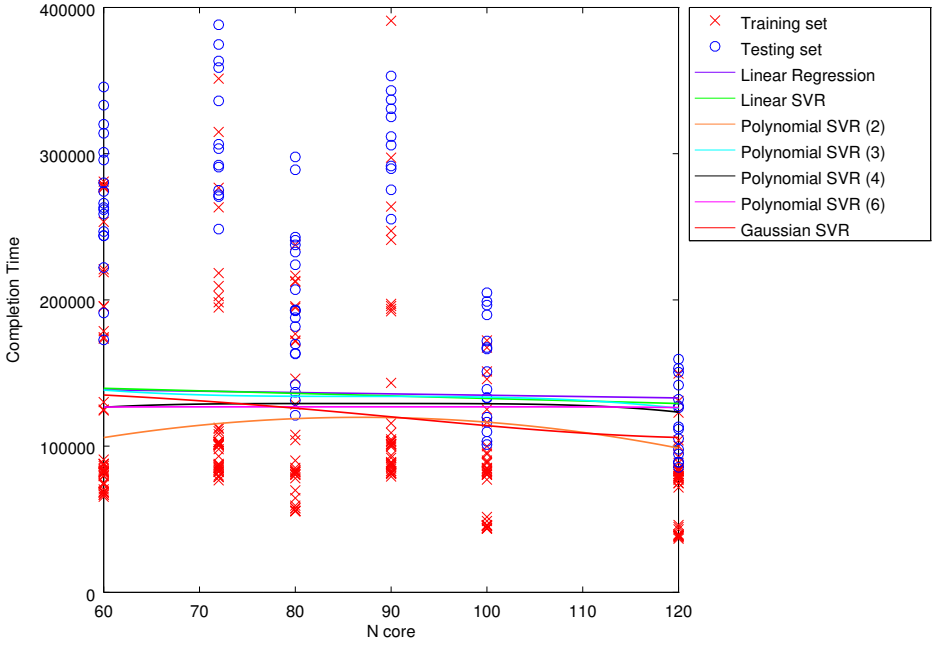

Figure 97: Completion time vs ncores for: R1,R2,R4 → R3 with datasize 250GB

5.5.2 R1,R2,R4 → R3 — Datasize 500GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.0914	0.9687	13457	0.0317
Linear SVR	0.0833	0.9741	13357	0.0308
Polynomial SVR (2)	0.6833	0.6864	118761	0.2933
Polynomial SVR (3)	0.2660	0.8449	42425	0.0947
Polynomial SVR (4)	0.7291	0.6105	129234	0.3055
Polynomial SVR (6)	0.7959	0.4006	136789	0.3104
Gaussian SVR	0.2844	0.8990	43387	0.1147

Table 98: Results for R1,R2,R4 → R3-500

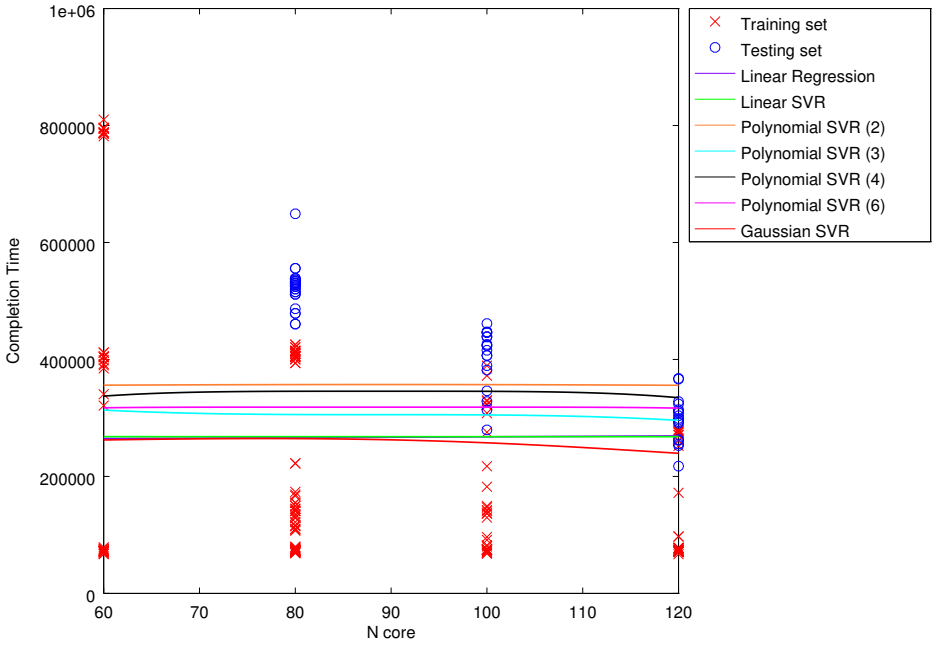

Figure 98: Completion time vs ncores for: R1,R2,R4 → R3 with datasize 500GB

5.5.3 R1,R2,R4 → R3 — Datasize 750GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.0581	0.9694	13358	0.0186
Linear SVR	0.0432	0.9924	10296	0.0152
Polynomial SVR (2)	0.4219	0.7036	106082	0.1502
Polynomial SVR (3)	0.2763	0.7323	53615	0.0742
Polynomial SVR (4)	0.6688	0.4095	168830	0.2349
Polynomial SVR (6)	0.6903	0.2073	166482	0.2269
Gaussian SVR	0.3779	0.5523	82501	0.1133

Table 99: Results for R1,R2,R4 → R3-750

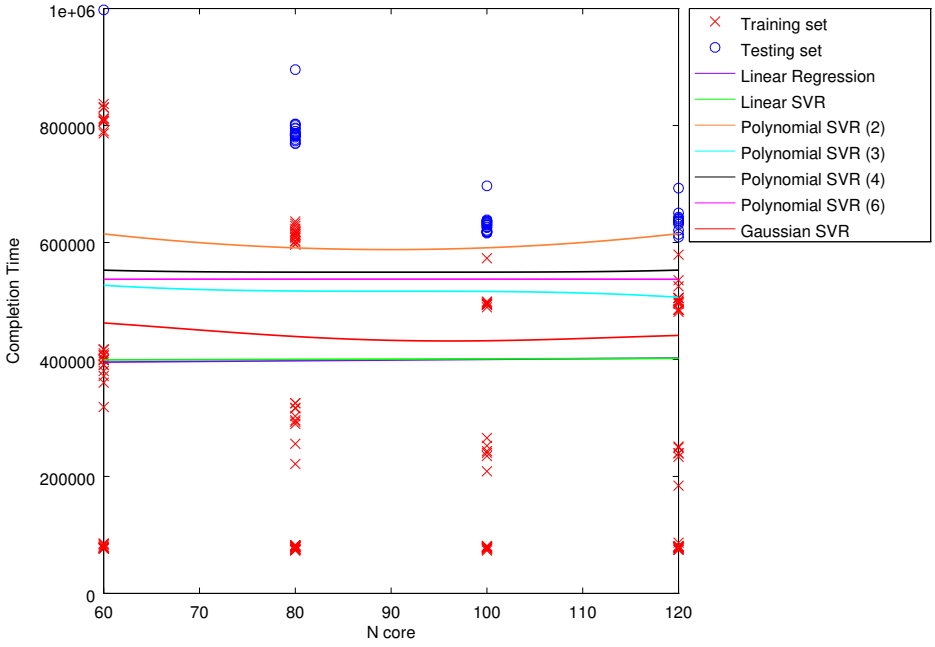

Figure 99: Completion time vs ncores for: R1,R2,R4 → R3 with datasize 750GB

5.5.4 R1,R2,R4 → R3 — Datasize 1000GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.5423	-0.7126	260530	0.2553
Linear SVR	0.5775	0.9500	275655	0.2718
Polynomial SVR (2)	0.5045	0.7467	202083	0.1809
Polynomial SVR (3)	0.4542	0.0345	148622	0.1290
Polynomial SVR (4)	0.5079	0.0024	181230	0.1545
Polynomial SVR (6)	0.5769	0.0175	203255	0.1711
Gaussian SVR	0.3920	0.7004	168171	0.1877

