package com.twitter.simclusters\_v2.scio.multi\_type\_graph.assemble\_multi\_type\_graph

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

object Config {

val RootMHPath: String = "manhattan\_sequence\_files/multi\_type\_graph/"

val RootThriftPath: String = "processed/multi\_type\_graph/"

val AdhocRootPath = "adhoc/multi\_type\_graph/"

val truncatedMultiTypeGraphMHOutputDir: String = "truncated\_graph\_mh"

val truncatedMultiTypeGraphThriftOutputDir: String = "truncated\_graph\_thrift"

val topKRightNounsMHOutputDir: String = "top\_k\_right\_nouns\_mh"

val topKRightNounsOutputDir: String = "top\_k\_right\_nouns"

val fullMultiTypeGraphThriftOutputDir: String = "full\_graph\_thrift"

val HalfLifeInDaysForFavScore = 100

val NumTopNounsForUnknownRightNodeType = 20

val GlobalDefaultMinFrequencyOfRightNodeType = 100

val TopKRightNounsForMHDump = 1000

// the topK most frequent nouns for each engagement type

val TopKConfig: Map[RightNodeType, Int] = Map(

RightNodeType.FollowUser -> 10000000, // 10M, current simclusters\_v2 has this value set to 20M, providing this the most weight

RightNodeType.FavUser -> 5000000,

RightNodeType.BlockUser -> 1000000,

RightNodeType.AbuseReportUser -> 1000000,

RightNodeType.SpamReportUser -> 1000000,

RightNodeType.FollowTopic -> 5000,

RightNodeType.SignUpCountry -> 200,

RightNodeType.ConsumedLanguage -> 50,

RightNodeType.FavTweet -> 500000,

RightNodeType.ReplyTweet -> 500000,

RightNodeType.RetweetTweet -> 500000,

RightNodeType.NotifOpenOrClickTweet -> 500000,

RightNodeType.SearchQuery -> 500000

)

val SampledEmployeeIds: Set[Long] =

Set()

}