

# CASESTUDY: AIRPORTS ANALYTICS Analysis with PIG

19CSE357 | Group 11

BIGDATA ANALYTICS

A large yellow triangle is positioned in the bottom right corner of the slide, pointing towards the top right.

RAVELLA ABHINAV



CB.EN.U4CSE19453

# 1. Find the names of active airlines in territory 'Unites States'?

```
2022-04-05 06:43:18,681 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapreduceLayer.MapReduceLauncher
2022-04-05 06:43:18,689 [main] WARN org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been in
2022-04-05 06:43:18,725 [main] INFO org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input files to pr
2022-04-05 06:43:18,725 [main] INFO org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input pat
(135 Airways)
(3D Aviation)
(40-Mile Air)
(Aloha Airlines)
(Alaska Island Air)
(American Airlines)
(Aviation Management Corporation)
(Atlantis Airlines (USA))
(Astro Air International)
(Allegiant Air)
(Airborne Express)
(ABX Air)
(Academy Airlines)
(Air Aurora)
(Advance Leasing Company)
(Aztec Worldwide Airlines)
(Aerodyne)
(Anderson Aviation)
(Alaska Central Express)
(Airway Express)
(Aerial Transit Company)
(Air Cson)
(Air Cargo America)
(Air Cargo Carriers)
(Atlantic Gulf Airlines)
(Atlantic Coast Airlines)
(Aero Flight Service)
(Aspen Helicopters)
(Alpine Air Express)
(Airlift International)
(Air Illinois)
(Ameristar Jet Charter)
(A2 Jet Leasing)
(AmeriJet International)
(Alpha Jet International)
(American Flyers)
(Aero Coach Aviation)
(Atlantic Southeast Airlines)
(Air Logistics)
(Air Lincoln)
(America West Airlines)
```

## Query:

- `active_airlines = filter finalair by active == 'Y';`
- `active_airlines_usa = filter active_airlines by territory == 'United States';`
- `find_active_airline_names_in_usa = foreach active_airlines_usa generate airline_name;`
- `dump find_active_airline_names_in_usa;`

## 2. Group airline id by active status;

### Query:

- ae = group finalair by airlineID;
- ans = limit ae 10;
- illustrate as;
- dump as;

```
(1,{{1,Private Flight,\N,-,N/A,,,Y}})
(2,{{2,135 Airways,\N,,GNL,GENERAL,United States,N}})
(3,{{3,1Time Airline,\N,IT,RNX,NEXTIME,South Africa,Y}})
(4,{{4,2 Sqn No 1 Elementary Flying Training School,\N,,WYT,,United Kingdom,N}})
(5,{{5,213 Flight Unit,\N,,TFU,,Russia,N}})
(6,{{6,223 Flight Unit State Airline,\N,,CHD,CHKALOVSK-AVIA,Russia,N}})
(7,{{7,224th Flight Unit,\N,,TTF,CARGO UNIT,Russia,N}})
(8,{{8,247 Jet Ltd,\N,,TWF,CLOUD RUNNER,United Kingdom,N}})
(9,{{9,3D Aviation,\N,,SEC,SECUREX,United States,N}})
(-1,{{-1,Unknown,\N,-,N/A,\N,\N,Y}})
grunt>
```

### 3. Split relation based on src airport id i.e, equal to DME to get exclusive info.

```
(2B,410,DME,4029,KZN,2990,)  
(2B,410,DME,4029,NBC,6969,)  
(2B,410,DME,4029,TGK,\N,)  
(2B,410,DME,4029,UUA,6160,)  
(3R,3545,DME,4029,ASF,2966,)  
(3R,3545,DME,4029,BHK,2980,)  
(3R,3545,DME,4029,EGO,6156,)  
(3R,3545,DME,4029,FEG,6149,)  
(3R,3545,DME,4029,KSQ,6153,)  
(3R,3545,DME,4029,NMA,6150,)  
grunt>
```

#### Query:

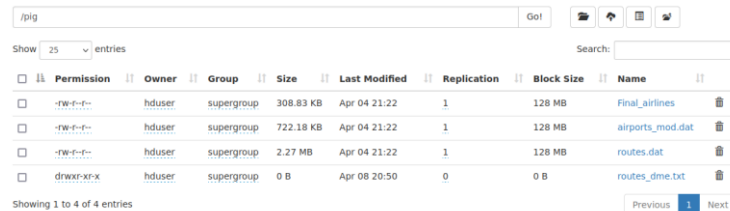
- split routes into ru\_dme if route\_source\_iata == 'DME', ru\_not\_dme if route\_source\_iata != 'DME';
- dump ru\_dme;
- dump ru\_not\_dme;

# 4. Store ru\_dme relation into text file separated by delimiter “,”.

Query:

- store ru\_dme into '/pig/routes\_dme.txt' using PigStorage(',');

Output:



Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name
-rw-r--r--	hduser	supergroup	308.83 KB	Apr 04 21:22	1	128 MB	<a href="#">final_airlines</a>
-rw-r--r--	hduser	supergroup	722.18 KB	Apr 04 21:22	1	128 MB	<a href="#">airports_mod.dat</a>
-rw-r--r--	hduser	supergroup	2.27 MB	Apr 04 21:22	1	128 MB	<a href="#">routes.dat</a>
drwxr-xr-x	hduser	supergroup	0 B	Apr 08 20:50	0	0 B	<a href="#">routes_dme.txt</a>

Showing 1 to 4 of 4 entries

# 5. JOIN Command:

```
:
ully read 67663 records from: "/pig/routes.dat"
ully read 6048 records from: "/pig/Final_airlines"
):
ully stored 10 records (150056224 bytes) in: "hdfs://localhost:54310/tmp/temp-130244988/tmp1067114587"
:
ords written : 10
tes written : 150056224
e Memory Manager spill count : 0
gs proactively spilled: 0
ords proactively spilled: 0

l1956056652_0009    ->    job_local635458340_0010,
l635458340_0010

05 10:10:40,531 [main] WARN org.apache.hadoop.metrics2.impl.MetricsSystemImpl - JobTracker metrics system alrea
05 10:10:40,532 [main] WARN org.apache.hadoop.metrics2.impl.MetricsSystemImpl - JobTracker metrics system alrea
05 10:10:40,534 [main] WARN org.apache.hadoop.metrics2.impl.MetricsSystemImpl - JobTracker metrics system alrea
05 10:10:40,557 [main] WARN org.apache.hadoop.metrics2.impl.MetricsSystemImpl - JobTracker metrics system alrea
05 10:10:40,572 [main] WARN org.apache.hadoop.metrics2.impl.MetricsSystemImpl - JobTracker metrics system alrea
05 10:10:40,575 [main] WARN org.apache.hadoop.metrics2.impl.MetricsSystemImpl - JobTracker metrics system alrea
05 10:10:40,586 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - Su
05 10:10:40,598 [main] WARN org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initia
05 10:10:40,602 [main] INFO org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input files to proces
05 10:10:40,602 [main] INFO org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths t
ocondor, \N,28,ARD,AEROCONDOR,Portugal,Y,28,410,DME,4029,KZN,2990,)
ocondor, \N,28,ARD,AEROCONDOR,Portugal,Y,28,410,DME,4029,TGK,\N,)
ocondor, \N,28,ARD,AEROCONDOR,Portugal,Y,28,410,DME,4029,UJA,6160,)
ocondor, \N,28,ARD,AEROCONDOR,Portugal,Y,28,410,EGD,6156,KGD,2952,)
ocondor, \N,28,ARD,AEROCONDOR,Portugal,Y,28,410,EGD,6156,KZN,2990,)
ocondor, \N,28,ARD,AEROCONDOR,Portugal,Y,28,410,GYD,2922,NBC,6969,)
ocondor, \N,28,ARD,AEROCONDOR,Portugal,Y,28,410,KGD,2952,EGO,6156,)
ocondor, \N,28,ARD,AEROCONDOR,Portugal,Y,28,410,KZN,2990,AER,2965,)
ocondor, \N,28,ARD,AEROCONDOR,Portugal,Y,28,410,KZN,2990,ASF,2966,)
ocondor, \N,28,ARD,AEROCONDOR,Portugal,Y,28,410,KZN,2990,EFY,2968,)
```

## Query:

- right\_join = JOIN finalair BY airline\_iata RIGHT OUTER, routes BY route\_iata;
- LIM = LIMIT right\_join 10;
- dump LIM;

