create\_datarecord\_datasets(

base\_name = "earlybird\_recap\_data\_records",

platform = "java8",

role = "timelines",

segment\_type = "partitioned",

tags = [

"bazel-compatible",

"bazel-compatible:migrated",

],

)

create\_datarecord\_datasets(

base\_name = "earlybird\_rectweet\_data\_records",

platform = "java8",

role = "timelines",

segment\_type = "partitioned",

tags = [

"bazel-compatible",

"bazel-compatible:migrated",

],

)

scala\_library(

name = "training\_data\_generation",

sources = ["\*.scala"],

platform = "java8",

strict\_deps = True,

tags = [

"bazel-compatible",

"bazel-compatible:migrated",

],

dependencies = [

":earlybird\_recap\_data\_records-java",

":earlybird\_rectweet\_data\_records-java",

"3rdparty/jvm/com/ibm/icu:icu4j",

"3rdparty/src/jvm/com/twitter/scalding:json",

"src/java/com/twitter/ml/api:api-base",

"src/java/com/twitter/ml/api/constant",

"src/java/com/twitter/ml/api/matcher",

"src/java/com/twitter/search/common/features",

"src/scala/com/twitter/ml/api:api-base",

"src/scala/com/twitter/ml/api/analytics",

"src/scala/com/twitter/ml/api/util",

"src/scala/com/twitter/scalding\_internal/dalv2",

"src/scala/com/twitter/scalding\_internal/dalv2/dataset",

"src/scala/com/twitter/scalding\_internal/job",

"src/scala/com/twitter/scalding\_internal/job/analytics\_batch",

"src/scala/com/twitter/timelines/prediction/features/common",

"src/scala/com/twitter/timelines/prediction/features/recap",

"src/thrift/com/twitter/ml/api:data-java",

"src/thrift/com/twitter/ml/api:dataset-analytics-java",

"timelines/data\_processing/ad\_hoc/earlybird\_ranking/common",

"timelines/data\_processing/ad\_hoc/recap/dataset\_utils",

"timelines/data\_processing/ad\_hoc/recap/offline\_execution",

"timelines/data\_processing/util/execution",

],

)

hadoop\_binary(

name = "bin",

basename = "earlybird\_training\_data\_generation-deploy",

main = "com.twitter.scalding.Tool",

platform = "java8",

runtime\_platform = "java8",

tags = [

"bazel-compatible",

"bazel-compatible:migrated",

"bazel-only",

],

dependencies = [

":training\_data\_generation",

],

)

hadoop\_binary(

name = "earlybird\_training\_data\_generation\_prod",

basename = "earlybird\_training\_data\_generation\_prod-deploy",

main = "com.twitter.timelines.data\_processing.ad\_hoc.earlybird\_ranking.training\_data\_generation.EarlybirdTrainingDataProdJob",

platform = "java8",

runtime\_platform = "java8",

tags = [

"bazel-compatible",

"bazel-compatible:migrated",

"bazel-only",

],

dependencies = [

":training\_data\_generation",

],

)