Table 100: Results for R1,R2,R4 → R3-1000

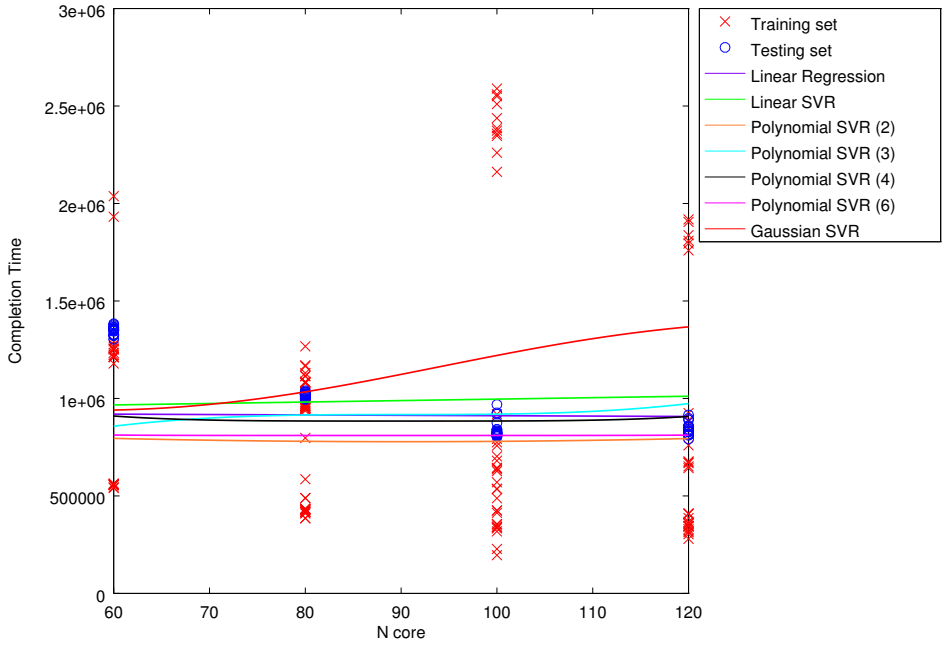

Figure 100: Completion time vs ncores for: R1,R2,R4 → R3 with datasize 1000GB

5.6 R2, R5 → R3 (Fast - Slow)

5.6.1 R2,R5 → R3 — Datasize 250GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	2.9358	-8.9645	225067	1.0067
Linear SVR	2.0189	0.1000	154238	0.6860
Polynomial SVR (2)	2.3413	0.0143	179635	0.8056
Polynomial SVR (3)	2.0566	0.0575	157367	0.6993
Polynomial SVR (4)	2.1957	0.0474	167202	0.7392
Polynomial SVR (6)	2.1578	0.0358	163842	0.7215
Gaussian SVR	2.0866	0.2880	160725	0.7205

Table 101: Results for R2,R5 → R3-250

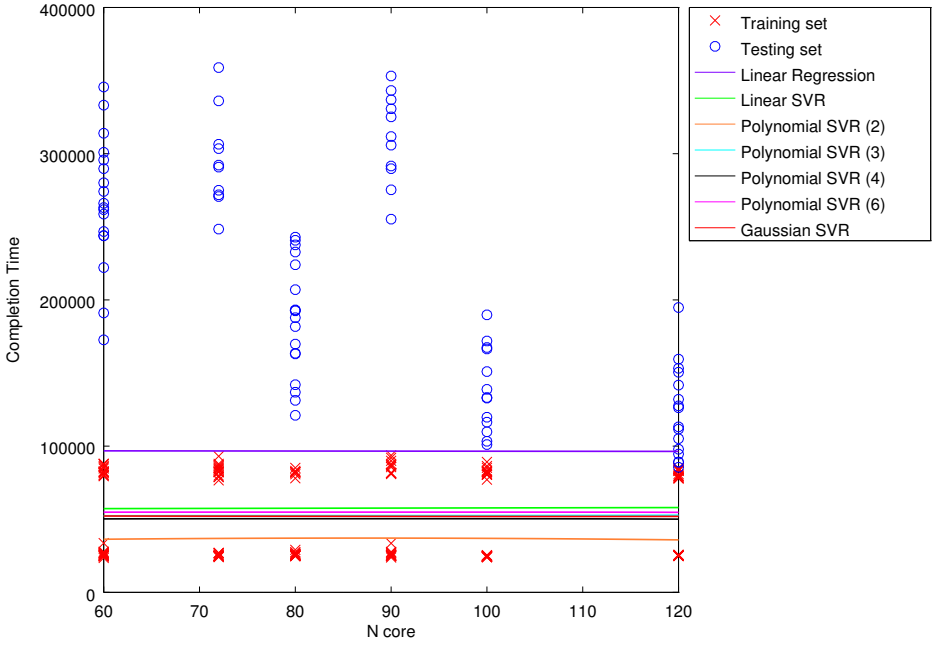

Figure 101: Completion time vs ncores for: R2,R5 → R3 with datasize 250GB

5.6.2 R2,R5 → R3 — Datasize 500GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	24.4317	-1732.1625	4082755	9.5054
Linear SVR	1.3901	0.0020	217391	0.4766
Polynomial SVR (2)	2.3279	0.8694	385016	0.8855
Polynomial SVR (3)	2.1472	0.6028	358017	0.8305
Polynomial SVR (4)	2.2733	0.4990	377968	0.8744
Polynomial SVR (6)	2.2643	0.2486	376874	0.8730
Gaussian SVR	2.2414	0.1919	373980	0.8683

Table 102: Results for R2,R5 → R3-500

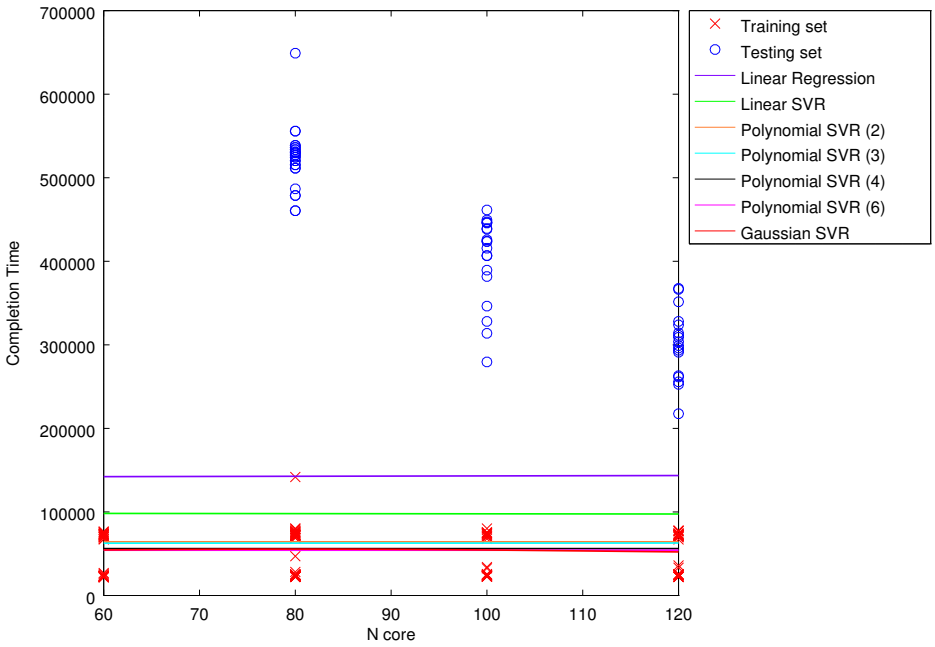

Figure 102: Completion time vs ncores for: R2,R5 → R3 with datasize 500GB

5.6.3 R2,R5 → R3 — Datasize 750GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	3.1762	-108.9756	905865	1.2847
Linear SVR	2.2593	0.0458	645146	0.9158
Polynomial SVR (2)	2.2682	0.1040	647657	0.9193
Polynomial SVR (3)	2.2586	0.3000	644988	0.9156
Polynomial SVR (4)	2.2552	0.0930	643893	0.9139
Polynomial SVR (6)	2.2549	0.0726	643840	0.9139
Gaussian SVR	2.2664	0.6737	647187	0.9187

Table 103: Results for R2,R5 → R3-750

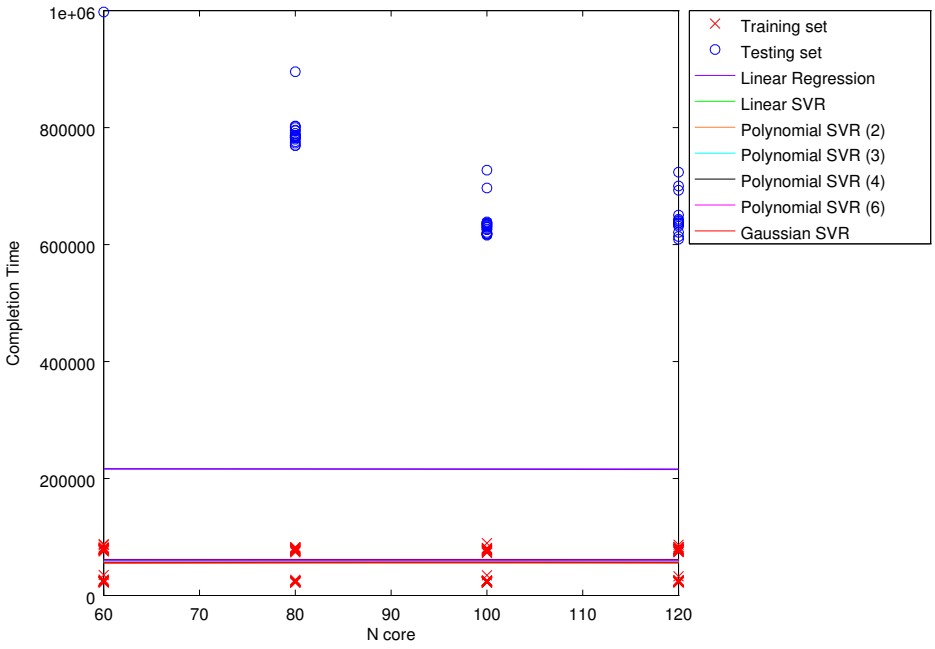

Figure 103: Completion time vs ncores for: R2,R5 → R3 with datasize 750GB

5.6.4 R2,R5 → R3 — Datasize 1000GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.4825	-0.2162	225269	0.2231
Linear SVR	1.5082	0.8905	700319	0.6897
Polynomial SVR (2)	1.8532	0.0474	856486	0.8417
Polynomial SVR (3)	1.7507	0.0017	810643	0.7966
Polynomial SVR (4)	1.8048	0.1072	838182	0.8256
Polynomial SVR (6)	1.8737	0.3832	871983	0.8603
Gaussian SVR	1.5367	0.8611	723592	0.7227

