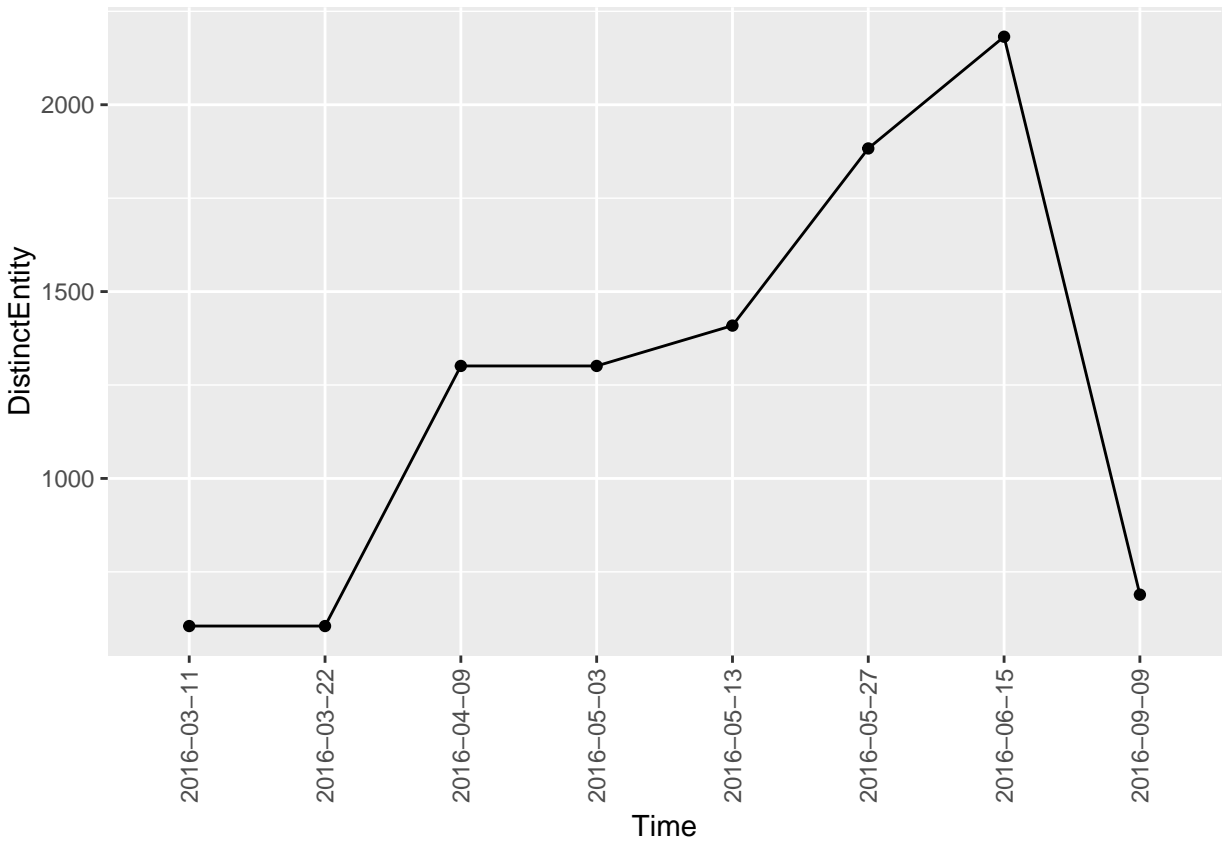