## Output:

S SHANTHAN REDDY



CB.EN.U4CSE19459



```
2022-04-05 08:45:11,606 [main] INFO org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total
2022-04-05 08:45:11,606 [main] INFO org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil
(163,Arab Agricultural Aviation Company,\N,,AGC,AGRICO,Egypt,N)
(317,AMC Airlines,\N,,AMV,,Egypt,Y)
(442,Air Sinai,\N,4D,ASD,AIR SINAI,Egypt,Y)
(610,Air Memphis,\N,,MHS,AIR MEMPHIS,Egypt,N)
(992,Alexandria Airlines,\N,,KHH,,Egypt,N)
(1015,Aleem,\N,,LEM,,Egypt,N)
(1020,Al Ahran Aviation,\N,,LHR,AL AHRAM,Egypt,N)
(1094,Air Cairo,\N,,MSC,,Egypt,N)
(1146,Al Farana Airline,\N,,PHR,PHARAOH,Egypt,N)
(1603,Cairo Air Transport Company,\N,,CCE,,Egypt,N)
(2143,Egyptair,\N,MS,MSR,EGYPTAIR,Egypt,Y)
(2144,Egyptair Cargo,\N,,MSX,EGYPTAIR CARGO,Egypt,N)
(2145,Egyptian Air Force,\N,,EGY,,Egypt,N)
(2146,Egyptian Aviation,\N,,EJX,,Egypt,N)
(2147,Egyptian Aviation Company,\N,,EMA,,Egypt,N)
(2369,Flash Airlines,\N,,FSH,FLASH,Egypt,N)
(2897,International Air Cargo Corporation,\N,,IAK,AIR CARGO EGYPT,Egypt,N)
(3314,Lotus Air,\N,,TAS,LOTUS FLOWER,Egypt,N)
(3333,Luxor Air,\N,,LXO,,Egypt,N)
(3449,Mediterranean Airways,\N,,MDY,,Egypt,N)
(3498,Midwest Airlines (Egypt),\N,MY,MWA,,Egypt,Y)
(3508,Mina Airline Company,\N,,NAB,,Egypt,N)
(3518,Misir Overseas Airways,\N,,MOS,,Egypt,N)
(3602,National Aviation Company,\N,,GTU,,Egypt,N)
(3610,National Overseas Airlines Company,\N,,NOL,NAT AIRLINE,Egypt,N)
(3664,Nile Valley Aviation Company,\N,,NVA,,Egypt,N)
(3802,Orca Air,\N,,ORK,ORCA TAXI,Egypt,N)
(3819,Orscom Tourist Installations Company,\N,,OAD,ORSCOM,Egypt,N)
(3876,Palestinian Airlines,\N,PF,PNW,PALESTINIAN,Egypt,N)
(4087,Pyramid Air Lines,\N,,PYR,PYAIR,Egypt,N)
(4141,Raslan Air Service,\N,,MWR,RASLAN,Egypt,N)
(4159,Red Sea Aviation,\N,,RDV,RED AVIATION,Egypt,N)
(4351,Scorpio,\N,,SCP,SCORPIO,Egypt,N)
(4419,Shorouk Air,\N,,SHK,,Egypt,N)
(4444,Saber Airlines,\N,,SIS,,Egypt,N)
(4528,Sun Air,\N,,SUR,,Egypt,N)
(4884,Tropicair Air Services,\N,,TCA,TROPICANA,Egypt,N)
(4982,Transmed Airlines,\N,,TMD,,Egypt,N)
(5075,Tristar Air,\N,,TSY,TRIPLE STAR,Egypt,N)
(5512,ZAS Airlines of Egypt,\N,,ZAS,ZAS AIRLINES,Egypt,N)
(11811,AlMasria Universal Airlines,,UJ,LMU,ALMASRIA,Egypt,Y)
(11813,EgyptAir Express,,MSE,EGYPTAIR EXPRESS,Egypt,Y)
(11815,Air Italy Egypt,,EUD,,Egypt,Y)
(11816,KoralBlue Airlines,,K7,KBR,KORAL BLUE,Egypt,Y)
(16628,Jusur airways,,JX,JSR,,Egypt,Y)
(16963,Air Arabia Egypt,,E5,RGB,,Egypt,Y)
```

1. `egypt = filter finalair by  
territory == 'Egypt';  
dump egypt;`

```
2022-04-05 08:49:51,998 [main] INFO org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not generate code.
2022-04-05 08:49:52,018 [main] WARN org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2022-04-05 08:49:52,020 [main] INFO org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2022-04-05 08:49:52,027 [main] INFO org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input files to process : 1
2022-04-05 08:49:52,027 [main] INFO org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to process : 1
2022-04-05 08:49:52,543 [main] INFO org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task 'attempt__0001_m_000001_1' to hdfs://localhost:54310/tmp/temp-368850805/tmp-61223068
2022-04-05 08:49:52,592 [main] WARN org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2022-04-05 08:49:52,635 [main] INFO org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input files to process : 1
2022-04-05 08:49:52,636 [main] INFO org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to process : 1
(PRIVATE FLIGHT,Y)
(135 AIRWAYS,N)
(1TIME AIRLINE,Y)
grunt> █
```

2. up = foreach finalair generate  
UPPER(airline\_name),active;  
a = limit up 3;  
dump a;

hduser@shanthan-VirtualBox: ~

```
(19599,Skyline Ulasim Ticaret A.S.,Skyline Ulasim Ticaret A.S.,KCU,Kocoglu,Turkey,Y,,,,,,,,)
(19619,Envoy Air,,ENV,Envoy,United States,Y,,,,,,,,)
(19651,CARICOM AIRWAYS (BARBADOS) INC.,CARICOM AIRWAYS,,CCB,,Barbados,Y,,,,,,,,)
(65,Airservices Australia,\N,,ADA,AUSCAL,Australia,N,,,,,,,,)
(64,Atlas Cargo Lines,\N,,ACY,ATLAS CARGOLINES,Morocco,N,,,,,,,,)
(63,Air Somalia,\N,,RSM,AIR SOMALIA,Somali Republic,N,,,,,,,,)
(61,Air Charters,\N,,ACX,PARAIR,Canada,N,,,,,,,,)
(59,Air Charter Service,\N,,ACV,,United Kingdom,N,,,,,,,,)
(58,Air Cargo Transportation System,\N,,ACU,AFRISPIRIT,Kenya,N,,,,,,,,)
(57,Air Aurora,\N,,AAI,BOREALIS,United States,N,,,,,,,,)
(56,Air Cess,\N,,ACS,,Liberia,N,,,,,,,,)
(54,Aero Comondu,\N,,ACO,AERO COMONDU,Mexico,N,,,,,,,,)
(53,Academy Airlines,\N,,ACD,ACADEMY,United States,N,,,,,,,,)
(52,Avcard Services,\N,,ACC,,United Kingdom,N,,,,,,,,)
(51,ATA Brasil,\N,,ABZ,ATA-BRAZIL,Brazil,N,,,,,,,,)
(47,Aerocenter,\N,,ACR, Escuela de Formacion de Pilotos Privados de Avion,AEROCENTER,N,,,,,,,,)
(46,Aerovias Bueno,\N,,ABU,AEROBUENO,Colombia,N,,,,,,,,)
(45,APSA Colombia,\N,,ABO,AEROEXPRESO,Colombia,N,,,,,,,,)
(44,Alberta Citylink,\N,,ABK,ALBERTA CITYLINK,Canada,N,,,,,,,,)
(43,Abaet,\N,,ABJ,Abaet,Brazil,N,,,,,,,,)
(41,Abakan-Avia,\N,,ABG,ABAKAN-AVIA,Russia,N,,,,,,,,)
(40,Aerial Oy,\N,,ABF,SKYWINGS,Finland,N,,,,,,,,)
(19745,Transilvania,,TNS,,Romania,Y,,,,,,,,)
(38,African Business and Transportations,\N,,ABB,AFRICAN BUSINESS-,Democratic Republic of the Congo,N,,,,,,,,)
(37,Artem-Avia,\N,,ABA,ARTEM-AVIA,Ukraine,N,,,,,,,,)
(36,Angus Aviation,\N,,AAZ,ANGUS,Canada,N,,,,,,,,)
(34,Afric'air Express,\N,,AAX,AFREX,Ivory Coast,N,,,,,,,,)
(31,Astro Air International,\N,,AAV,ASTRO-PHIL,United States,N,,,,,,,,)
(30,Australia Asia Airlines,\N,,AAU,AUSTASIA,Australia,N,,,,,,,,)
(19776,Grand Cru Airlines,,GCA,,Lithuania,Y,,,,,,,,)
(27,Aerovista Airlines,\N,,AAP,AEROVISTA GROUP,United Arab Emirates,N,,,,,,,,)
(26,Atlantis Airlines (USA),\N,,AAO,ATLANTIS AIR,United States,N,,,,,,,,)
(25,Aviation Management Corporation,\N,,AAM,AM CORP,United States,N,,,,,,,,)
(19785,Go2Sky,,RLX,RELAX,Slovakia,Y,,,,,,,,)
(23,Alaska Island Air,\N,,AAK,ALASKA ISLAND,United States,N,,,,,,,,)
(20,Aero Albatros,\N,,ABM,ALBATROS ESPANA,Spain,N,,,,,,,,)
(19,Aero Biniza,\N,,BZS,BINIZA,Mexico,N,,,,,,,,)
(18,Aero Servicios Ejecutivos Internacionales,\N,,SII,ASEISA,Mexico,N,,,,,,,,)
(17,Aero Aviation Centre Ltd.,\N,,AAD,SUNRISE,Canada,N,,,,,,,,)
(16,Army Air Corps,\N,,AAC,ARMYAIR,United Kingdom,N,,,,,,,,)
(12,611897 Alberta Limited,\N,,THD,DONUT,Canada,N,,,,,,,,)
(11,4D Air,\N,,QRT,QUARTET,Thailand,N,,,,,,,,)
(9,3D Aviation,\N,,SEC,SECUREX,United States,N,,,,,,,,)
(8,247 Jet Ltd,\N,,TWF,CLOUD RUNNER,United Kingdom,N,,,,,,,,)
(7,224th Flight Unit,\N,,TTF,CARGO UNIT,Russia,N,,,,,,,,)
(6,223 Flight Unit State Airline,\N,,CHD,CHKALOVSK-AVIA,Russia,N,,,,,,,,)
(5,213 Flight Unit,\N,,TFU,,Russia,N,,,,,,,,)
(4,2 Sqn No 1 Elementary Flying Training School,\N,,WYT,,United Kingdom,N,,,,,,,,)
```

