

Number Customer objects	2.4M	4.8M	9.6M	14.4M	19.2M	24M
Kryo data size	41.5GB	83.1GB	167.2GB	251.1GB	333.2	416.2GB
	Customers per Supplier					
PlinyCompute: hot storage	00:11	00:19	00:35	00:51	01:08	01:21
Spark: hot HDFS	01:04	01:53	03:24	04:54	06:25	08:16
Spark: in-RAM deserialized RDD	00:16	00:29	00:56	01:21	02:18	03:56
	top- $k$ Jaccard					
PlinyCompute: hot storage	00:03	00:03	00:04	00:05	00:05	00:06
Spark: hot HDFS	00:56	01:38	03:01	04:01	05:22	06:34
Spark: in-RAM deserialized RDD	00:08	00:12	00:21	00:32	01:11	02:38