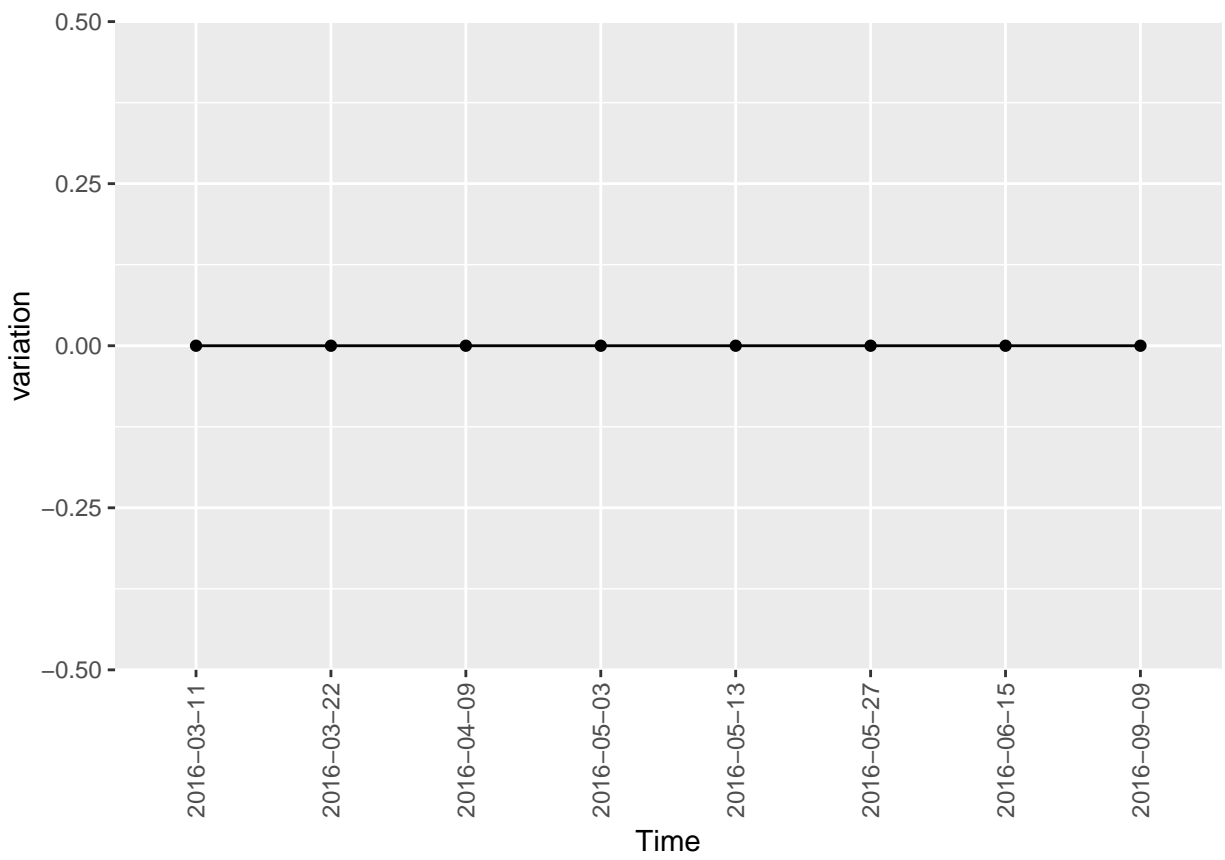# Data Quality Metrics