3. outer\_left = JOIN  
finalair BY  
airline\_iata LEFT  
OUTER, routes BY  
route\_iata;  
dump outer\_left;

```
(19358,Caribbean Wings)
(19359,FLYJET)
(19361,Snowbird Airlines)
(19363,Russkie Krylya)
(19367,Kharkiv Airlines)
(19423,Key Air)
(19433,XAIR USA)
(19451,Air Costa)
(19459,Simrik Airlines)
(19465,Global Freightways)
(19473,XPT0)
(19474,Royal Flight)
(19525,BBN-Airways)
(19531,Tomsk-Avia)
(19538,Vintage Props and Jets)
(19541,Malawian Airlines)
(19548,Yeti Airlines )
(19567,Avilu)
(19599,Skyline Ulasim Ticaret A.S.)
(19619,Envoy Air)
(19651,CARICOM AIRWAYS (BARBADOS) INC.)
(19674,Rainbow Air (RAI))
(19675,Rainbow Air Canada)
(19676,Rainbow Air Polynesia)
(19677,Rainbow Air Euro)
(19678,Rainbow Air US)
(19745,Transilvania)
(19751,Dobrolet)
(19774,Spike Airlines)
(19776,Grand Cru Airlines)
(19785,Go2Sky)
(19803,All Argentina)
(19804,All America)
(19805,All Asia)
(19806,All Africa)
(19807,Regionalia México)
(19808,All Europe)
(19809,All Spain)
(19810,Regional Air Iceland)
(19811,British Air Ferries)
(19812,Voestar)
(19813,All Colombia)
(19814,Regionalia Uruguay)
(19815,Regionalia Venezuela)
(19827,Regionalia Chile)
(19828,Vuela Cuba)
(19830,All Australia)
```

```
4. order_data = ORDER
    finalair by airlineID
    DESC;
    Disp=FOREACH finalair
    GENERATE
    airlineID,airline_name;
    Dump disp;
```

```
hduser@shanthan-VirtualBox: ~  
(Turks and Caicos Islands,Caribbean Wings)  
(Poland,FLYJET)  
(Finland,Snowbird Airlines)  
(Russia,Russkie Krylya)  
(Ukraine,Kharkiv Airlines)  
(United States,Key Air)  
(United States,XAIR USA)  
(India,Air Costa)  
(Nepal,Simrik Airlines)  
(United States,Global Freightways)  
(Portugal,XPT0)  
(Russia,Royal Flight)  
(United Kingdom,BBN-Airways)  
(Russia,Tomsk-Avia)  
(United States,Vintage Props and Jets)  
(Malawi,Malawian Airlines)  
(Nepal,Yeti Airlines )  
(Switzerland,Avilu)  
(Turkey,Skyline Ulasim Ticaret A.S.)  
(United States,Envoy Air)  
(Barbados,CARICOM AIRWAYS (BARBADOS) INC.)  
(United States,Rainbow Air (RAI))  
(Canada,Rainbow Air Canada)  
(United States,Rainbow Air Polynesia)  
(United Kingdom,Rainbow Air Euro)  
(United States,Rainbow Air US)  
(Romania,Transilvania)  
(Russia,Dobrolet)  
(United States,Spike Airlines)  
(Lithuania,Grand Cru Airlines)  
(Slovakia,Go2Sky)  
(Argentina,All Argentina)  
(United States,All America)  
(China,All Asia)  
(South Africa,All Africa)  
(Mexico,Regionalia México)  
(United Kingdom,All Europe)  
(Spain,All Spain)  
(Iceland,Regional Air Iceland)  
(United Kingdom,British Air Ferries)  
(Brazil,Voestar)  
(Colombia,All Colombia)  
(Uruguay,Regionalia Uruguay)  
(Venezuela,Regionalia Venezuela)  
(Chile,Regionalia Chile)  
(Cuba,Vuela Cuba)  
(Australia,All Australia)  
(Spain,Fly Europa)  
(Portugal,Fly Europa)
```

CB.EN.U4CSE19459

```
5. group_data = GROUP  
   finalair by territory;  
   display = FOREACH  
   finalair GENERATE  
   territory,airline_name;  
   Dump display;
```

P SAI KOUSHIK



CB.EN.U4CSE19449



```
(5037,Transportacion Aerea De Queretaro,\N,,TQR,TRANSQUERETARO,Mexico,N)
(5081,Transportaciones Y Servicios Aereos,\N,,TTR,TRANSPORTACIONES,Mexico,N)
(5082,Transporte Aereo Tecnico Ejecutivo,\N,,TTS,TECNICO,Mexico,N)
(5088,Taxi Aereo Turistico,\N,,TUO,TURISTICO,Mexico,N)
(5108,Taxi Aereo Cozatl,\N,,TXL,TAXI COZATL,Mexico,N)
(5109,Taxi Aereo de Mexico,\N,,TXM,TAXIMEX,Mexico,N)
(5111,Taxis Aereos de Sinaloa,\N,,TXO,TAXIS SINALOA,Mexico,N)
(5112,Taxirey,\N,,TXR,TAXIREY,Mexico,N)
(5119,Transporte Aereo Ernesto Saenz,\N,,TZE,TRANSPORTE SAENZ,Mexico,N)
(5135,Transporte Ejecutivo Aereo,\N,,EAR,EJECUTIVO-AEREO,Mexico,N)
(5162,Transportacion Aerea Del Mar De Cortes,\N,,MCT,TRANS CORTES,Mexico,N)
(5163,Transporte Aero MGM,\N,,MGM,AERO EMM-GEE-EMM,Mexico,N)
(5165,Transportes Aereos Amparo,\N,,MPO,AMPARO,Mexico,N)
(5167,Transportes Aereos Mexiquenses,\N,,MXQ,MEXIQUENSES,Mexico,N)
(5181,Transportes Aereos Sierra,\N,,RRT,SIERRA ALTA,Mexico,N)
(5185,Transportes Aereos Sierra Madre,\N,,SEI,TRANSPORTE SIERRA,Mexico,N)
(5196,Taxi de Veracruz,\N,,VRC,VERACRUZ,Mexico,N)
(5198,Taxi Aero Del Norte,\N,,XNR,TAXI NORTE,Mexico,N)
(5247,Universal Jet Rental de Mexico,\N,,UJR,UNIVERSAL JET,Mexico,N)
(5278,Uvavemex,\N,,UVM,UVAVEMEX,Mexico,N)
(5299,Vanguardia en Aviacion en Colima,\N,,VGC,VANGUARDIA COLIMA,Mexico,N)
(5305,Verataxis,\N,,VTX,VERATAXIS,Mexico,N)
(5325,Volaris,\N,Y4,VOI,VOLARIS,Mexico,Y)
(5334,VZ Flights,\N,,VFT,ZETA FLIGHTS,Mexico,N)
(5340,VIP Servicios Aereos Ejecutivos,\N,,VIC,VIP-EJECUTIVO,Mexico,N)
(5341,VIP Empresarial,\N,,VIE,VIP EMPRESARIAL,Mexico,N)
(5348,Viajes Ejecutivos Mexicanos,\N,,VJM,VIAJES MEXICANOS,Mexico,N)
(5375,Vuelos Especializados Tollocan,\N,,VTC,VUELOS TOLLOCAN,Mexico,N)
(5376,Vuelos Corporativos de Tehuacan,\N,,VTH,VUELOS TEHUACAN,Mexico,N)
(5381,Vuela Bus,\N,,VUS,VUELA BUS,Mexico,N)
(5481,Xabre Aerolineas,\N,,XAB,AERO XABRE,Mexico,N)
(10715,Click (Mexicana),,QA,\N,,Mexico,Y)
(10955,MexicanaLink,,I6,MXI,LINK,Mexico,Y)
(16724,Red Jet Mexico,,4X,\N,,Mexico,Y)
(17885,Interjet (ABC Aerolineas),,40,\N,INTERJET,Mexico,Y)
(19807,Regionalia México,Regionalia México,N4,J88,,Mexico,Y)
grunt> █
```