Table 104: Results for R2,R5 → R3-1000

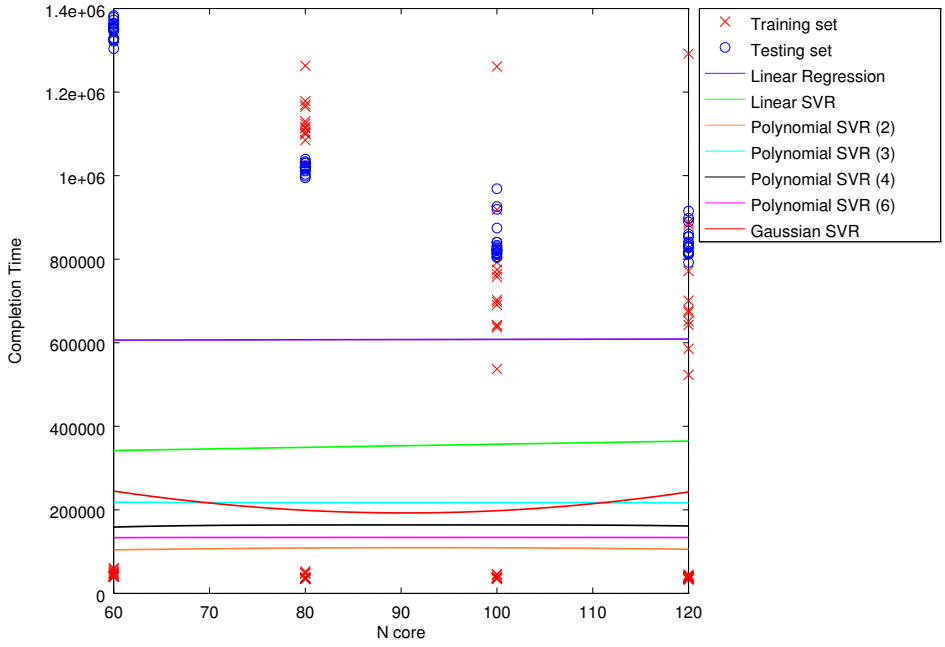

Figure 104: Completion time vs ncores for: R2,R5 → R3 with datasize 1000GB

5.7 R3, R4 → R2 (Fast - Slow)

5.7.1 R3,R4 → R2 — Datasize 250GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.7320	-324.7047	67111	0.8046
Linear SVR	0.6184	0.3343	56625	0.6788
Polynomial SVR (2)	2.2958	0.1994	209908	2.5228
Polynomial SVR (3)	0.6067	0.0155	47847	0.5734
Polynomial SVR (4)	2.0229	0.0101	180533	2.1680
Polynomial SVR (6)	1.7103	0.0153	146100	1.7568
Gaussian SVR	1.0720	0.2857	98273	1.1766

Table 105: Results for R3,R4 → R2-250

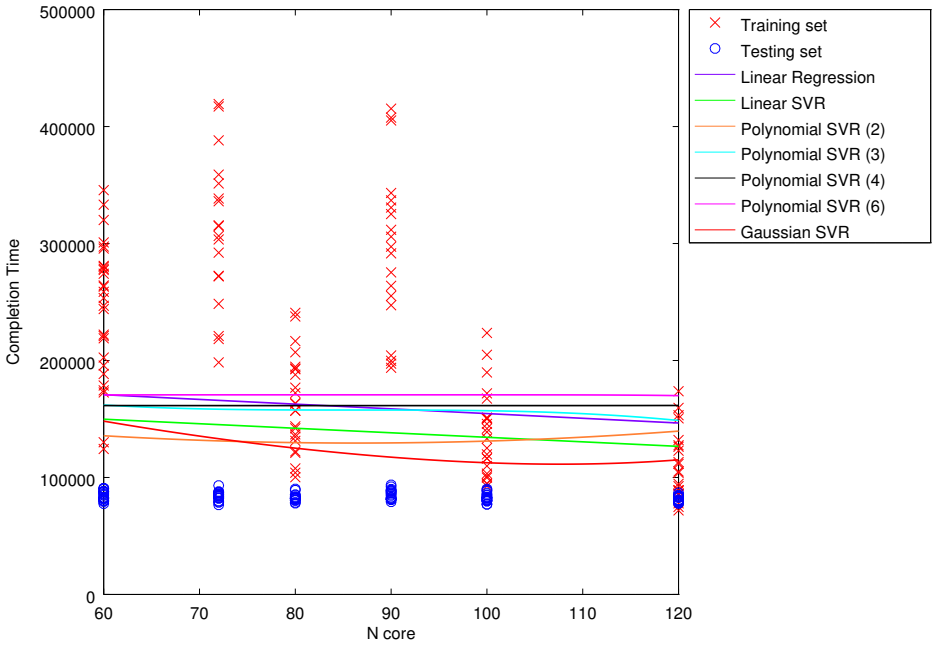

Figure 105: Completion time vs ncores for: R3,R4 → R2 with datasize 250GB

5.7.2 R3,R4 → R2 — Datasize 500GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.2274	-43.9565	50511	0.6952
Linear SVR	0.4418	0.0006	96855	1.3309
Polynomial SVR (2)	1.8749	0.0345	420366	5.7616
Polynomial SVR (3)	0.9866	0.0290	213001	2.9014
Polynomial SVR (4)	1.8453	0.0260	410413	5.6274
Polynomial SVR (6)	1.7365	0.0248	385605	5.2880
Gaussian SVR	1.4964	0.0442	337994	4.6265

Table 106: Results for R3,R4 → R2-500

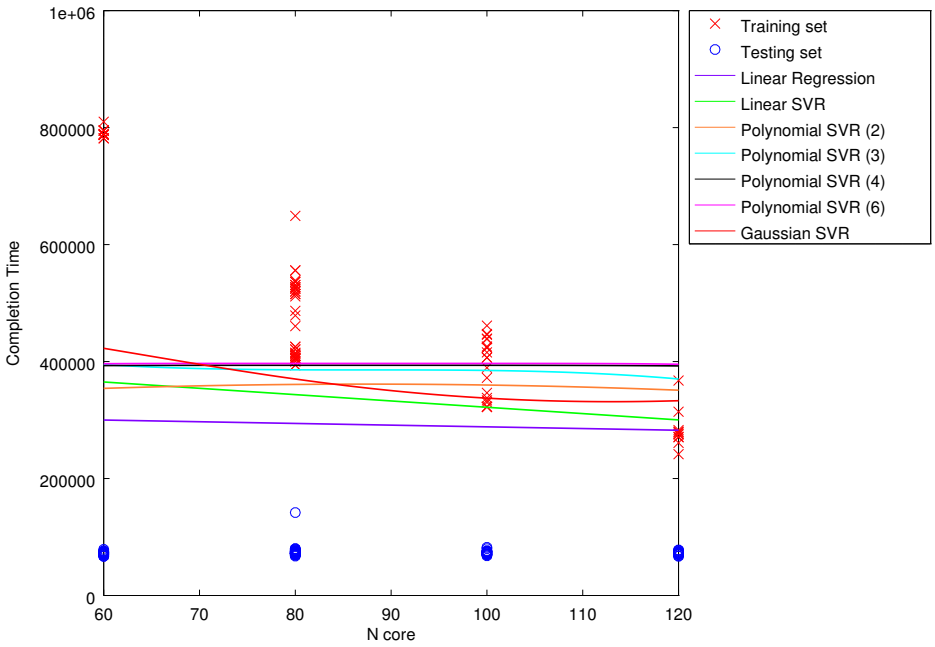

Figure 106: Completion time vs ncores for: R3,R4 → R2 with datasize 500GB

5.7.3 R3,R4 → R2 — Datasize 750GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.4540	-3592.1168	146163	1.8730
Linear SVR	0.4483	0.0421	137083	1.7535
Polynomial SVR (2)	2.4760	0.1503	788379	10.1194
Polynomial SVR (3)	1.1909	0.3274	304007	3.8427
Polynomial SVR (4)	2.6951	0.2837	803028	10.3423
Polynomial SVR (6)	3.3215	0.2411	930987	11.9433
Gaussian SVR	1.9111	0.1845	615634	7.8865