Rifat

Tuesday, October 25, 2016

In this document presented, a detials analysis on the 3city KB dataset. The analysis based on two types Events and places. Details on each type summary data presented end of the document.
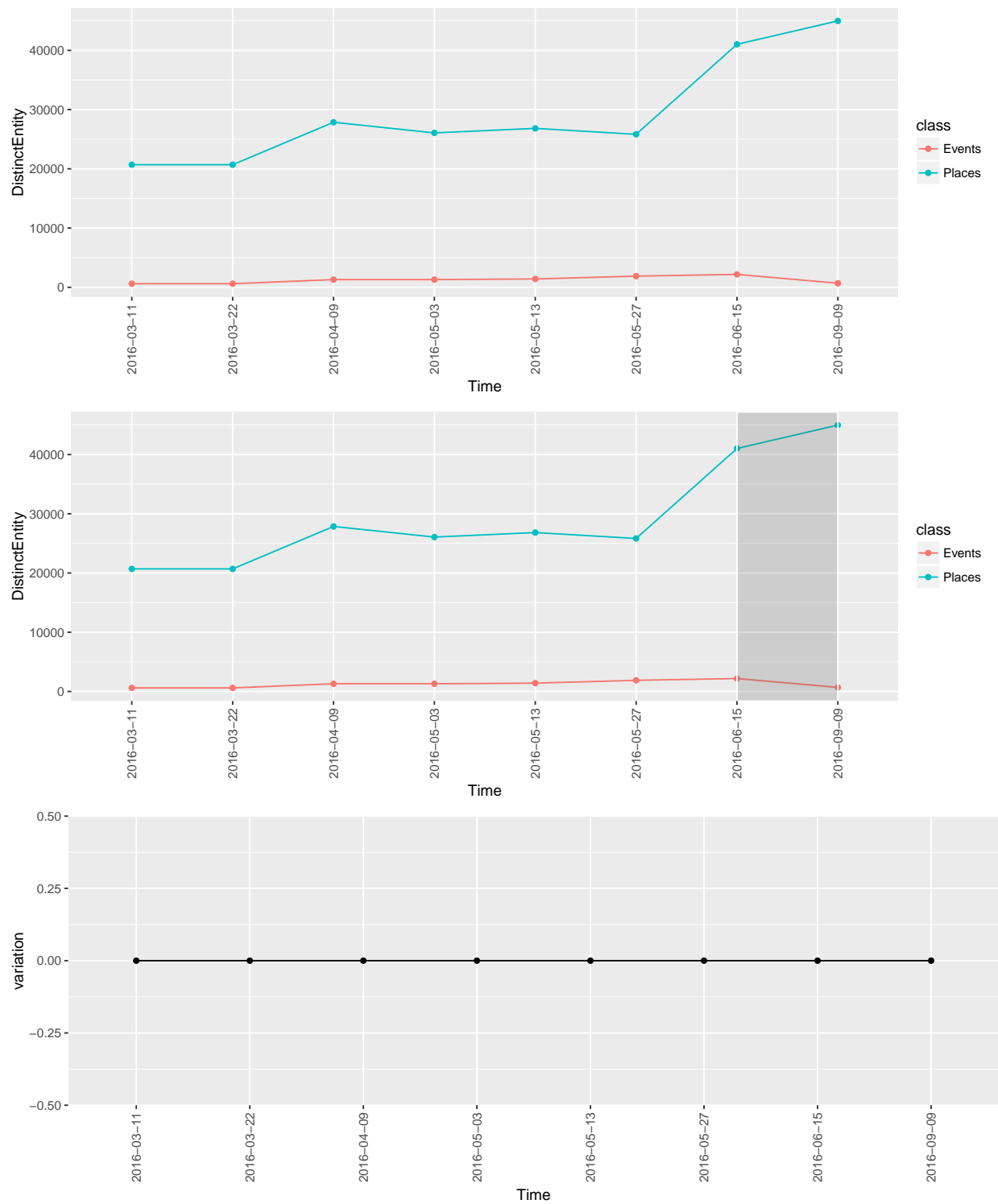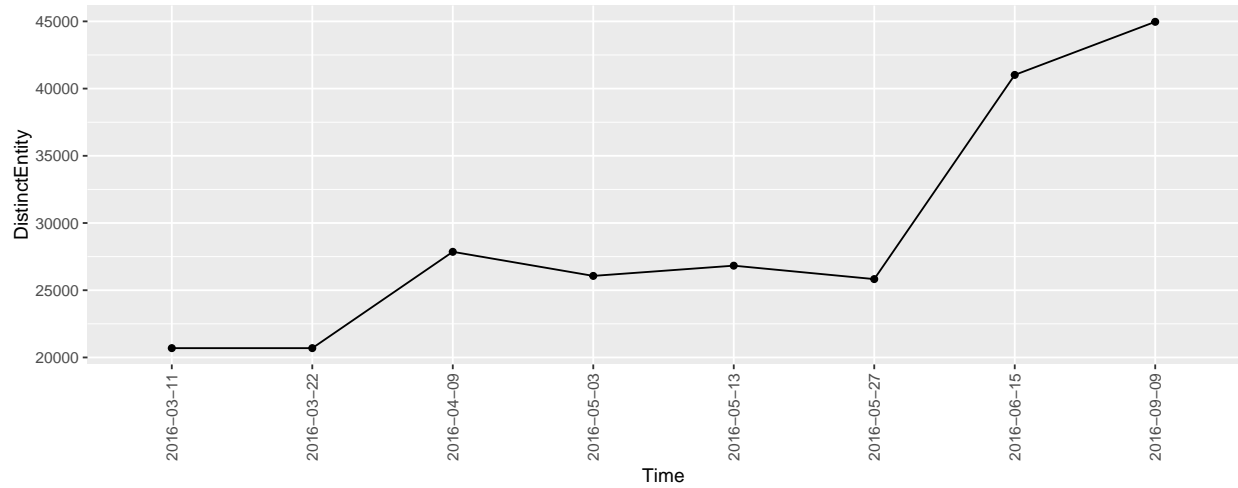
Entity Level