1. SPLIT finalair into finalair1 if  
territory=='Mexico', finalair2 if  
territory=='Canada';

dump finalair1;

```
(5161,Trans Atlantis,\N,,LTA,LANTRA,Canada,N)
(5292,Vacationair,\N,,VAC,VACATIONAIR,Canada,N)
(5307,Voyageur Airways,\N,VC,VAL,VOYAGEUR,Canada,N)
(5313,VIA Rail Canada,\N,2R,,,Canada,N)
(5349,Vistajet,\N,,VJT,VISTA,Canada,N)
(5371,Vision Airways Corporation,\N,,VSN,VISION,Canada,N)
(5391,Walsten Air Services,\N,,WAS,WALSTEN,Canada,N)
(5393,Wapiti Aviation,\N,,WPT,WAPITI,Canada,N)
(5396,Wasaya Airways,\N,,WSG,WASAYA,Canada,N)
(5411,West Coast Air,\N,80,,,Canada,Y)
(5415,West Wind Aviation,\N,,WEW,WESTWIND,Canada,N)
(5416,WestJet,\N,WS,WJA,WESTJET,Canada,Y)
(5420,Westcoast Energy,\N,,BLK,BLUE FLAME,Canada,N)
(5425,Western Arctic Air,\N,,WAL,WESTERN ARCTIC,Canada,N)
(5427,Western Express Air Lines,\N,,WES,WEST EX,Canada,N)
(5433,Westpoint Air,\N,,WTP,WESTPOINT,Canada,N)
(5437,White River Air Services,\N,,WRA,,Canada,N)
(5466,World Weatherwatch,\N,,XWW,,Canada,N)
(5469,Worldwide Air Charter Systems,\N,,CSW,CHARTER SYSTEMS,Canada,N)
(5521,Zip,\N,3J,WZP,ZIPPER,Canada,N)
(5523,Zoom Airlines,\N,Z4,OOM,ZOOM,Canada,Y)
(9828,Ciel Canadien,,YC,YCC,Ciel,Canada,Y)
(9829,Canadian National Airways,,CN,YCP,CaNational,Canada,Y)
(10673,CanXpress,,C1,CA1,CAX,Canada,Y)
(10683,CanXplorer,,C2,CAP,,Canada,Y)
(10735,World Experience Airline,WEA,W1,WE1,WEA,Canada,Y)
(10945,Enerjet,,,ENJ,ENERJET AIR,Canada,Y)
(11922,Worldways,,WG,WGC,,Canada,N)
(14725,Air Atlantic,,9A,\N,Atlantic,Canada,N)
(16110,Harbour Air (Priv),,H3,\N,,Canada,Y)
(16329,Sunwing,,WG,\N,sunwing,Canada,Y)
(16459,Sky Regional,Air Canada Express,RS,\N,Sky Regional,Canada,Y)
(16721,Red Jet Canada,,QY,\N,,Canada,Y)
(16753,Canadian World,,10,CNN,Canadian,Canada,N)
(19337,Transair,,,TTZ,,Canada,Y)
(19675,Rainbow Air Canada,Rainbow Air CAN,RY,RAY,Rainbow CAN,Canada,Y)
grunt> █
```

# Dump finalair2;





```
(5161,Trans Atlantis,\N,,LTA,LANTRA,Canada,N)
(5292,Vacationair,\N,,VAC,VACATIONAIR,Canada,N)
(5307,Voyageur Airways,\N,VC,VAL,VOYAGEUR,Canada,N)
(5313,VIA Rail Canada,\N,2R,,,Canada,N)
(5349,Vistajet,\N,,VJT,VISTA,Canada,N)
(5371,Vision Airways Corporation,\N,,VSN,VISION,Canada,N)
(5391,Walsten Air Services,\N,,WAS,WALSTEN,Canada,N)
(5393,Wapiti Aviation,\N,,WPT,WAPITI,Canada,N)
(5396,Wasaya Airways,\N,,WSG,WASAYA,Canada,N)
(5411,West Coast Air,\N,80,,,Canada,Y)
(5415,West Wind Aviation,\N,,WEW,WESTWIND,Canada,N)
(5416,WestJet,\N,WS,WJA,WESTJET,Canada,Y)
(5420,Westcoast Energy,\N,,BLK,BLUE FLAME,Canada,N)
(5425,Western Arctic Air,\N,,WAL,WESTERN ARCTIC,Canada,N)
(5427,Western Express Air Lines,\N,,WES,WEST EX,Canada,N)
(5433,Westpoint Air,\N,,WTP,WESTPOINT,Canada,N)
(5437,White River Air Services,\N,,WRA,,Canada,N)
(5466,World Weatherwatch,\N,,XWW,,Canada,N)
(5469,Worldwide Air Charter Systems,\N,,CSW,CHARTER SYSTEMS,Canada,N)
(5521,Zip,\N,3J,WZP,ZIPPER,Canada,N)
(5523,Zoom Airlines,\N,Z4,OOM,ZOOM,Canada,Y)
(9828,Ciel Canadien,,YC,YCC,Ciel,Canada,Y)
(9829,Canadian National Airways,,CN,YCP,CaNational,Canada,Y)
(10673,CanXpress,,C1,CA1,CAX,Canada,Y)
(10683,CanXplorer,,C2,CAP,,Canada,Y)
(10735,World Experience Airline,WEA,W1,WE1,WEA,Canada,Y)
(10945,Enerjet,,ENJ,ENERJET AIR,Canada,Y)
(11922,Worldways,,WG,WGC,,Canada,N)
(14725,Air Atlantic,,9A,\N,Atlantic,Canada,N)
(16110,Harbour Air (Priv),,H3,\N,,Canada,Y)
(16329,Sunwing,,WG,\N,sunwing,Canada,Y)
(16459,Sky Regional,Air Canada Express,RS,\N,Sky Regional,Canada,Y)
(16721,Red Jet Canada,,QY,\N,,Canada,Y)
(16753,Canadian World,,10,CNN,Canadian,Canada,N)
(19337,Transair,,,TTZ,,Canada,Y)
(19675,Rainbow Air Canada,Rainbow Air CAN,RY,RAY,Rainbow CAN,Canada,Y)
grunt> █
```

2. finalair3 = UNION finalair1,  
finalair2;

dump finalair3;

```
(5223,University Air Squadron,\N,,UAA,,United Kingdom,N)
(5224,University Air Squadron,\N,,UAH,,United Kingdom,N)
(5226,UK HEMS,\N,,HLE,HELIMED,United Kingdom,N)
(5249,UK International Airlines,\N,,UKI,KHALIQ,United Kingdom,N)
(5253,United Kingdom Home Office,\N,,UKP,POLICE,United Kingdom,N)
(5306,Veritair Ltd,\N,,BTP,NET RAIL,United Kingdom,N)
(5314,Victoria Aviation,\N,,ENV,ENDEAVOUR,United Kingdom,N)
(5347,Virgin Atlantic Airways,\N,,VS,VIR,VIRGIN,United Kingdom,Y)
(5365,Vertair,\N,,VRA,VERITAIR,United Kingdom,N)
(5369,Vickers Limited,\N,,VSB,VICKERS,United Kingdom,N)
(5395,Warwickshire Aerocentre Ltd.,\N,,ATX,AIRTAX,United Kingdom,N)
(5414,West Freugh DTEO,\N,,TEE,TEEBIRD,United Kingdom,N)
(5430,Westflight Aviation,\N,,WSL,WEST LINE,United Kingdom,N)
(5432,Westland Helicopters,\N,,WHE,WESTLAND,United Kingdom,N)
(5444,Williams Grand Prix Engineering,\N,,WGP,GRAND PRIX,United Kingdom,N)
(5464,Woodgate Executive Air Charter,\N,,CWY,CAUSEWAY,United Kingdom,N)
(5475,Wycombe Air Centre,\N,,WYC,WYCOMBE,United Kingdom,N)
(5482,Xclusive Jet Charter Limited,\N,,XJC,EXCLUSIVE JET,United Kingdom,N)
(5486,Xjet Limited,\N,,XJT,XRAY,United Kingdom,N)
(6860,Dan-Air London,,DAN,,United Kingdom,N)
(8523,Inter European Airways,,,IEA,,United Kingdom,N)
(10318,Air UK,,UK,\N,,United Kingdom,N)
(10334,Suckling Airways,,CB,\N,,United Kingdom,N)
(10642,Jc royal.britannica,,,JRB,,United Kingdom,Y)
(10646,Birmingham European,,VB,\N,,United Kingdom,N)
(13388,Court Line,,??,???,United Kingdom,N)
(13936,Royal European Airlines,,69,\N,,United Kingdom,Y)
(16260,AUOS,AUOS,,AUK,AUOS,United Kingdom,N)
(16261,CB Airways UK ( Interliging Flights ),,1F,CIF,,United Kingdom,Y)
(16881,Air Cudlua,Air Cudlua,,CUD,Cudlua,United Kingdom,Y)
(17666,Go Fly (United Kingdom),,,GOE,Go Flight,United Kingdom,N)
(18681,Fly Zoom,,,UKZ,,United Kingdom,N)
(19525,BBN-Airways,BlackBurn,,EGH,BBN,United Kingdom,Y)
(19677,Rainbow Air Euro,Rainbow Air EU,RU,RUE,Rainbow Air,United Kingdom,Y)
(19808,All Europe,All Europe,N9,N99,,United Kingdom,Y)
(19811,British Air Ferries,,??,??!,United Kingdom,N)
grunt> █
```