Table 107: Results for R3,R4 → R2-750

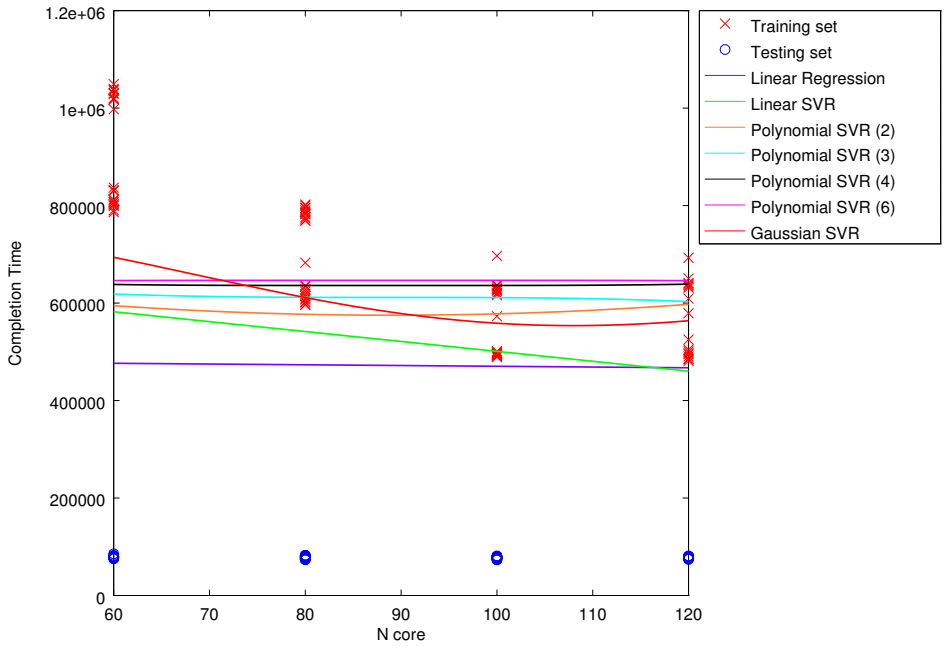

Figure 107: Completion time vs ncores for: R3,R4 → R2 with datasize 750GB

5.7.4 R3,R4 → R2 — Datasize 1000GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	11.2761	-583.0303	3410161	3.2357
Linear SVR	7.7162	0.9822	3657876	4.2191
Polynomial SVR (2)	8.0373	0.9441	2766694	2.6339
Polynomial SVR (3)	10.5398	0.8977	3750682	3.6335
Polynomial SVR (4)	10.3551	0.8329	3230811	2.9786
Polynomial SVR (6)	7.2376	0.6977	2298685	2.2241
Gaussian SVR	2.4222	0.5863	1124917	1.5115

Table 108: Results for R3,R4 → R2-1000

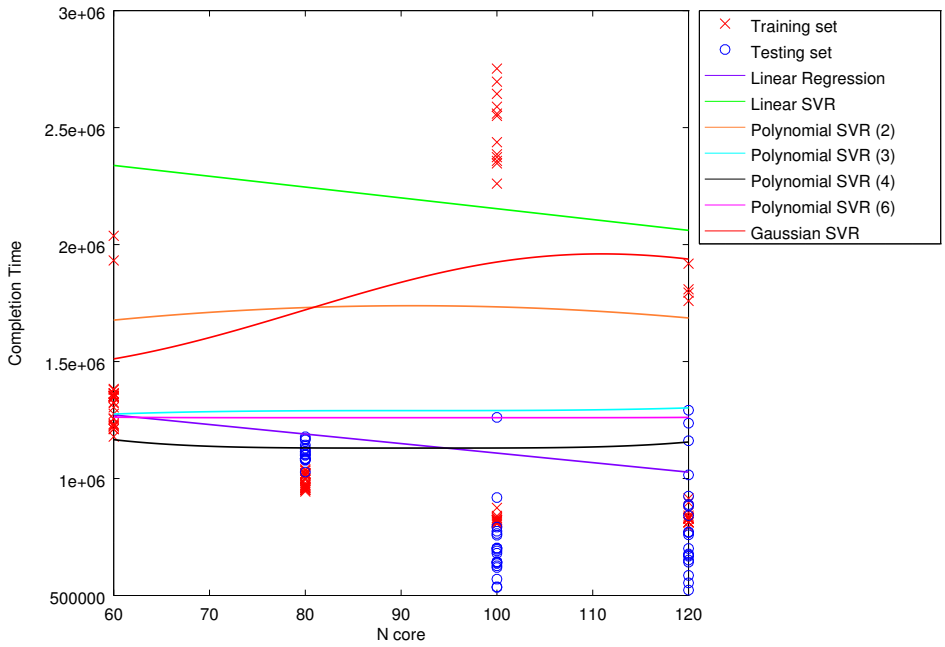

Figure 108: Completion time vs ncores for: R3,R4 → R2 with datasize 1000GB

S

6 Fixed Datasize, all the features,

6.1 Query Q2

6.1.1 Query Q2 — Datasize 40GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	1.6276	-0.9999	32359	0.1341
Linear SVR	0.3547	0.9266	6650	0.0274
Polynomial SVR (2)	0.8078	0.6069	15924	0.0649
Polynomial SVR (3)	0.5883	0.7715	10921	0.0446
Polynomial SVR (4)	1.0508	0.1893	19767	0.0789
Polynomial SVR (6)	1.1443	0.0177	21436	0.0850
Gaussian SVR	0.5819	0.8864	10799	0.0422

Table 109: Results for Q2-40GB

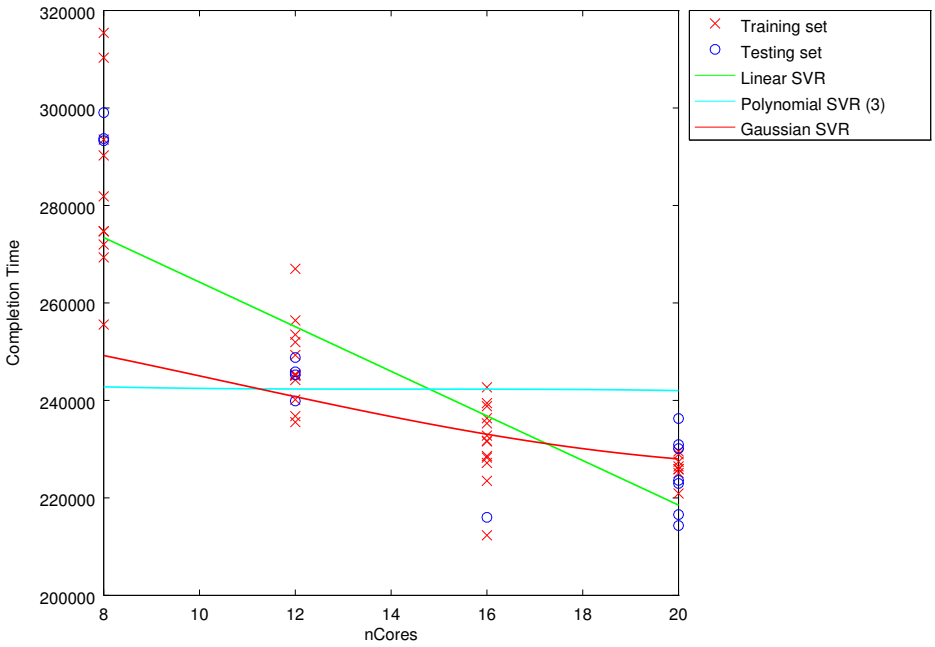

Figure 109: Completion time vs ncores for query Q2 with datasize 40

6.1.2 Query Q2 — Datasize 50GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.7974	0.4045	26855	0.0982
Linear SVR	0.3127	0.9112	10937	0.0419
Polynomial SVR (2)	0.7898	0.4304	28540	0.1044
Polynomial SVR (3)	0.7750	0.4572	30331	0.1124
Polynomial SVR (4)	1.2471	0.2083	37425	0.1433
Polynomial SVR (6)	0.9696	0.2704	37310	0.1350
Gaussian SVR	0.5052	0.8391	17201	0.0663

Table 110: Results for Q2-50GB

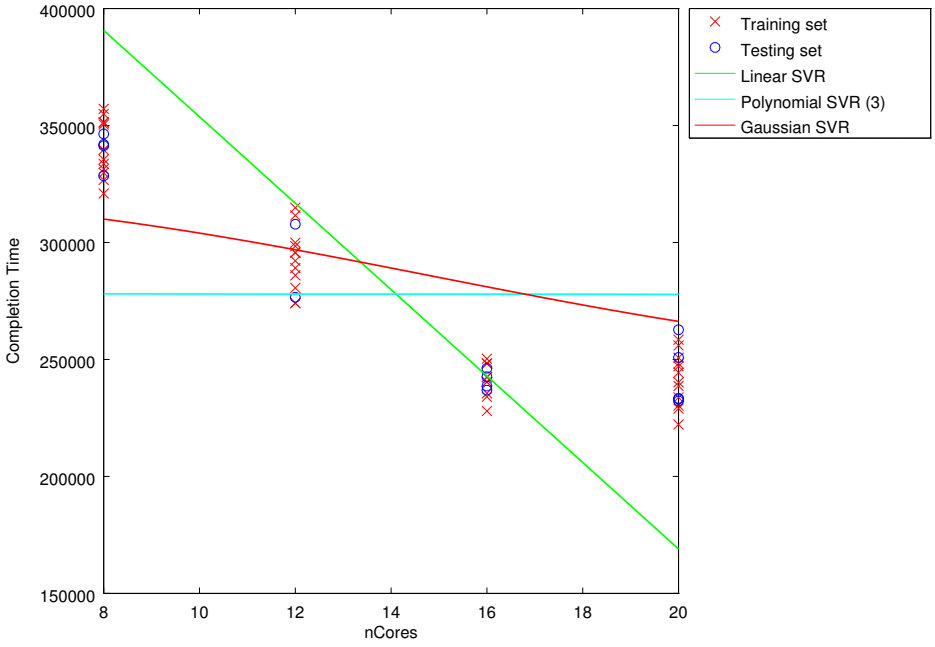

Figure 110: Completion time vs ncores for query Q2 with datasize 50

6.2 Query Q3

6.2.1 Query Q3 — Datasize 50GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.1821	0.9385	5147	0.0238
Linear SVR	0.3589	0.8794	10758	0.0502
Polynomial SVR (2)	0.7194	0.0867	23497	0.1173
Polynomial SVR (3)	0.4076	0.7912	11214	0.0514
Polynomial SVR (4)	0.5708	0.5249	14385	0.0683
Polynomial SVR (6)	0.6185	0.4444	17186	0.0817
Gaussian SVR	0.1952	0.9502	5697	0.0269