Persistency

Events persistency:

Events Entity persistency=0
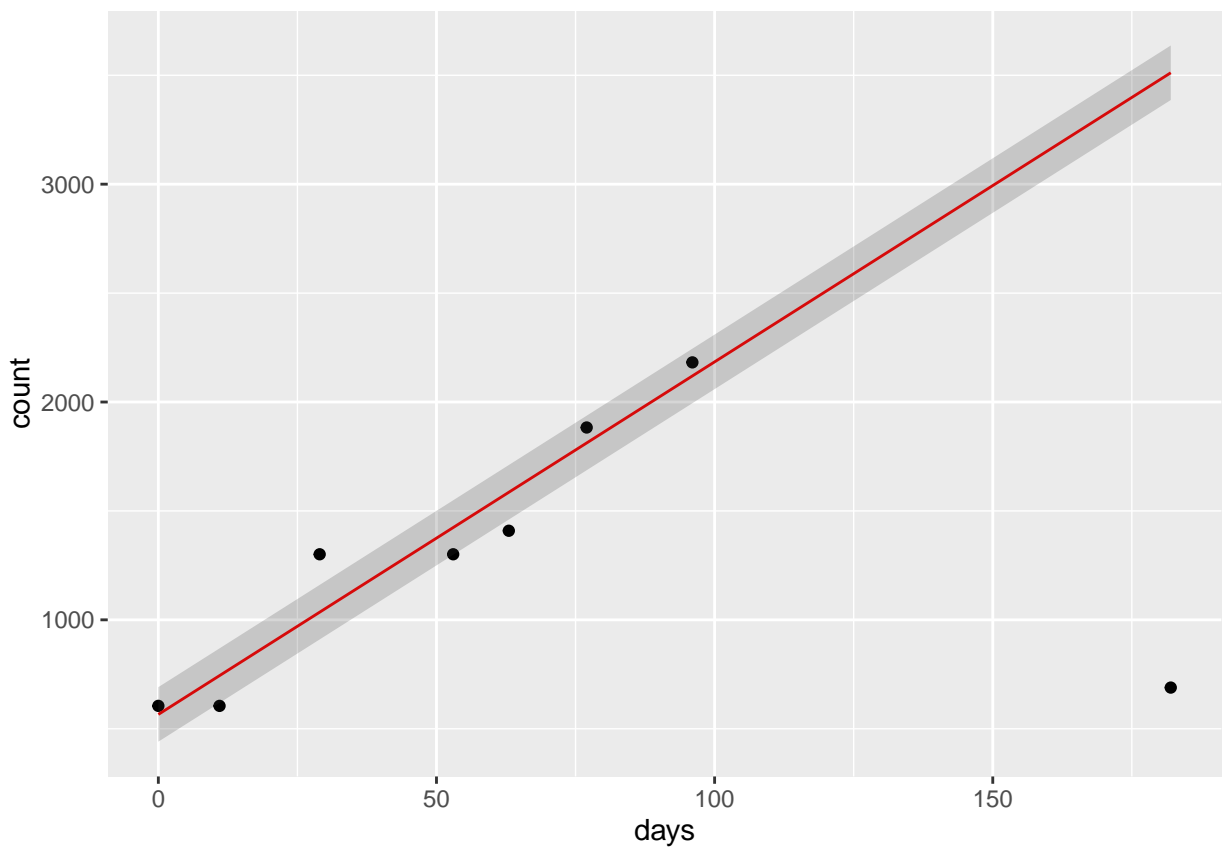
Places persistency:

