

Impala Queries:

Queries to create tables in Impala for the Solar(UPV and DPV) and Landcover_solar dataset

1. create table upv(latitude string, longitude string, month int, sum double) row format delimited fields terminated by ',' location '/user/cloudera/project/upv/';
2. create table dpv(latitude string, longitude string, month int, sum double) row format delimited fields terminated by ',' location '/user/cloudera/project/dpv/';
3. create table landsolar(longitude string, latitude string, landcover int) row format delimited fields terminated by ',' location '/user/cloudera/project/landcover_solar/';

Query for Distributed PV (DPV) that can be used for solar power generation in houses and buildings and hence will include landcover like Urban, CroPLands and Crop, Natural Veg. Mosaic.

4. select dpv.latitude, dpv.longitude, dpv.month, dpv.sum from dpv, landsolar where
CAST(dpv.latitude as double) = CAST(landsolar.latitude as double) and CAST(dpv.longitude as double) = CAST(landsolar.longitude as double) and landsolar.landcover in (12,13,14);

Result of Analytic:

File Edit View Search Terminal Help				
40.55	-122.35	4	8770.04000000003	
40.65	-112.05	5	226882.9000000005	
41.05	-111.95	2	66521.10000000001	
41.15	-112.05	9	98258.99999999968	
41.25	-111.95	11	52746.20000000001	
41.35	-111.95	10	76365.49999999994	
42.35	-122.85	10	676.1089999999971	
42.95	-71.05	11	12638.60000000003	
43.55	-112.05	3	133.9710000000005	
43.95	-91.15	2	16308.40000000004	
44.05	-123.05	7	12882.78000000003	
44.55	-123.35	10	3004.980000000002	
44.65	-93.05	10	26242.09999999998	
44.75	-91.25	11	21212.70000000001	
44.75	-93.05	11	17864.19999999999	
44.85	-123.05	10	3723.510000000004	
44.95	-93.15	1	17791.10000000006	
44.95	-93.15	12	15038.30000000001	
44.95	-93.35	10	30123.79999999997	
47.05	-122.25	3	3937.120000000012	
47.05	-122.35	2	2635.200000000021	
47.15	-122.25	2	2796.689999999999	
47.15	-122.35	1	1412.329999999992	
47.15	-122.35	12	1624.189999999995	
47.25	-122.35	11	1553.329999999991	
47.25	-122.45	10	15937.4	
47.35	-122.25	11	2424.999999999997	
47.45	-122.25	10	5928.340000000016	
48.15	-122.15	2	5004.590000000016	
+-----+-----+-----+-----+				
Fetched 16644 row(s) in 4.06s				
[quickstart.cloudera:21000] > █				

Query for Utility-scale PV (UPV) that can be used for Solar power farms, large space is required and hence will include Closed Shrubland, Savannas, Desert, Barren and Grasslands. The query for this is given below:

5. select upv.latitude, upv.longitude, upv.month, upv.sum from upv, landsolar where
CAST(upv.latitude as double) = CAST(landsolar.latitude as double) and CAST(upv.longitude as double) = CAST(landsolar.longitude as double) and landsolar.landcover in (6,9,10,16);

Result for the analytic:

File	Edit	View	Search	Terminal	Help
43.25	-100.15	6	192749.19999999997		
43.25	-100.55	2	128220.60000000001		
43.25	-101.45	2	15487.900000000004		
43.25	-101.75	10	17672.299999999999		
43.25	-102.05	5	244945.20000000004		
43.25	-102.45	1	55285.400000000002		
43.25	-102.45	12	39444.300000000003		
43.25	-102.65	10	70311.999999999983		
43.35	-102.35	1	15816.1		
43.35	-102.35	12	11953.300000000001		
43.35	-102.45	11	91566.900000000011		
43.45	-102.35	11	26262.9		
43.45	-102.45	10	137940.9		
43.75	-100.05	2	60455.199999999992		
43.85	-101.15	10	52384.800000000003		
44.15	-121.35	5	51618.500000000011		
44.35	-101.15	3	48311.000000000001		
45.25	-104.15	11	80257.599999999995		
45.35	-104.15	10	176517.3		
46.75	-120.05	1	98976.300000000005		
46.75	-120.05	12	75070.499999999977		
46.75	-120.15	11	67593.499999999984		
46.75	-120.25	10	70816.700000000013		
46.85	-120.05	11	28541.599999999999		
46.85	-120.15	10	175182.30000000004		
47.35	-120.35	1	35758.199999999992		
47.35	-120.35	12	27738.199999999999		
47.45	-120.25	1	67589.399999999973		
47.45	-120.25	12	49670.199999999985		
+-----+-----+-----+-----+-----+					
Fetched 4152 row(s) in 1.60s					
[quickstart.cloudera:21000] > █					