3. A = filter finalair by  
territory == 'United Kingdom';

dump A;



```

), (LX, 4559, ZRH, 1678, VYZ, 193, Y), (LX, 4559, ZRH, 1678, ZAG, 1208, Y), (LY, 2150, ZRH, 1678, TLV, 1590, ), (S4, 4304, ZRH, 1678, LIS, 1638, Y), (FB, 1548, ZRH, 1678, SOF, 1194, ), (AF, 137, ZRH, 1678, CDG, 1382, ), (AT, 4248, ZRH, 1678, CMN, 1074, ), (RJ, 4259, ZRH, 1678, VIE, 1613, ), (RJ, 4259, ZRH, 1678, GVA, 1665, ), (RJ, 4259, ZRH, 1678, AMM, 2170, ), (MS, 2143, ZRH, 1678, CAI, 1128, Y), (AB, 214, ZRH, 1678, ACE, 1055, ), (AB, 214, ZRH, 1678, ALC, 1212, ), (AB, 214, ZRH, 1678, BDS, 1506, ), (AB, 214, ZRH, 1678, CTA, 1509, ), (AB, 214, ZRH, 1678, DUS, 345, ), (AB, 214, ZRH, 1678, FNC, 4091, ), (AB, 214, ZRH, 1678, FUE, 1051, ), (AB, 214, ZRH, 1678, GWT, 410, ), (AB, 214, ZRH, 1678, HEL, 421, ), (AB, 214, ZRH, 1678, HER, 1452, ), (AB, 214, ZRH, 1678, IBZ, 1225, ), (AB, 214, ZRH, 1678, LCY, 503, Y), (AB, 214, ZRH, 1678, LHR, 507, Y), (AB, 214, ZRH, 1678, LPA, 1054, ), (AB, 214, ZRH, 1678, MAD, 1229, Y), (AB, 214, ZRH, 1678, NAP, 1561, ), (AB, 214, ZRH, 1678, OLB, 1520, ), (AB, 214, ZRH, 1678, PMI, 3998, ), (AB, 214, ZRH, 1678, SUF, 1508, ), (AB, 214, ZRH, 1678, TFS, 1056, ), (AB, 214, ZRH, 1678, TXL, 351, ), (AB, 214, ZRH, 1678, VIE, 1613, Y), (PS, 5282, ZRH, 1678, KBP, 2939, ), (US, 5265, ZRH, 1678, SIN, 3316, ), (US, 5265, ZRH, 1678, PHL, 3752, ), (US, 5265, ZRH, 1678, LIS, 1638, Y), (US, 5265, ZRH, 1678, JFK, 3797, ), (US, 5265, ZRH, 1678, IST, 1701, Y), (SI, 4438, ZRH, 1678, JER, 499, ), (QS, 5097, ZRH, 1678, PRG, 1587, ), (QR, 4091, ZRH, 1678, DOH, 2241, ), (FI, 2835, ZRH, 1678, KEF, 16, ), (QF, 4089, ZRH, 1678, DXB, 2188, Y), (NH, 324, ZRH, 1678, NRT, 2279, Y), (A I, 218, ZRH, 1678, DEL, 3093, Y), (AI, 218, ZRH, 1678, BOM, 2997, Y), (OZ, 28, ZRH, 1678, IST, 1701, ), (OK, 1946, ZRH, 1678, PRG, 1587, ), (OS, 491, ZRH, 1678, VIE, 1613, Y), (A3, 96, ZRH, 1678, LIS, 1638, ), (4T, 1401, ZRH, 1678, AYT, 1688, ))
(ZRJ, {(WT, \N, ZRJ, 5547, MSA, 5544, ), (WT, \N, ZRJ, 5547, YAC, 5478, )})
(ZSA, {(UP, 1403, ZSA, 1956, NAS, 1953, ), (PL, 179, ZSA, 1956, NAS, 1953, ), (AC, 330, ZSA, 1956, PLS, 1754, ), (NK, 4687, ZSA, 1956, FLL, 3533, )})
(ZSE, {(MK, 1057, ZSE, 917, MRU, 893, ), (AF, 137, ZSE, 917, MRU, 893, Y), (UU, 1191, ZSE, 917, MRU, 893, )})
(ZSJ, {(WT, \N, ZSJ, 5548, YRL, 5527, ), (WT, \N, ZSJ, 5548, KEW, 8355, ), (WT, \N, ZSJ, 5548, YPM, 5521, )})
(ZTB, {(WJ, 5399, ZTB, 5473, ZLT, 5474, ), (WJ, 5399, ZTB, 5473, YHR, 5503, )})
(ZTH, {(AB, 214, ZTH, 1488, MUC, 346, ), (AB, 214, ZTH, 1488, VIE, 1613, ), (DE, 1868, ZTH, 1488, MAN, 478, ), (A3, 96, ZTH, 1488, ATH, 3941, ), (U2, 2297, ZTH, 1488, LGW, 502, ), (GQ, 1492, ZTH, 1488, KIT, 1454, ), (TO, 8745, ZTH, 1488, LGW, 502, ), (HQ, 2684, ZTH, 1488, BRU, 302, ), (HG, 3661, ZTH, 1488, VIE, 1613, ), (LS, 3026, ZTH, 1488, LBA, 517, ), (HV, 5039, ZTH, 1488, AMS, 580, ), (FR, 4296, ZTH, 1488, CRL, 304, ), (OA, 3776, ZTH, 1488, ATH, 3941, ), (AB, 214, ZTH, 1488, DUS, 345, ), (LS, 3026, ZTH, 1488, MAN, 478, ), (GQ, 1492, ZTH, 1488, EFL, 1455, )})
(ZUH, {(CA, 751, ZUH, 6355, HYN, 6384, ), (CA, 751, ZUH, 6355, FOC, 3385, Y), (PN, 12978, ZUH, 6355, CGO, 3375, ), (CZ, 1767, ZUH, 6355, HGH, 3386, ), (CA, 751, ZUH, 6355, CKG, 3393, ), (CZ, 1767, ZUH, 6355, HFE, 3389, ), (PN, 12978, ZUH, 6355, CKG, 3393, ), (CZ, 1767, ZUH, 6355, HAK, 4120, ), (CZ, 1767, ZUH, 6355, FOC, 3385, Y), (CA, 751, ZUH, 6355, CTU, 3395, ), (FM, 4609, ZUH, 6355, PVG, 3406, ), (GS, 690, ZUH, 6355, KWL, 3372, ), (9C, 15999, ZUH, 6355, PVG, 3406, ), (HO, 3081, ZUH, 6355, SHA, 3391, ), (9C, 15999, ZUH, 6355, SHA, 3391, ), (CZ, 1767, ZUH, 6355, YIW, 6394, ), (CZ, 1767, ZUH, 6355, XMN, 3383, Y), (MU, 1758, ZUH, 6355, PVG, 3406, ), (CZ, 1767, ZUH, 6355, WUH, 3376, ), (MF, 5484, ZUH, 6355, FOC, 3385, ), (MF, 5484, ZUH, 6355, XMN, 3383, ), (MU, 1758, ZUH, 6355, SHA, 3391, ), (HO, 3081, ZUH, 6355, PVG, 3406, ), (ZH, 4611, ZUH, 6355, PEK, 3364, Y), (ZH, 4611, ZUH, 6355, NNG, 3373, ), (ZH, 4611, ZUH, 6355, HYN, 6384, ), (GS, 15837, ZUH, 6355, MUZ, 6744, ), (ZH, 4611, ZUH, 6355, FOC, 3385, ), (ZH, 4611, ZUH, 6355, CKG, 3393, Y), (CZ, 1767, ZUH, 6355, KWE, 3394, ), (GS, 690, ZUH, 6355, KOW, 6379, ), (CA, 751, ZUH, 6355, XMN, 3383, Y), (CA, 751, ZUH, 6355, WNZ, 6392, Y), (CZ, 1767, ZUH, 6355, KMG, 3382, ), (GS, 690, ZUH, 6355, HAK, 4120, ), (CZ, 1767, ZUH, 6355, CSX, 3371, ), (CZ, 1767, ZUH, 6355, SYX, 4030, ), (CZ, 1767, ZUH, 6355, SWA, 4302, ), (BK, 4026, ZUH, 6355, SYX, 4030, ), (BK, 4026, ZUH, 6355, TSN, 3368, ), (CZ, 1767, ZUH, 6355, CGO, 3375, ), (GS, 15837, ZUH, 6355, LZH, 6356, ), (3U, 4608, ZUH, 6355, HAK, 4120, ), (3U, 4608, ZUH, 6355, CZX, 4109, ), (S C, 4599, ZUH, 6355, XMN, 3383, ), (NS, 16136, ZUH, 6355, NGB, 3387, ), (SC, 4599, ZUH, 6355, WNZ, 6392, ), (CZ, 1767, ZUH, 6355, SHA, 3391, ), (CZ, 1767, ZUH, 6355, PVG, 3406, ), (CZ, 1767, ZUH, 6355, PEK, 3364, ), (SC, 4599, ZUH, 6355, NNG, 3373, ), (SC, 4599, ZUH, 6355, KWE, 3394, ), (CA, 751, ZUH, 6355, KWE, 3394, Y), (CA, 751, ZUH, 6355, KHN, 3384, Y), (CZ, 1767, ZUH, 6355, MXZ, 6354, ), (CZ, 1767, ZUH, 6355, KHN, 3384, ), (SC, 4599, ZUH, 6355, KHN, 3384, ), (CA, 751, ZUH, 6355, PEK, 3364, ), (CA, 751, ZUH, 6355, NNG, 3373, ), (ZH, 4611, ZUH, 6355, CTU, 3395, Y)})
(ZUM, {(PB, \N, ZUM, 5550, YMK, 161, ), (PB, \N, ZUM, 5550, YZR, 188, )})
(ZVK, {(QV, 3233, ZVK, 3118, BKK, 3885, ), (QV, 3233, ZVK, 3118, PKZ, 3116, ), (QV, 3233, ZVK, 3118, VTE, 3120, )})
(ZYI, {(GS, 690, ZYI, \N, XMN, 3383, ), (GS, 690, ZYI, \N, CSX, 3371, ), (GS, 690, ZYI, \N, HAK, 4120, ), (GS, 690, ZYI, \N, SHA, 3391, ), (GS, 690, ZYI, \N, XIY, 3379, ), (ZH, 4611, ZYI, \N, PEK, 3364, Y), (CZ, 1767, ZYI, \N, SZX, 3374, ), (EU, 2193, ZYI, \N, CTU, 3395, ), (CZ, 1767, ZYI, \N, PEK, 3364, ), (MU, 1758, ZYI, \N, WUH, 3376, ), (MU, 1758, ZYI, \N, NKG, 3388, ), (9C, 15999, ZYI, \N, KMG, 3382, ), (CZ, 1767, ZYI, \N, CAN, 3370, ), (CA, 751, ZYI, \N, PEK, 3364, )})
(ZYL, {(RX, 19676, ZYL, 3074, DAC, 3076, ), (VQ, 11948, ZYL, 3074, DAC, 3076, ), (BG, 1359, ZYL, 3074, DAC, 3076, ), (4H, 8463, ZYL, 3074, DAC, 3076, )})
grunt> █