Places Entity persistency=1

Stability

Events Stability:

The normalized distance(ND)=(abs(Last TimeSeries Value - Predicted Value)/mean(abs(Residuals))= 22.5481796

Stability is 1 if ND<1 or Stability is 0 if ND>=1

Stability=0

Places Stability:
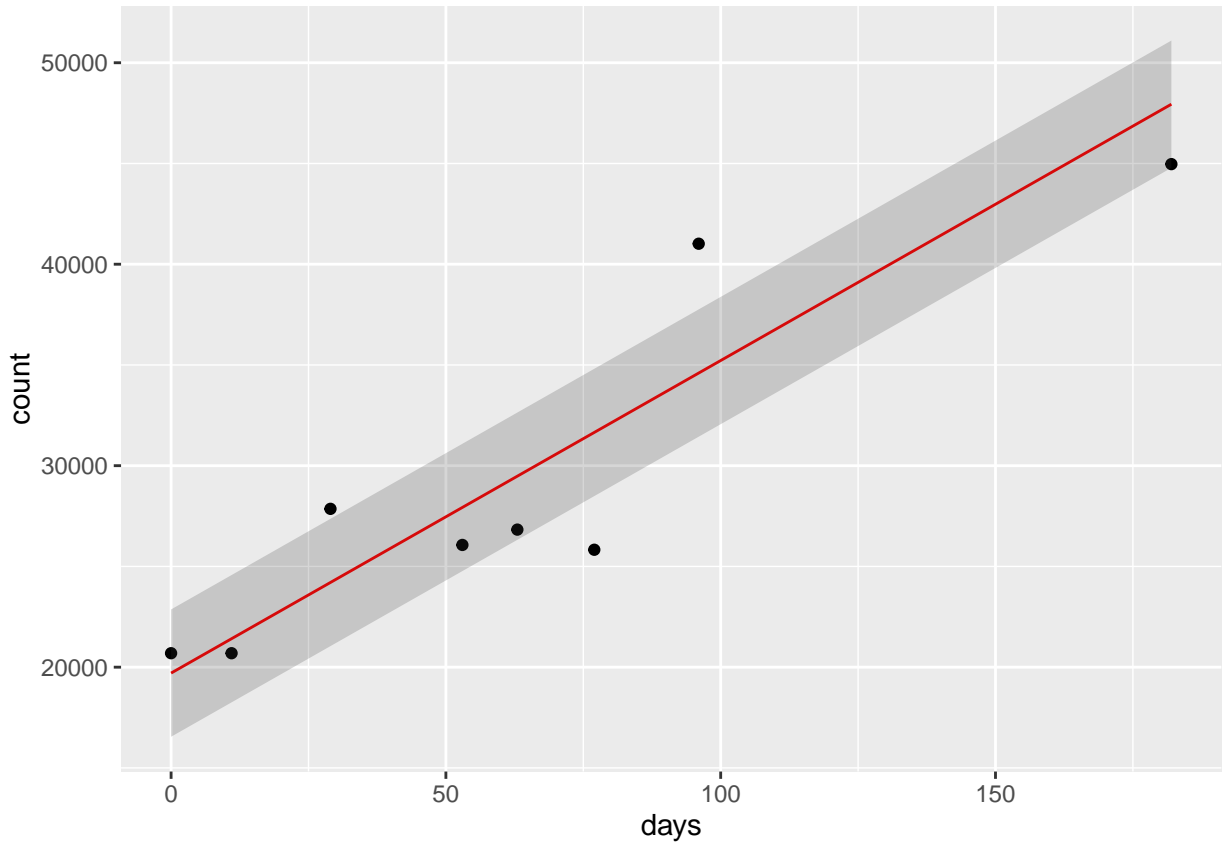
The normalized distance(ND)=(abs(Last TimeSeries Value - Predicted Value)/mean(abs(Residuals))= 0.9411357