Table 111: Results for Q3-50GB

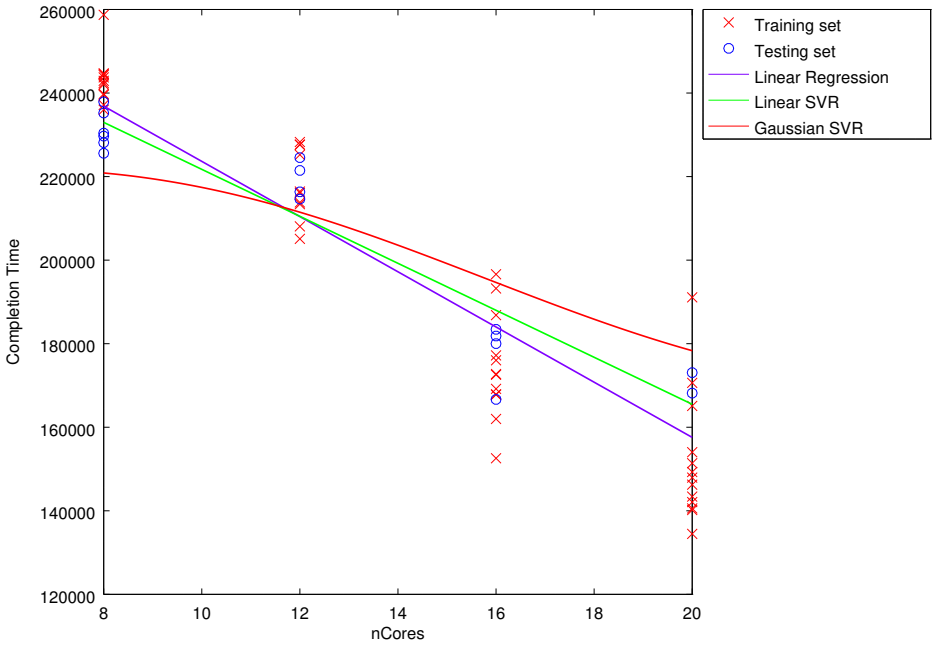

Figure 111: Completion time vs ncores for query Q3 with datasize 50

6.3 Query Q4

6.3.1 Query Q4 — Datasize 50GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.0990	0.9881	4688	0.0154
Linear SVR	0.1198	0.9849	6748	0.0217
Polynomial SVR (2)	0.7944	0.3575	53287	0.1731
Polynomial SVR (3)	0.4045	0.8300	25556	0.0869
Polynomial SVR (4)	0.7117	0.5033	42806	0.1298
Polynomial SVR (6)	0.7431	0.4922	45187	0.1372
Gaussian SVR	0.1867	0.9590	11207	0.0387

Table 112: Results for Q4-50GB

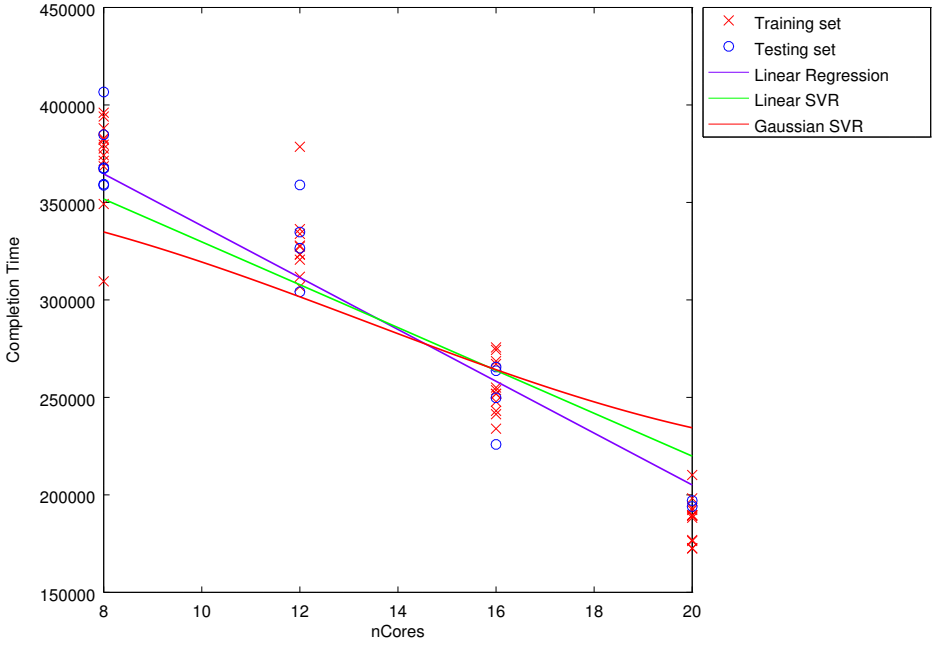

Figure 112: Completion time vs ncores for query Q4 with datasize 50

7 Fixed Datasize, only nCores,

7.1 Query Q2

7.1.1 Query Q2 — Datasize 40GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.4732	0.7213	11516	0.0462
Linear SVR	0.4773	0.7465	10981	0.0436
Polynomial SVR (2)	0.9249	0.1892	21428	0.0832
Polynomial SVR (3)	0.5122	0.6857	11435	0.0460
Polynomial SVR (4)	0.9083	0.2025	21353	0.0831
Polynomial SVR (6)	0.9293	0.2194	21908	0.0848
Gaussian SVR	0.4556	0.7493	9715	0.0382

Table 113: Results for Q2-40GB

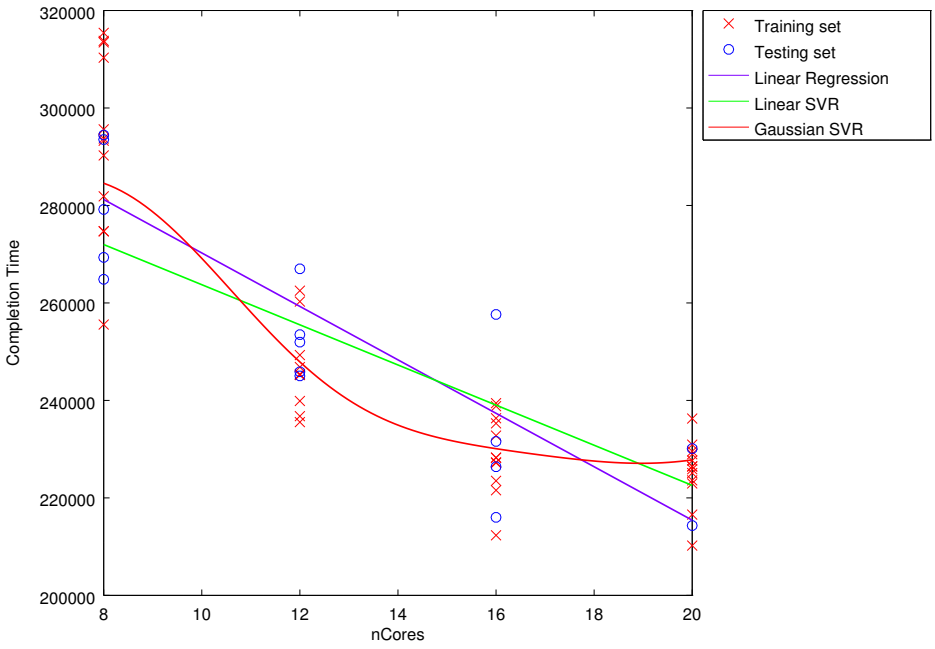

Figure 113: Completion time vs ncores for query Q2 with datasize 40

7.1.2 Query Q2 — Datasize 50GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.4165	0.8333	13809	0.0505
Linear SVR	0.4199	0.8411	14497	0.0514
Polynomial SVR (2)	1.0202	0.0000	38508	0.1365
Polynomial SVR (3)	0.5115	0.7521	16040	0.0601
Polynomial SVR (4)	1.0202	0.0000	38508	0.1365
Polynomial SVR (6)	1.0202	0.0000	38508	0.1365
Gaussian SVR	0.2887	0.9208	10435	0.0374