```

4. group\_data = GROUP  
routes  
by route\_source\_iata;  
dump group\_data;

## Browse Directory

/home/hduser/output

Go!



Show 25 entries

Search:

<input type="checkbox"/>	Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name	
<input type="checkbox"/>	-rw-r--r--	hduser	supergroup	0 B	Apr 08 21:18	1	128 MB	<a href="#">_SUCCESS</a>	
<input type="checkbox"/>	-rw-r--r--	hduser	supergroup	22.74 KB	Apr 08 21:18	1	128 MB	<a href="#">part-m-00000</a>	

Showing 1 to 2 of 2 entries

Previous 1 Next

5. store A into '/home/hduser/output'  
using PigStorage (',' );

```
hadoop fs -cat  
/home/hduser/output/part-m-00000
```

DIVESH KOSURI



CB.EN.U4CSE19422

```

2022-04-05 00:08:47,482 [main] WARN org.apache.hadoop.metrics2.impl.MetricsSystemImpl - JobTracker m
2022-04-05 00:08:47,483 [main] WARN org.apache.hadoop.metrics2.impl.MetricsSystemImpl - JobTracker m
2022-04-05 00:08:47,486 [main] WARN org.apache.hadoop.metrics2.impl.MetricsSystemImpl - JobTracker m
2022-04-05 00:08:47,489 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.Map
2022-04-05 00:08:47,490 [main] WARN org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has
2022-04-05 00:08:47,496 [main] INFO org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total in
2022-04-05 00:08:47,496 [main] INFO org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil -
(2,135 Airways,\N,,GNL,GENERAL,United States,N)
(9,3D Aviation,\N,,SEC,SECUREX,United States,N)
(10,40-Mile Air,\N,Q5,MLA,MILE-AIR,United States,Y)
(22,Aloha Airlines,\N,AQ,AAH,ALOHA,United States,Y)
(23,Alaska Island Air,\N,,AAK,ALASKA ISLAND,United States,N)
(24,American Airlines,\N,AA,AAL,AMERICAN,United States,Y)
(25,Aviation Management Corporation,\N,,AAM,AM CORP,United States,N)
(26,Atlantis Airlines (USA),\N,,AAO,ATLANTIS AIR,United States,N)
(31,Astro Air International,\N,,AAV,ASTRO-PHIL,United States,N)
(35,Allegiant Air,\N,G4,AAV,ALLEGIAN,United States,Y)
(49,Airborne Express,\N,GB,ABX,ABEX,United States,N)
(50,ABX Air,\N,GB,ABX,ABEX,United States,N)
(53,Academy Airlines,\N,,ACD,ACADEMY,United States,N)
(57,Air Aurora,\N,,AAI,BOREALIS,United States,N)
(66,Advance Leasing Company,\N,4G,,,United States,N)
(67,Aztec Worldwide Airlines,\N,7A,,,United States,N)
(81,Aerodyne,\N,,ADY,AERODYNE,United States,N)
(89,Anderson Aviation,\N,,ADX,ANDAX,United States,N)
(109,Alaska Central Express,\N,KO,AER,ACE AIR,United States,Y)
(115,Airway Express,\N,,AEX,AVCO,United States,N)
(117,Aerial Transit Company,\N,,AEZ,AERIAL TRANZ,United States,N)
(141,Air Guam,\N,,AGM,AIR GUAM,United States,N)
(145,Air Cargo America,\N,,MVM,PEGASUS,United States,N)
(149,Air Cargo Carriers,\N,2Q,SNC,NIGHT CARGO,United States,Y)

```

```

HadoopVersion  PigVersion  UserId  StartedAt      FinishedAt      Features
3.3.1          0.17.0         hduser  2022-04-05 00:08:39  2022-04-05 00:08:47  FILTER

Success!

Job Stats (time in seconds):
JobId  Maps  Reduces  MaxMapTime  MinMapTime  AvgMapTime  MedianMapTime  MaxReduceTime  MinReduceTime  AvgReduceTime
job_local968342233_0002  1      0      0      n/a      n/a      n/a      n/a      0      0      0

Input(s):
Successfully read 6048 records (14562740 bytes) from: "/CaseStudy/Final_airlines"

Output(s):
Successfully stored 1080 records (12018703 bytes) in: "hdfs://localhost:54310/tmp/temp-1344964502/tmp522233176"

Counters:
Total records written : 1080
Total bytes written : 12018703
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local968342233_0002

```

## Query 1:

Query1 = FILTER finalairs  
BY territory == 'United  
States';  
Dump Query1;

```
(ZL,4178,SYD,3361,MYA,6302,)  
(ZL,4178,SYD,3361,NRA,6307,)  
(ZL,4178,SYD,3361,NTL,4320,)  
(ZL,4178,SYD,3361,OAG,6793,)  
(ZL,4178,SYD,3361,PKE,6317,)  
(ZL,4178,SYD,3361,TRO,6794,)  
(ZL,4178,SYD,3361,WGA,3363,)  
(ZL,4178,TRO,6794,GFN,6792,)  
(ZL,4178,TRO,6794,SYD,3361,)
```

```
2022-04-05 00:16:38,479 [LocalJobRunner Map Task Executor #0] INFO org.apache.hadoop.mapred.LocalJobRunner - map  
2022-04-05 00:16:38,479 [LocalJobRunner Map Task Executor #0] INFO org.apache.hadoop.mapred.Task - Task 'attempt_local344417590_0005_m_000000_0'  
2022-04-05 00:16:38,481 [LocalJobRunner Map Task Executor #0] INFO org.apache.hadoop.mapred.Task - Final Counters for attempt_local344417590_0  
File System Counters  
  FILE: Number of bytes read=2158  
  FILE: Number of bytes written=32777528  
  FILE: Number of read operations=0  
  FILE: Number of large read operations=0  
  FILE: Number of write operations=0  
  HDFS: Number of bytes read=44860897  
  HDFS: Number of bytes written=38020006  
  HDFS: Number of read operations=493  
  HDFS: Number of large read operations=0  
  HDFS: Number of write operations=65  
  HDFS: Number of bytes read erasure-coded=0  
Map-Reduce Framework  
  Map input records=67663  
  Map output records=67663  
  Input split bytes=370  
  Spilled Records=0  
  Failed Shuffles=0  
  Merged Map outputs=0  
  GC time elapsed (ms)=60  
  Total committed heap usage (bytes)=396886016  
File Input Format Counters  
  Bytes Read=0  
File Output Format Counters  
  Bytes Written=0
```