Stability is 1 if ND<1 or Stability is 0 if ND>=1

Stability=1

Predicate Level

Correctness

Events Correctness:

Time	DistinctEntity	Predicate	freq	evDatPctNormalizedFreq

Places Correctness:

	Time	DistinctEntity	Predicate	freq	NormalizedFreq
2	2016-03-11	20692	http://data.linkedevents.org/def/location# businessTypeTop	19393	0.9372221

	Time	DistinctEntity	Predicate	freq	NormalizedFreq
4	2016-03-11	20692	http://linkedevents.org/ontology/poster	17777	0.8591243
7	2016-03-11	20692	http://schema.org/aggregateRating	435	0.0210226
8	2016-03-11	20692	http://schema.org/description	311	0.0150300
9	2016-03-11	20692	http://schema.org/image	2473	0.1195148
10	2016-03-11	20692	http://schema.org/interactionCount	6919	0.3343804
11	2016-03-11	20692	http://schema.org/location	20666	0.9987435
12	2016-03-11	20692	http://schema.org/openingHours	2363	0.1141987
13	2016-03-11	20692	http://schema.org/priceRange	15	0.0007249
14	2016-03-11	20692	http://schema.org/review	2211	0.1068529
15	2016-03-11	20692	http://schema.org/telephone	18325	0.8856080
16	2016-03-11	20692	http://schema.org/url	4309	0.2082447
22	2016-03-11	20692	http://www.w3.org/2002/07/owl#sameAs	20597	0.9954089
25	2016-03-22	20692	http://data.linkedevents.org/def/location#businessTypeTop	19393	0.9372221
27	2016-03-22	20692	http://linkedevents.org/ontology/poster	17777	0.8591243
30	2016-03-22	20692	http://schema.org/aggregateRating	435	0.0210226
31	2016-03-22	20692	http://schema.org/description	311	0.0150300
32	2016-03-22	20692	http://schema.org/image	2473	0.1195148
33	2016-03-22	20692	http://schema.org/interactionCount	6919	0.3343804
34	2016-03-22	20692	http://schema.org/location	20666	0.9987435
35	2016-03-22	20692	http://schema.org/openingHours	2363	0.1141987
36	2016-03-22	20692	http://schema.org/priceRange	15	0.0007249
37	2016-03-22	20692	http://schema.org/review	2211	0.1068529
38	2016-03-22	20692	http://schema.org/telephone	18325	0.8856080
39	2016-03-22	20692	http://schema.org/url	4309	0.2082447
45	2016-03-22	20692	http://www.w3.org/2002/07/owl#sameAs	20597	0.9954089
48	2016-04-09	27858	http://data.linkedevents.org/def/location#businessTypeTop	26473	0.9502836

	Time	DistinctEntity	Predicate	freq	NormalizedFreq
50	2016-04-09	27858	http://linkedevents.org/ontology/poster	23897	0.8578146
53	2016-04-09	27858	http://schema.org/aggregateRating	798	0.0286453
54	2016-04-09	27858	http://schema.org/description	345	0.0123842
55	2016-04-09	27858	http://schema.org/image	13397	0.4809032
56	2016-04-09	27858	http://schema.org/interactionCount	8999	0.3230311
57	2016-04-09	27858	http://schema.org/location	27793	0.9976667
58	2016-04-09	27858	http://schema.org/openingHours	3173	0.1138991
59	2016-04-09	27858	http://schema.org/priceRange	25	0.0008974
60	2016-04-09	27858	http://schema.org/review	4332	0.1555029
61	2016-04-09	27858	http://schema.org/telephone	22411	0.8044727
62	2016-04-09	27858	http://schema.org/url	5432	0.1949889
68	2016-04-09	27858	http://www.w3.org/2002/07/owl#sameAs	27735	0.9955848
71	2016-05-03	26066	http://data.linkedevents.org/def/location#businessTypeTop	25080	0.9621729
73	2016-05-03	26066	http://linkedevents.org/ontology/poster	22546	0.8649582
76	2016-05-03	26066	http://schema.org/aggregateRating	672	0.0257807
77	2016-05-03	26066	http://schema.org/description	314	0.0120463
78	2016-05-03	26066	http://schema.org/image	12214	0.4685798
79	2016-05-03	26066	http://schema.org/interactionCount	7985	0.3063378
80	2016-05-03	26066	http://schema.org/location	26011	0.9978900
81	2016-05-03	26066	http://schema.org/openingHours	2935	0.1125988
82	2016-05-03	26066	http://schema.org/priceRange	22	0.0008440
83	2016-05-03	26066	http://schema.org/review	3939	0.1511164
84	2016-05-03	26066	http://schema.org/telephone	20938	0.8032686
85	2016-05-03	26066	http://schema.org/url	5049	0.1937006
91	2016-05-03	26066	http://www.w3.org/2002/07/owl#sameAs	25957	0.9958183
96	2016-05-13	26827	http://linkedevents.org/ontology/poster	23781	0.8864577

