1.Write a pig script to find the top 20 destination people travel the most:

TravelData = load '/user/acadgild/Desktop/TravelData.txt' using PigStorage('\t') as (CityPair:chararray, fromLocation:chararray, toLocation:chararray, Producttype:int, Adults:int, Seniors:int, Children:int, Youth:int, Infant:int, DateofTravel:chararray, TimeofTravel:chararray, DateofReturn:chararray, TimeofReturn:int, Price:float, Hotel:chararray);

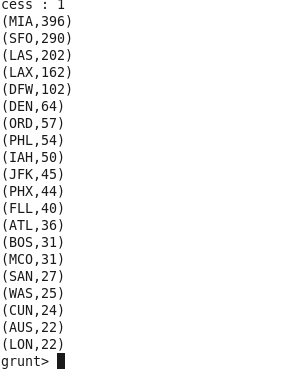
group\_destination = group TravelData by toLocation;

count\_destination = foreach group\_destination generate group, COUNT(TravelData.toLocation)as cnt;

orderby = order count\_destination by cnt DESC;

top\_destination = limit orderby 20;

dump top\_destination;



2. Write a pig script to find the top 20 locations from where people travel the most

TravelData = load '/user/acadgild/Desktop/TravelData.txt' using PigStorage('\t') as (CityPair:chararray, fromLocation:chararray, toLocation:chararray, Producttype:int, Adults:int, Seniors:int, Children:int, Youth:int, Infant:int, DateofTravel:chararray, TimeofTravel:chararray, DateofReturn:chararray, TimeofReturn:int, Price:float, Hotel:chararray);

group\_source = group TravelData by fromLocation;

count\_source = foreach group\_source generate group, COUNT(TravelData.fromLocation)as cnt;

orderby = order count\_source by cnt DESC;

top\_source = limit orderby 20;

dump top\_source;

