package com.twitter.simclusters\_v2.scio

package multi\_type\_graph.multi\_type\_graph\_sims

import com.twitter.dal.client.dataset.SnapshotDALDataset

import com.twitter.simclusters\_v2.hdfs\_sources.RightNodeSimHashScioScalaDataset

import com.twitter.simclusters\_v2.thriftscala.RightNodeSimHashSketch

/\*\*

Build:

./bazel bundle src/scala/com/twitter/simclusters\_v2/scio/multi\_type\_graph/multi\_type\_graph\_sims:multi-type-graph-sim-hash-scio-adhoc-app

To kick off an adhoc run:

bin/d6w create \

${GCP\_PROJECT\_NAME}/us-central1/multi-type-graph-sim-hash-scio-adhoc-app \

src/scala/com/twitter/simclusters\_v2/scio/multi\_type\_graph/multi\_type\_graph\_sims/sim-hash-scio-adhoc.d6w \

--jar dist/multi-type-graph-sim-hash-scio-adhoc-app.jar \

--bind=profile.project=${GCP\_PROJECT\_NAME} \

--bind=profile.user\_name=${USER} \

--bind=profile.date="2021-12-01" \

--bind=profile.machine="n2d-highmem-16" --ignore-existing

\*/

object RightNodeSimHashScioAdhocApp extends RightNodeSimHashScioBaseApp {

override val isAdhoc: Boolean = true

override val rightNodeSimHashSnapshotDataset: SnapshotDALDataset[RightNodeSimHashSketch] =

RightNodeSimHashScioAdhocScalaDataset

}

/\*\*

To deploy the job:

bin/d6w schedule \

${GCP\_PROJECT\_NAME}/us-central1/multi-type-graph-sim-hash-scio-batch-app \

src/scala/com/twitter/simclusters\_v2/scio/multi\_type\_graph/multi\_type\_graph\_sims/sim-hash-scio-batch.d6w \

--bind=profile.project=${GCP\_PROJECT\_NAME} \

--bind=profile.user\_name=recos-platform \

--bind=profile.date="2021-12-01" \

--bind=profile.machine="n2d-highmem-16"

\*/

object RightNodeSimHashScioBatchApp extends RightNodeSimHashScioBaseApp {

override val isAdhoc: Boolean = false

override val rightNodeSimHashSnapshotDataset: SnapshotDALDataset[RightNodeSimHashSketch] =

RightNodeSimHashScioScalaDataset

}