	Time	DistinctEntity	Predicate	freq	NormalizedFreq
99	2016-05-13	26827	http://schema.org/aggregateRating	669	0.0249376
100	2016-05-13	26827	http://schema.org/description	319	0.0118910
101	2016-05-13	26827	http://schema.org/image	12491	0.4656130
102	2016-05-13	26827	http://schema.org/interactionCount	6944	0.2588437
103	2016-05-13	26827	http://schema.org/location	26770	0.9978753
104	2016-05-13	26827	http://schema.org/openingHours	3233	0.1205129
105	2016-05-13	26827	http://schema.org/priceRange	24	0.0008946
106	2016-05-13	26827	http://schema.org/review	4026	0.1500727
107	2016-05-13	26827	http://schema.org/telephone	21567	0.8039289
108	2016-05-13	26827	http://schema.org/url	5511	0.2054274
114	2016-05-13	26827	http://www.w3.org/2002/07/owl#sameAs	26719	0.9959742
119	2016-05-27	25828	http://linkedevents.org/ontology/poster	22904	0.8867895
122	2016-05-27	25828	http://schema.org/aggregateRating	646	0.0250116
123	2016-05-27	25828	http://schema.org/description	301	0.0116540
124	2016-05-27	25828	http://schema.org/image	12084	0.4678643
125	2016-05-27	25828	http://schema.org/interactionCount	6656	0.2577048
126	2016-05-27	25828	http://schema.org/location	25776	0.9979867
127	2016-05-27	25828	http://schema.org/openingHours	3107	0.1202958
128	2016-05-27	25828	http://schema.org/priceRange	22	0.0008518
129	2016-05-27	25828	http://schema.org/review	3862	0.1495276
130	2016-05-27	25828	http://schema.org/telephone	20767	0.8040499
131	2016-05-27	25828	http://schema.org/url	5308	0.2055134
137	2016-05-27	25828	http://www.w3.org/2002/07/owl#sameAs	25725	0.9960121
142	2016-06-15	41018	http://linkedevents.org/ontology/poster	37570	0.9159393
145	2016-06-15	41018	http://schema.org/aggregateRating	1280	0.0312058
146	2016-06-15	41018	http://schema.org/description	435	0.0106051