Table 114: Results for Q2-50GB

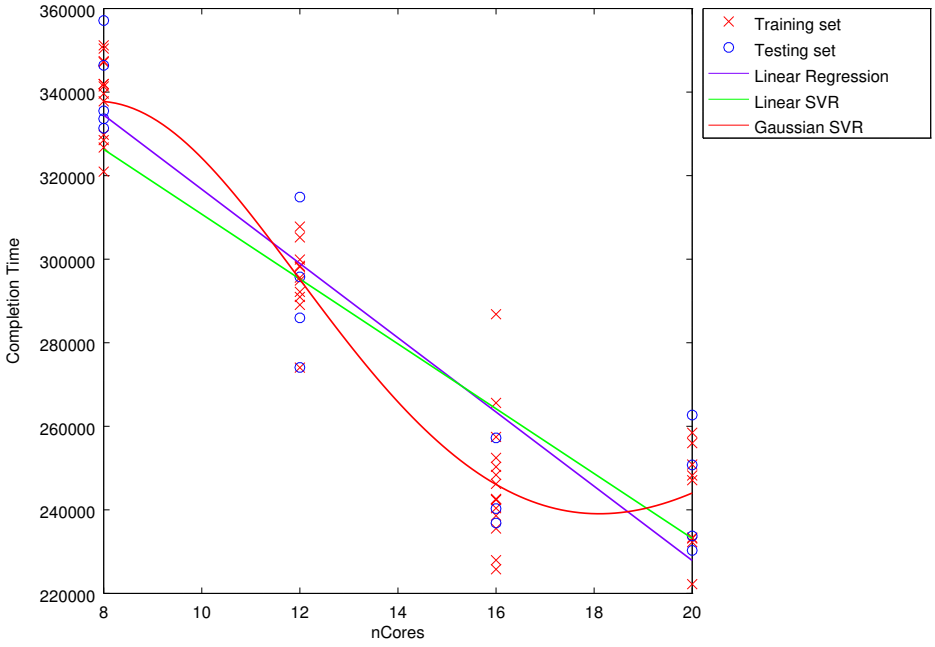

Figure 114: Completion time vs ncores for query Q2 with datasize 50

7.2 Query Q3

7.2.1 Query Q3 — Datasize 50GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.3813	0.7992	11506	0.0597
Linear SVR	0.3626	0.9342	10534	0.0556
Polynomial SVR (2)	0.8629	-0.0000	27382	0.1376
Polynomial SVR (3)	0.4068	0.8520	11280	0.0585
Polynomial SVR (4)	0.8629	-0.0000	27382	0.1376
Polynomial SVR (6)	0.8629	-0.0000	27382	0.1376
Gaussian SVR	0.2917	0.9418	8764	0.0471

Table 115: Results for Q3-50GB

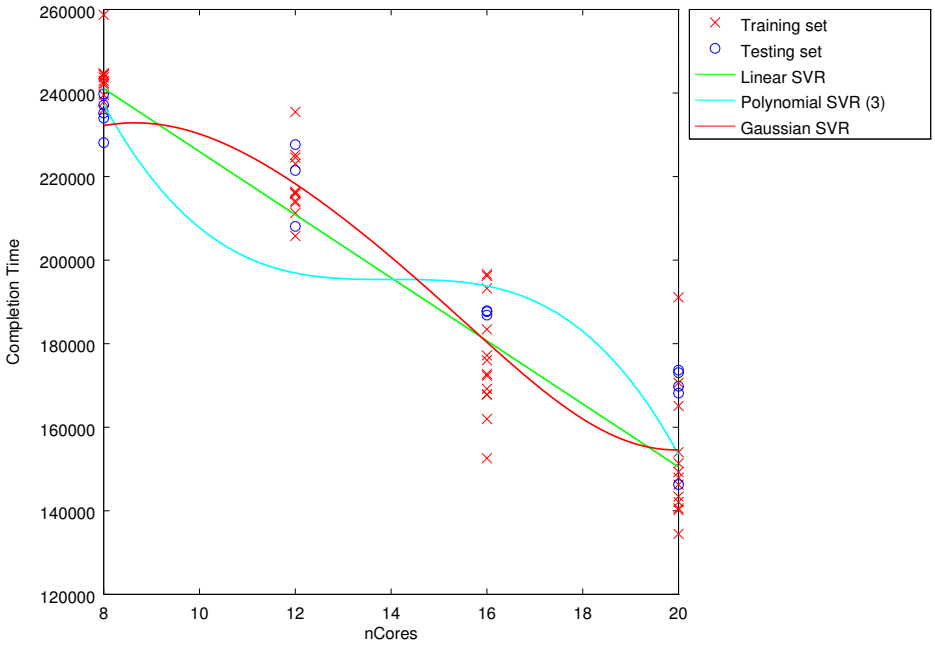

Figure 115: Completion time vs ncores for query Q3 with datasize 50

7.3 Query Q4

7.3.1 Query Q4 — Datasize 50GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.3026	0.9203	14054	0.0471
Linear SVR	0.3036	0.9208	14764	0.0491
Polynomial SVR (2)	1.0691	0.0077	71807	0.2703
Polynomial SVR (3)	0.4319	0.8457	25075	0.0836
Polynomial SVR (4)	1.0698	0.0077	71807	0.2722
Polynomial SVR (6)	1.0698	0.0077	71807	0.2722
Gaussian SVR	0.2880	0.9306	14033	0.0477

Table 116: Results for Q4-50GB

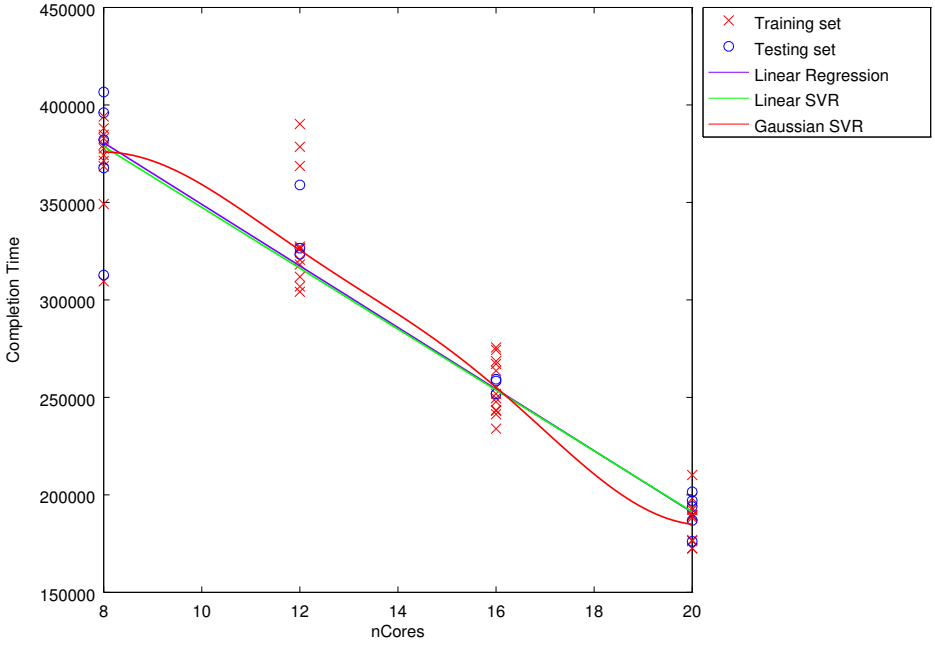

Figure 116: Completion time vs ncores for query Q4 with datasize 50

8 Fixed Datasize, all features with 1/nCores,

8.1 Query Q2

8.1.1 Query Q2 — Datasize 40GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	1.1724	-0.0377	24035	0.1012
Linear SVR	0.3713	0.9310	7170	0.0293
Polynomial SVR (2)	0.7597	0.7219	15618	0.0647
Polynomial SVR (3)	0.5592	0.8204	10138	0.0414
Polynomial SVR (4)	0.9967	0.4517	18992	0.0756
Polynomial SVR (6)	1.1150	0.0733	20847	0.0824
Gaussian SVR	0.5455	0.9096	10367	0.0410

Table 117: Results for Q2-40GB

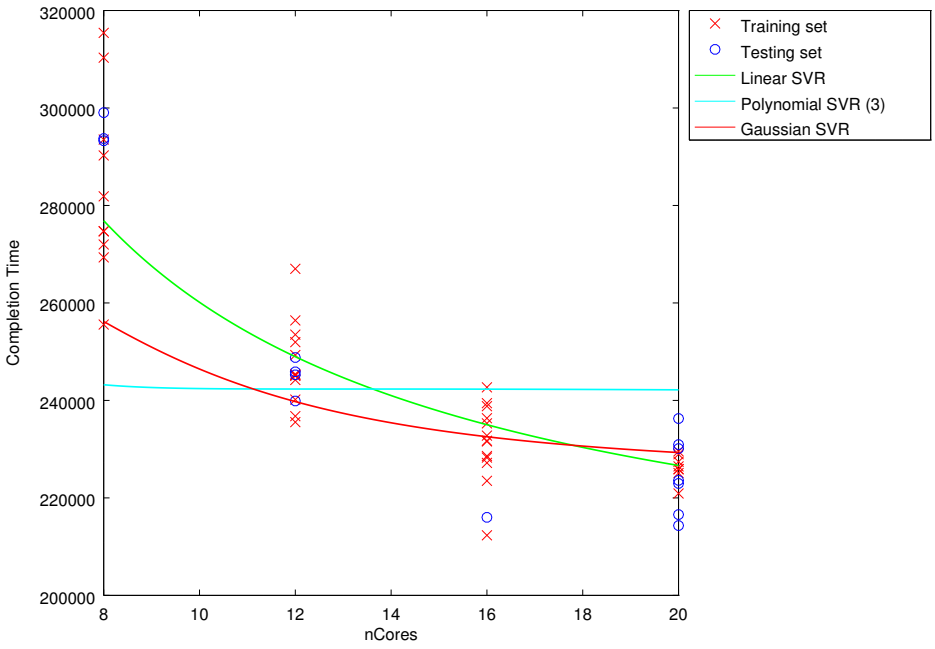

Figure 117: Completion time vs ncores for query Q2 with datasize 40

8.1.2 Query Q2 — Datasize 50GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.9401	0.1721	35276	0.1336
Linear SVR	0.4863	0.7826	15991	0.0631
Polynomial SVR (2)	0.7756	0.4662	28112	0.1025
Polynomial SVR (3)	0.7518	0.4994	26075	0.0987
Polynomial SVR (4)	1.1967	0.2245	36195	0.1382
Polynomial SVR (6)	1.1080	0.1845	37971	0.1412
Gaussian SVR	0.4857	0.8544	16546	0.0638

