**Pig commands:**

**$pig -x mapreduce**

**Usecase1:**

Airports = load '/user/edureka\_445719/sangeethadata/Airdataset/airports\_mod.dat' using PigStorage (',') as (airport\_id:int,name:chararray,city: chararray, country: chararray,IATA: chararray,ICAO:chararray,latitude: chararray,longitude: chararray,altitude:int,timezone:int, DST: chararray);

Country = foreach Airports generate name,country;

Filitered = Filter Country by country == 'India';

Airportsi = foreach Filitered generate name;

Dump Airortsi;

store Airportsi into 'sangeethadata/Airdataset/usecase1';

**usecase2:**

Airlines = load '/user/edureka\_445719/sangeethadata/Airdataset/Final\_airlines'using PigStorage (',') as (airline\_id:int,name:chararray,alias: chararray,IATA: chararray,ICAO: chararray,country: chararray);

Routes= load '/user/edureka\_445719/sangeethadata/Airdataset/routes.dat' using PigStorage (',') as (airline:chararray,airline\_id:int,Source\_airport: chararray,Source\_airport\_ID:int, Destination\_airport: chararray,Destination\_airport\_ID:int,Codeshare: chararray,Stops:int,Equipment: chararray);

FR = foreach Routes generate airline\_id,Stops== 0;

describe FS;

ARJ = join FA by airline\_id ,FR by airline\_id;

describe ARJ;

grp =group ARJ by FA::name;

describe grp;

FF = foreach grp generate group;

describe FF;

dump FF;

store FF into 'sangeethadata/Airdataset/usecase2';

**Usecase3:**

Airlines = load '/user/edureka\_445719/sangeethadata/Airdataset/Final\_airlines'using PigStorage (',') as (airline\_id:int,name:chararray,alias: chararray,IATA: chararray,ICAO: chararray,country: chararray);

Routes= load '/user/edureka\_445719/sangeethadata/Airdataset/routes.dat' using PigStorage (',') as (airline:chararray,airline\_id:int,Source\_airport:chararray,Source\_airport\_ID:int, Destination\_airport: chararray,Destination\_airport\_ID:int,Codeshare: chararray,Stops:int,Equipment: chararray);

FA = Foreach Airlines generate airline\_id,name;

FR = foreach Routes generate airline\_id,Codeshare;

FFR = filter FR by Codeshare == 'Y';

describe FFR;

ARJ = join FA by airline\_id,FFR by airline\_id;

describe ARJ;

DARJ = Distinct ARJ;

G = group DARJ by (name,Codeshare);

describe G;

grouped = foreach G generate flatten(group) as (name,Codeshare);

store grouped in 'sangeethadata/Airdataset/usecase3';

Usecase 4:

Airports = load '/user/edureka\_445719/sangeethadata/Airdataset/airports\_mod.dat' using PigStorage (',') as (airport\_id:int,name:chararray,city: chararray, country: chararray,IATA: chararray,ICAO:chararray,latitude: chararray,longitude: chararray,altitude:int,timezone:int, DST: chararray);

FA = foreach Airports generate name,country;

describe FA;

grouped = GROUP FA BY country;

describe grouped;

FG = foreach grouped generate group,COUNT(FA.name) as airports\_count;

describe FG;

sort = order FG by airports\_count desc;

FC = limit sort by 1;

describe FC;

dump FC;

store FC in 'sangeethadata/Airdataset/usecase4'

**Usecase:5**

Airlines = load '/user/edureka\_445719/sangeethadata/Airdataset/Final\_airlines'using PigStorage (',') as (airline\_id:int,name:chararray,alias: chararray,IATA: chararray,ICAO: chararray,country: chararray,active:chararray);

describe Airlines;

FC = FOREACH Airlines generate name,active,country;

FL = FILTER FC by country == 'United States' AND active == 'Y';

describe FL;

dump FL;

STORE FL into 'sangeethadata/Airdataset/usecase6';