	Time	DistinctEntity	Predicate	freq	NormalizedFreq
147	2016-06-15	41018	http://schema.org/image	23470	0.5721878
148	2016-06-15	41018	http://schema.org/interactionCount	7607	0.1854552
149	2016-06-15	41018	http://schema.org/location	40903	0.9971964
150	2016-06-15	41018	http://schema.org/openingHours	6459	0.1574675
151	2016-06-15	41018	http://schema.org/priceRange	67	0.0016334
152	2016-06-15	41018	http://schema.org/review	9564	0.2331659
153	2016-06-15	41018	http://schema.org/telephone	33143	0.8080111
154	2016-06-15	41018	http://schema.org/url	9960	0.2428202
160	2016-06-15	41018	http://www.w3.org/2002/07/owl#sameAs	39743	0.9689161
162	2016-06-15	41018	http://data.ign.fr/def/topo#importance	66	0.0016090
163	2016-06-15	41018	http://vocab.org/transit/terms/route	2320	0.0565605
164	2016-06-15	41018	http://www.w3.org/2000/01/rdf-schema#seeAlso	122	0.0029743
168	2016-09-09	44968	http://linkedevents.org/ontology/poster	8323	0.1850872
171	2016-09-09	44968	http://schema.org/aggregateRating	1720	0.0382494
172	2016-09-09	44968	http://schema.org/description	600	0.0133428
173	2016-09-09	44968	http://schema.org/image	25290	0.5623999
174	2016-09-09	44968	http://schema.org/interactionCount	11844	0.2633873
175	2016-09-09	44968	http://schema.org/location	44692	0.9938623
176	2016-09-09	44968	http://schema.org/openingHours	6281	0.1396771
177	2016-09-09	44968	http://schema.org/priceRange	60	0.0013343
178	2016-09-09	44968	http://schema.org/review	10697	0.2378803
179	2016-09-09	44968	http://schema.org/telephone	35371	0.7865816
180	2016-09-09	44968	http://schema.org/url	9982	0.2219801
186	2016-09-09	44968	http://www.w3.org/2002/07/owl#sameAs	43538	0.9681996
188	2016-09-09	44968	http://data.ign.fr/def/topo#importance	67	0.0014899
189	2016-09-09	44968	http://vocab.org/transit/terms/route	2523	0.0561066

	Time	DistinctEntity	Predicate	freq	NormalizedFreq
190	2016-09-09	44968	http://www.w3.org/2000/01/rdf-schema#seeAlso	124	0.0027575

Completeness

Events Completeness:

Predicate	t20160909	t20160615	NormFreqDiff
http://linkedevents.org/ontology/atPlace	0.6153846	0.7479377	-0.1325531
http://linkedevents.org/ontology/atTime	0.7111756	0.9230064	-0.2118308
http://linkedevents.org/ontology/involvedAgent	0.0609579	0.1219065	-0.0609486
http://linkedevents.org/ontology/poster	0.5515239	1.0022915	-0.4507675
http://purl.org/dc/elements/1.1/issued	0.9927431	1.1342805	-0.1415374
http://purl.org/dc/elements/1.1/subject	1.4804064	1.7241063	-0.2436999
http://www.w3.org/2000/01/rdf-schema#label	0.9956459	1.0032081	-0.0075622
http://www.w3.org/2002/07/owl#sameAs	0.6371553	0.8143905	-0.1772352
NA	NA	NA	NA
NA	NA	NA	NA

Places Completeness:

Predicate	t20160909	t20160615	NormFreqDiff
http://data.linkedevents.org/def/location#businessType	2.7888721	2.8952167	-0.1063446
http://data.linkedevents.org/def/location#businessTypeTop	1.0357365	1.0871081	-0.0513716
http://linkedevents.org/ontology/poster	0.1850872	0.9159393	-0.7308522
http://schema.org/image	0.5623999	0.5721878	-0.0097879
http://schema.org/location	0.9938623	0.9971964	-0.0033341
http://schema.org/openingHours	0.1396771	0.1574675	-0.0177903
http://schema.org/priceRange	0.0013343	0.0016334	-0.0002991
http://schema.org/telephone	0.7865816	0.8080111	-0.0214295
http://schema.org/url	0.2219801	0.2428202	-0.0208402
http://www.w3.org/1999/02/22-rdf-syntax-ns#type	2.9716910	2.9719635	-0.0002725
http://www.w3.org/2002/07/owl#sameAs	0.9681996	0.9689161	-0.0007165
http://data.ign.fr/def/topo#importance	0.0014899	0.0016090	-0.0001191
http://vocab.org/transit/terms/route	0.0561066	0.0565605	-0.0004540
http://www.w3.org/2000/01/rdf-schema#seeAlso	0.0027575	0.0029743	-0.0002168
NA	NA	NA	NA
NA	NA	NA	NA