Table 118: Results for Q2-50GB

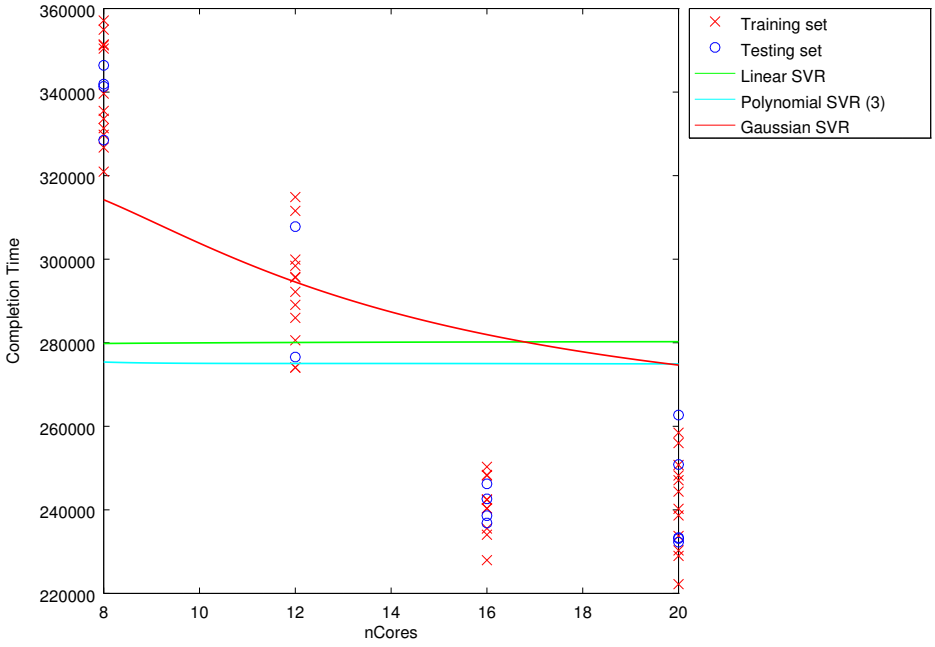

Figure 118: Completion time vs ncores for query Q2 with datasize 50

8.2 Query Q3

8.2.1 Query Q3 — Datasize 50GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.1821	0.9385	5147	0.0238
Linear SVR	0.3698	0.7939	9957	0.0472
Polynomial SVR (2)	0.6657	0.2370	22068	0.1086
Polynomial SVR (3)	0.4115	0.7344	10516	0.0487
Polynomial SVR (4)	0.6275	0.4275	16290	0.0778
Polynomial SVR (6)	0.6407	0.4513	16807	0.0790
Gaussian SVR	0.2121	0.9383	6227	0.0295

Table 119: Results for Q3-50GB

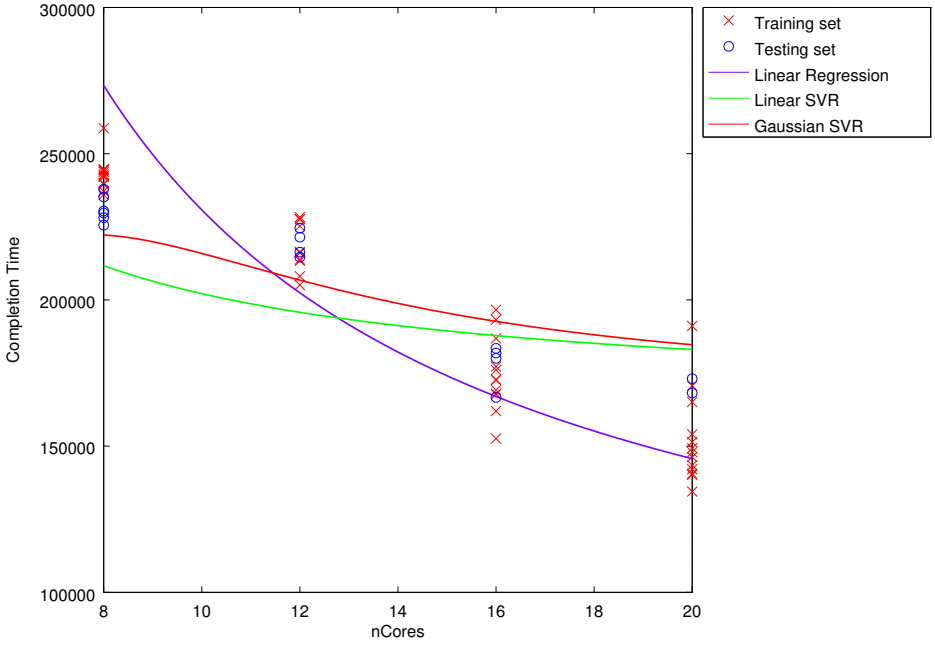

Figure 119: Completion time vs ncores for query Q3 with datasize 50

8.3 Query Q4

8.3.1 Query Q4 — Datasize 50GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.0990	0.9881	4688	0.0154
Linear SVR	0.1853	0.9602	11231	0.0364
Polynomial SVR (2)	0.6609	0.5640	41416	0.1327
Polynomial SVR (3)	0.3972	0.8135	23263	0.0776
Polynomial SVR (4)	0.4891	0.7383	28158	0.0915
Polynomial SVR (6)	0.6421	0.6013	39270	0.1209
Gaussian SVR	0.2128	0.9463	12910	0.0451

Table 120: Results for Q4-50GB

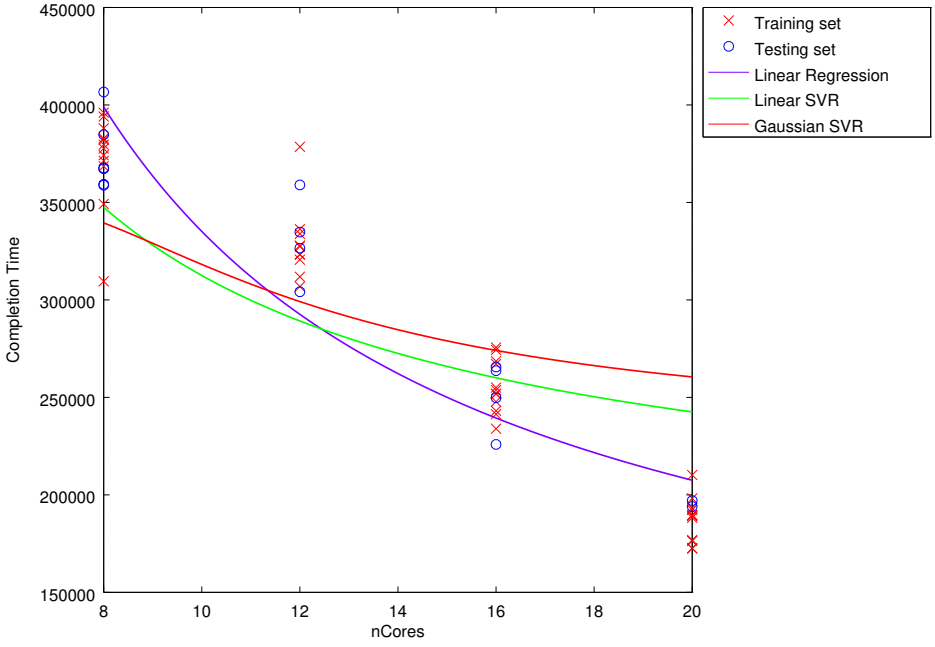

Figure 120: Completion time vs ncores for query Q4 with datasize 50

9 Fixed Datasize, only 1/nCores,

9.1 Query Q2

9.1.1 Query Q2 — Datasize 40GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.4755	0.7186	9942	0.0394
Linear SVR	0.4356	0.7670	9580	0.0378
Polynomial SVR (2)	0.7545	0.3885	17726	0.0705
Polynomial SVR (3)	0.5410	0.6751	12134	0.0488
Polynomial SVR (4)	0.6630	0.5285	15720	0.0630
Polynomial SVR (6)	0.6016	0.5722	14553	0.0587
Gaussian SVR	0.4536	0.7500	9672	0.0381