Query to determine locations with highest UPV power generation potential (without combining with landcover data)

6. select latitude, longitude, sum(sum) as power from upv group by latitude, longitude order by power DESC limit 10;

Snapshot for the query:

```
[quickstart.cloudera:21000] > select latitude,longitude, sum(sum) as power from
upv group by latitude,longitude order by power DESC limit 10;
Query: select latitude,longitude, sum(sum) as power from upv group by latitude,l
ongitude order by power DESC limit 10
+-----+-----+-----+
| latitude | longitude | power |
+-----+-----+-----+
| 26.95    | -80.85    | 6239727.599999998 |
| 27.25    | -80.65    | 6030665.399999999 |
| 26.55    | -80.45    | 5954478.899999995 |
| 26.45    | -80.25    | 5935230.399999997 |
| 25.35    | -80.45    | 5903128.900000002 |
| 36.45    | -115.95   | 5601073.699999998 |
| 36.35    | -115.95   | 5564719.799999999 |
| 34.65    | -115.15   | 5464587.300000002 |
| 34.85    | -116.75   | 5429355.299999997 |
| 25.25    | -80.65    | 5323555.799999998 |
+-----+-----+-----+
Fetched 10 row(s) in 1.06s
[quickstart.cloudera:21000] > █
```

Result of Analytic:

latitude	longitude	power	Location
26.95	-80.85	6239727.6	FL, USA
27.25	-80.65	6030665.4	FL, USA
26.55	-80.45	5954478.9	FL, USA
26.45	-80.25	5935230.4	FL, USA
25.35	-80.45	5903128.9	FL, USA
36.45	-115.95	5601073.7	NV, USA
36.35	-115.95	5564719.8	NV, USA
34.65	-115.15	5464587.3	CA, USA
34.85	-116.75	5429355.3	CA, USA
36.95	-115.15	5158296.7	NV, USA
36.35	-119.95	4895116.1	NV, USA

Query to determine locations with highest UPV power generation potential (without combining with landcover data)

7. select latitude, longitude, sum(sum) as power from upv group by latitude, longitude order by power DESC limit 10;

Snapshot for the query:

```
[quickstart.cloudera:21000] > select latitude,longitude, sum(sum) as power from
dpv group by latitude,longitude order by power DESC limit 10;
Query: select latitude,longitude, sum(sum) as power from dpv group by latitude,l
ongitude order by power DESC limit 10
+-----+-----+-----+
| latitude | longitude | power |
+-----+-----+-----+
| 33.35    | -111.85   | 2340599.000000003 |
| 33.25    | -111.85   | 2331590 |
| 34.25    | -118.35   | 2263532.500000002 |
| 34.05    | -118.55   | 2263496.900000001 |
| 34.15    | -118.65   | 2254228.600000002 |
| 34.25    | -118.55   | 2250810.7 |
| 34.05    | -118.45   | 2249373.400000005 |
| 34.15    | -118.55   | 2245489.100000005 |
| 33.55    | -112.25   | 2243462.900000003 |
| 34.15    | -118.35   | 2233602.000000005 |
+-----+-----+-----+
Fetched 10 row(s) in 0.93s
[quickstart.cloudera:21000] > █
```

Result of analytic:

latitude	longitude	power	Location
33.35	-111.85	2340599	AZ, USA
33.25	-111.85	2331590	AZ, USA
34.25	-118.35	2263532.5	CA, USA
34.05	-118.55	2263496.9	CA, USA
34.15	-118.65	2254228.6	CA, USA
34.25	-118.55	2250810.7	CA, USA
34.05	-118.45	2249373.4	CA, USA
34.15	-118.55	2245489.1	CA, USA
33.55	-112.25	2243462.9	AZ, USA
34.15	-118.35	2233602	CA, USA