## Query2 : DISTINCT

Query2 = DISTINCT  
routes;  
DUMP Query2;

```

2022-04-05 00:28:13,840 [LocalJobRunner Map Task Executor #0] INFO org.apache.hadoop.mapred.Task
File System Counters
    FILE: Number of bytes read=6834859
    FILE: Number of bytes written=55729492
    FILE: Number of read operations=0
    FILE: Number of large read operations=0
    FILE: Number of write operations=0
    HDFS: Number of bytes read=72931082
    HDFS: Number of bytes written=58111619
    HDFS: Number of read operations=641
    HDFS: Number of large read operations=0
    HDFS: Number of write operations=89
    HDFS: Number of bytes read erasure-coded=0
Map-Reduce Framework
    Map input records=67663
    Map output records=67663
    Map output bytes=2843656
    Map output materialized bytes=2978988
    Input split bytes=370
    Combine input records=0
    Spilled Records=67663
    Failed Shuffles=0
    Merged Map outputs=0
    GC time elapsed (ms)=20
    Total committed heap usage (bytes)=652214272
MultiInputCounters
    Input records from _1_routes.dat=67663
File Input Format Counters
    Bytes Read=0

```

```

Regional Express,\N,ZL,RXA,REX,Australia,Y,ZL,4178,NT
Regional Express,\N,ZL,RXA,REX,Australia,Y,ZL,4178,NT
Regional Express,\N,ZL,RXA,REX,Australia,Y,ZL,4178,OA
Regional Express,\N,ZL,RXA,REX,Australia,Y,ZL,4178,PK
Regional Express,\N,ZL,RXA,REX,Australia,Y,ZL,4178,PK
Regional Express,\N,ZL,RXA,REX,Australia,Y,ZL,4178,PL
Regional Express,\N,ZL,RXA,REX,Australia,Y,ZL,4178,RC
,Apache Air,Apache,ZM,IWA,APACHE,United States,Y,ZM,1
,Apache Air,Apache,ZM,IWA,APACHE,United States,Y,ZM,1
,Apache Air,Apache,ZM,IWA,APACHE,United States,Y,ZM,1

```

## Query 3: JOINS

Query3 = JOIN finalairs BY  
 airline\_iata, routes BY  
 route\_iata;  
 DUMP Query3;



```

2022-04-05 00:35:44.620 [LocalJobRunner Map Task Executor #0] INFO org.apache.hadoop.mapred.task - Final Counters for attempt_local27796562_0000_0: Counters: 21
File System Counters
  FILE: Number of bytes read=13666332
  FILE: Number of bytes written=4613702
  FILE: Number of read operations=0
  FILE: Number of large read operations=0
  FILE: Number of write operations=0
  HDFS: Number of bytes read=8958661
  HDFS: Number of bytes written=7641007
  HDFS: Number of read operations=733
  HDFS: Number of large read operations=0
  HDFS: Number of write operations=104
  HDFS: Number of bytes read erasure-coded=0
Map-Reduce Framework
  Map input records=67663
  Map output records=67663
  Input split bytes=370
  Spilled Records=0
  Failed Shuffles=0
  Merged Map outputs=0
  GC time elapsed (ms)=5
  Total committed heap usage (bytes)=513277952
File Input Format Counters
  Bytes Read=0
File Output Format Counters
  Bytes Written=0

```

```

(4178,TRO,3361)
(4178,TSV,3324)
(4178,TSV,6337)
(4178,WGA,3339)
(4178,WGA,3361)
(4178,WIN,6289)
(4178,WIN,3330)
(4178,WYA,3341)
(19016,DME,2912)
(19016,FRU,4029)
(19016,FRU,2913)

```

## Query 4: FOREACH and GENERATE

Query4 = FOREACH routes GENERATE  
 route\_airid, route\_source\_iata, route\_stops;  
 DUMP Query4;

```
hduser@DiveshKosuri:~$ hadoop fs -mkdir /home/hduser/OutPut
2022-04-08 22:12:22,188 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
mkdir: /home/hduser/OutPut: File exists
hduser@DiveshKosuri:~$ hadoop fs -cat /home/hduser/OutPut/part-m-00000
2022-04-08 22:13:35,862 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
2,135 Airways,\N,,GNL,GENERAL,United States,N
9,30 Aviation,\N,,SEC,SECUREX,United States,N
10,40-Mile Air,\N,Q5,MLA,MILE-AIR,United States,Y
22,Aloha Airlines,\N,AQ,AAH,ALOHA,United States,Y
23,Alaska Island Air,\N,,AAK,ALASKA ISLAND,United States,N
24,American Airlines,\N,AA,AAL,AMERICAN,United States,Y
25,Aviation Management Corporation,\N,,AAM,AM CORP,United States,N
26,Atlantis Airlines (USA),\N,,AAQ,ATLANTIS AIR,United States,N
31,Astro Air International,\N,,AAV,ASTRO-PHIL,United States,N
35,Allegiant Air,\N,G4,AAV,ALLEGIANTE,United States,Y
49,Airborne Express,\N,GB,ABX,ABEX,United States,N
50,ABX Air,\N,GB,ABX,ABEX,United States,N
53,Academy Airlines,\N,,ACD,ACADEMY,United States,N
57,Air Aurora,\N,,AAI,BOREALIS,United States,N
66,Advance Leasing Company,\N,4G,,,United States,N
67,Aztec Worldwide Airlines,\N,7A,,,United States,N
81,Aerodyne,\N,,ADY,AERODYNE,United States,N
89,Anderson Aviation,\N,,ADX,ANDAX,United States,N
109,Alaska Central Express,\N,KO,AER,ACE AIR,United States,Y
115,Airway Express,\N,,AEX,AVCO,United States,N
117,Aerial Transit Company,\N,,AEZ,AERIAL TRANZ,United States,N
141,Air Guam,\N,,AGM,AIR GUAM,United States,N
145,Air Cargo America,\N,,MVM,PEGASUS,United States,N
149,Air Cargo Carriers,\N,2Q,SNC,NIIGHT CARGO,United States,Y
164,Atlantic Gulf Airlines,\N,,AGF,ATLANTIC GULF,United States,N
177,Atlantic Coast Airlines,\N,,BLR,BLUE RIDGE,United States,N
182,Aero Flight Service,\N,,AGV,FLIGHT GROUP,United States,N
188,Aspen Helicopters,\N,,AHF,ASPEL,United States,N
205,Alpine Air Express,\N,SA,AIP,ALPINE AIR,United States,Y
210,Airlift International,\N,,AIR,AIRLIFT,United States,Y
217,Air Illinois,\N,,AIL,AIR ILLINOIS,United States,N
236,Ameristar Jet Charter,\N,,AJI,AMERISTAR,United States,N
237,A2 Jet Leasing,\N,,AJJ,ATLANTIC JET,United States,N
```

```
2022-04-08 22:31:41,493 [pool-24-thread-1] INFO org.apache.hadoop.mapred.LocalJobRunner - Finishing task: attempt_local1641374742_0007_r_000000_0
2022-04-08 22:31:41,495 [Thread-S10] INFO org.apache.hadoop.mapred.LocalJobRunner - reduce task executor complete.
2022-04-08 22:31:41,799 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - Running jobs are [job_local1641374742_0007]
2022-04-08 22:31:46,327 [main] WARN org.apache.hadoop.metrics2Impl.MetricsSystemImpl - JobTracker metrics system already initialized!
2022-04-08 22:31:46,332 [main] WARN org.apache.hadoop.metrics2Impl.MetricsSystemImpl - JobTracker metrics system already initialized!
2022-04-08 22:31:46,335 [main] WARN org.apache.hadoop.metrics2Impl.MetricsSystemImpl - JobTracker metrics system already initialized!
2022-04-08 22:31:46,359 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100% complete
2022-04-08 22:31:46,366 [main] INFO org.apache.pig.tools.pigstats.SimplePigStats - Script Statistics:

HadoopVersion PigVersion UserId StartedAt FinishedAt Features
3.3.1 0.17.0 hduser 2022-04-08 22:31:33 2022-04-08 22:31:46 GROUP_BY,LIMIT

Success!

Job Stats (time in seconds):
JobId Maps Reduces MaxMapTime MinMapTime AvgMapTime MedianMapTime MaxReduceTime MinReduceTime AvgReduceTime MedianReduceTime Alias Feature Outputs
job_local1641374742_0007 1 1 n/a n/a n/a n/a n/a n/a n/a Query7,Query8,FinalAirs GROUP_BY hdfs://localhost:54310/tmp/tmp1955396190/tmp1969260916,
Input(s):
Successfully read 6048 records (83000370 bytes) from: "/CaseStudy/Final airlines"
```