Table 121: Results for Q2-40GB

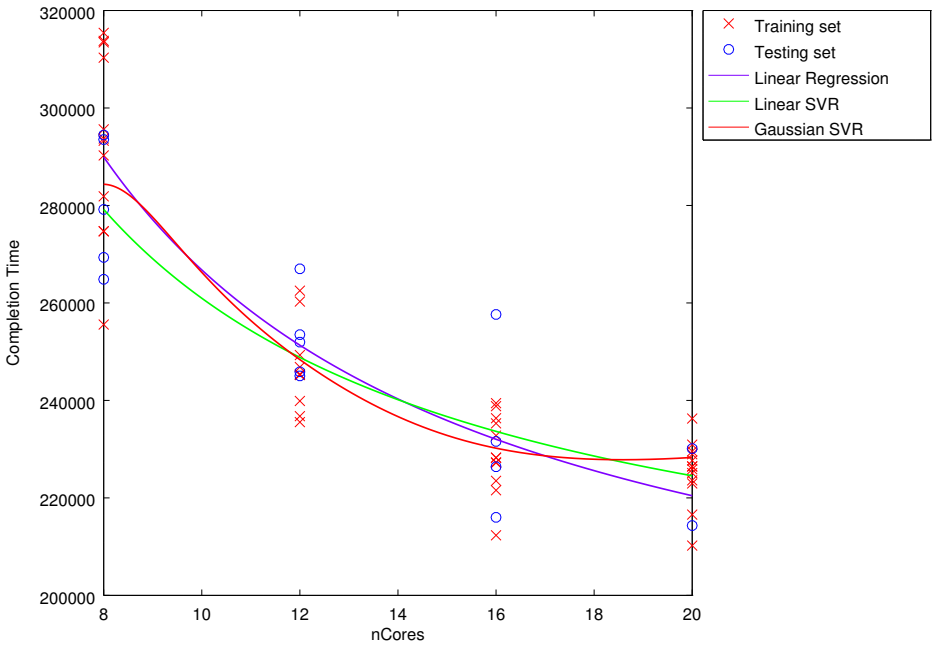

Figure 121: Completion time vs ncores for query Q2 with datasize 40

9.1.2 Query Q2 — Datasize 50GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.3294	0.8958	11482	0.0409
Linear SVR	0.3301	0.8973	11520	0.0410
Polynomial SVR (2)	0.8382	0.4016	32740	0.1153
Polynomial SVR (3)	0.4515	0.8074	15713	0.0568
Polynomial SVR (4)	0.6385	0.6123	21200	0.0775
Polynomial SVR (6)	0.5805	0.6807	18767	0.0683
Gaussian SVR	0.2887	0.9208	10434	0.0374

Table 122: Results for Q2-50GB

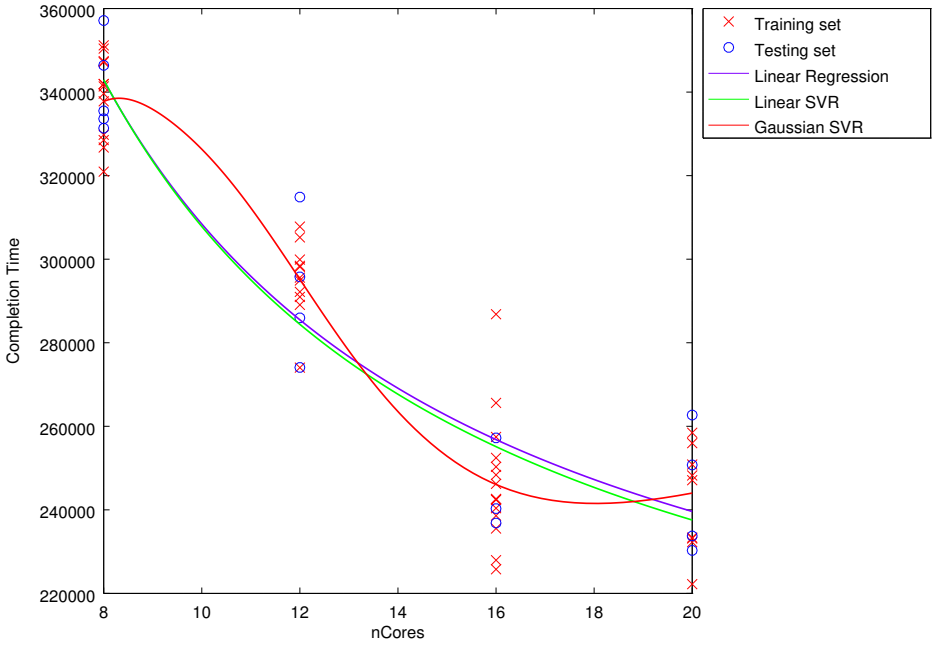

Figure 122: Completion time vs ncores for query Q2 with datasize 50

9.2 Query Q3

9.2.1 Query Q3 — Datasize 50GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.4364	0.7371	14301	0.0707
Linear SVR	0.3581	0.8593	10636	0.0527
Polynomial SVR (2)	0.7692	0.1922	22882	0.1221
Polynomial SVR (3)	0.4719	0.7356	12895	0.0635
Polynomial SVR (4)	0.6625	0.4085	18813	0.0980
Polynomial SVR (6)	0.6111	0.5019	16636	0.0864
Gaussian SVR	0.2815	0.9430	8413	0.0456

Table 123: Results for Q3-50GB

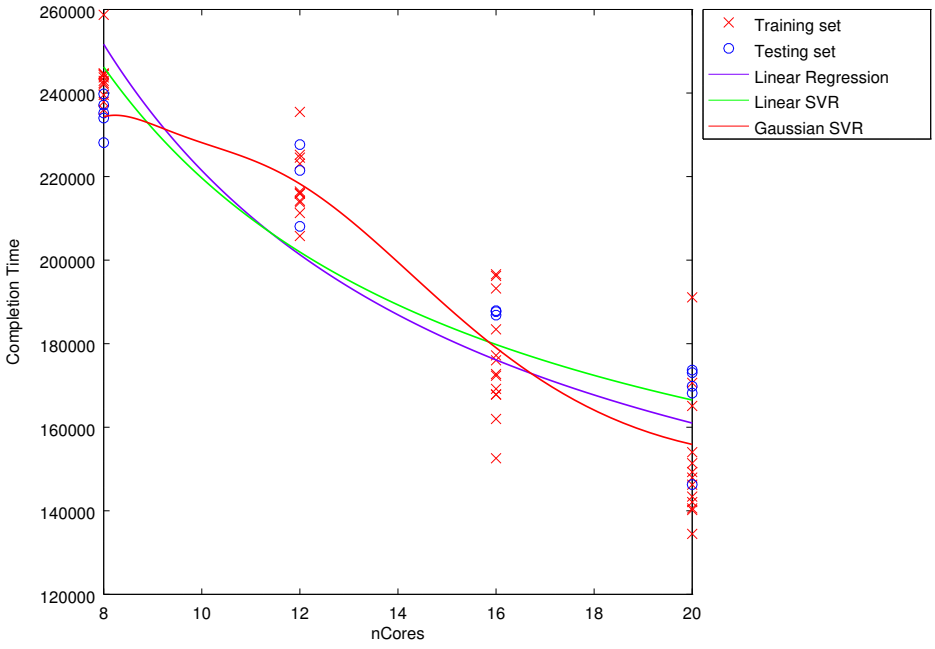

Figure 123: Completion time vs ncores for query Q3 with datasize 50

9.3 Query Q4

9.3.1 Query Q4 — Datasize 50GB

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.4439	0.8284	26027	0.0969
Linear SVR	0.4490	0.8300	27044	0.1029
Polynomial SVR (2)	1.0085	0.1628	66229	0.2716
Polynomial SVR (3)	0.6029	0.7055	37412	0.1473
Polynomial SVR (4)	0.8773	0.3698	55542	0.2300
Polynomial SVR (6)	0.7938	0.4631	47441	0.1942
Gaussian SVR	0.2888	0.9308	14254	0.0484

Table 124: Results for Q4-50GB

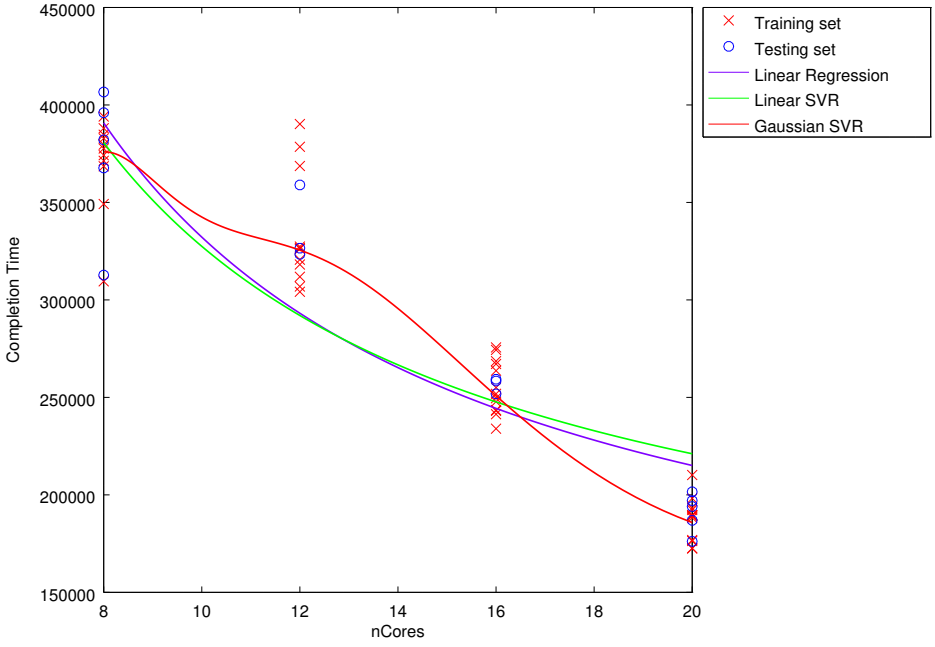

Figure 124: Completion time vs ncores for query Q4 with datasize 50