## Query 6: STORE

`hadoop fs -mkdir  
/home/hduser/OutPut`

`STORE Query1 INTO  
'/home/hduser/OutPut' USING  
PigStorage (',' );`

`hadoop fs -ls  
/home/hduser/OutPut`

`hadoop fs -cat  
/home/hduser/OutPut/part-m-  
00000`

```
Total Length = 316243
Input split[0]:
  Length = 316243
  ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
  Locations:
```

## Query 7: Group

Query7 = GROUP finalairs  
by territory;  
DUMP Query7;

```
Democratic Republic of the Congo,((486,Air Transport Service,N,,ATS,,Democratic Republic of the Congo,N),(396,Air Plan International,N,,APV,AIR PLAN,Democratic Republic of the Congo,N),(764,Air Chartr  
Services,N,,CHB,ZAIRE CHARTER,Democratic Republic of the Congo,N),(4338,Scibe Airlift,N,,582,SCIBE AIRLIFT,Democratic Republic of the Congo,N),(1953,DAS Airlines,N,,RKC,DAS CONGO,Democratic Republic  
of the Congo,N),(888,African Company Airlines,N,,FNF,AFRICOMPANY,Democratic Republic of the Congo,N),(3146,Kinshasa Airways,N,,KAS,KINSHASA AIRWAYS,Democratic Republic of the Congo,N),(1479,Bravo Air  
Congo,N,,K6,BRV,BRAVO,Democratic Republic of the Congo,N),(1502,Blue Airlines,N,,BUL,BLUE AIRLINES,Democratic Republic of the Congo,N),(3261,Lignes Aeriennes Congolaises,N,,LCC,CONGOLAISE,Democratic Re  
public of the Congo,N),(2756,Hewa Bora Airways,N,,EO,ALX,ALLCONGO,Democratic Republic of the Congo,N),(38,African Business and Transportations,N,,ABB,AFRICAN BUSINESS,Democratic Republic of the Congo,N  
,(4482,Shahair,N,,SHB,SHMAIR,Democratic Republic of the Congo,N),(1718,Cetraca Aviation Service,N,,CEA,CETRACA,Democratic Republic of the Congo,N),(2418,Funtsha Aviation Service,N,,FUN,FUNTSHI,Democr  
atic Republic of the Congo,N),(2215,Espace Aviation Services,N,,EAS,ESPACE,Democratic Republic of the Congo,N),(2284,Enterprise World Airways,N,,EWS,WORLD ENTERPRISE,Democratic Republic of the Congo,N  
,(3382,Mallia Airlift,N,,MLC,MALLIA,Democratic Republic of the Congo,N),(758,Africa One,N,,CFR,,Democratic Republic of the Congo,N))  
Saint Vincent and the Grenadines,((4534,St. Vincent Grenadines Air (1998),N,,SVG,GRENADINES,Saint Vincent and the Grenadines,N),(11724,SVG Air,,SVG,Grenadines,Saint Vincent and the Grenadines,Y)))  
Democratic People's Republic of Korea,((1998,Air Koryo,N,,KOS,AIR KORYO,Democratic People's Republic of Korea,Y)))  
,,((415,Aeroflot de Costa Rica,N,,ARJ,,S.A.,,N),(1,Private flight,N,,N/A,,Y),(4438,Scorpio Aviation,N,,85,,N),(4854,Sirin,N,,R1,,N),(4051,Princess Air,N,,80,,N),(3562,Multiflight,N,,MFT,YORKAIR  
,N),(5199,Transports et Travaux A,N,,OF,,N),(2804,ICAO,N,,YYY,,N),(2649,Guyana Airways 2000,N,,GY,,N),(2636,Gulf African Airlines - Gambia,N,,GUF,GULF AFRICAN,,N),(2080,Electronic Data Systems,N
```

```

File System Counters
  FILE: Number of bytes read=8091982
  FILE: Number of bytes written=34351155
  FILE: Number of read operations=0
  FILE: Number of large read operations=0
  FILE: Number of write operations=0
  HDFS: Number of bytes read=29671070
  HDFS: Number of bytes written=26409654
  HDFS: Number of read operations=570
  HDFS: Number of large read operations=0
  HDFS: Number of write operations=52
  HDFS: Number of bytes read erasure-coded=0
Map-Reduce Framework
  Combine input records=0
  Combine output records=0
  Reduce input groups=277
  Reduce shuffle bytes=431550
  Reduce input records=6048
  Reduce output records=4
  Spilled Records=6048
  Shuffled Maps =1
  Failed Shuffles=0
  Merged Map outputs=1
  GC time elapsed (ms)=0
  Total committed heap usage (bytes)=697303040
Shuffle Errors
  BAD_ID=0
  CONNECTION=0
  IO_ERROR=0
  WRONG_LENGTH=0
  WRONG_MAP=0
  WRONG_REDUCE=0
File Output Format Counters
  Bytes Written=0

```

```

2022-04-08 22:31:46,403 [main] INFO org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to process : 1
(5,((1390,British Gulf International,(N,BGI,BRITISH GULF,S,N),(2290,Express International Cargo,(N,EIC,EXCARGO,S,N),(3591,Natalco Air Lines,(N,NC0,NATALCO,S,N),(3286,Linhas Aereas Santonenses,(N,SMS,
SANTONENSES,S,N),(3289,LASTP,(N,0N,LASTP,S,N)))
((N,((5533,Tyrolean Airways,(N,N,TYR,TYROLEAN,(N,Y),(5556,buzz,(N,UK,BUZ,(N,N,N),(-1,Unknown,(N,-,N/A,(N,N,Y)))
(LAP,((1207,LAP Colombia - Lineas Aereas Petroleras,(N,APT,S.A.,LAP,N)))
(ACON,((4456,Southern Jersey Airways,(N,ALC,Inc.,ACON,N)))
(ASUR,((1851,Compania Boliviana de Transporte Aereo Privado Aerosur,(N,ASU,S.A.,ASUR,N)))
(ATCO,((400,Aerolatias,(N,AQA,S.A.,ATCO,N)))
(AUSA,((513,Aerollneas Uruguayas,(N,ROY,S.A.,AUSA,N)))
(Chad,((3577,Nada Air Service,(N,NAD,NADA AIR,Chad,N),(894,Air Affaires Tchad,(N,FTC,AFFAIRES TCHAD,Chad,N),(1733,Charl Aviation Services,(N,S8,CSU,CHARI SERVICE,Chad,N),(3262,Lignes Aeriennes Du Tchad
,(N,LKD,LATCHAD,Chad,N),(4887,Tchad Airlines,(N,TCD,TCHADLINES,Chad,N),(4942,Toumai Air Tchad,(N,THE,TOUMAI AIR,Chad,N),(1312,Air Horizon,(N,TPK,TCHAD-HORIZON,Chad,N),(433,Air Tchad,(N,HTT,HOTEL TAN
GO,Chad,N)))

```

# Query 8: Limit

Query8 = LIMIT  
Query7 15;

```

UDF_Divesh.java x PrimeCheck.java
1 import java.io.IOException;
2 import org.apache.pig.EvalFunc;
3 import org.apache.pig.data.Tuple;
4
5 import java.io.IOException;
6 import org.apache.pig.EvalFunc;
7 import org.apache.pig.data.Tuple;
8
9 public class UDF_Divesh extends EvalFunc<String>{
10
11     public String exec(Tuple input) throws IOException {
12         if (input == null || input.size() == 0)
13             return null;
14         String str = (String)input.get(0);
15         return str.toUpperCase();
16     }
17 }

```

```

grunt> REGISTER '/home/hduser/BigData/JARs/UDF_Divesh.jar';
grunt> PIG_Input = LOAD '/CaseStudy/PigInput.txt' using PigStorage() AS (line:chararray);
grunt> PIG_UDF = FOREACH PIG_Input GENERATE UDF_Divesh(line);

```

## QUERY 9: UDF

REGISTER

'/home/hduser/BigData/JARs/UDF\_Divesh.jar';

PIG\_Input = LOAD

'/CaseStudy/PigInput.txt'  
using PigStorage() AS  
(line:chararray);

PIG\_UDF =

FOREACH PIG\_Input GENERATE  
UDF\_Divesh(line);