

A Mini-Project in Big Data Analytics on Google Cloud in Hive

BE COMPS

TITLE:

Group Members: Ajinkya Darshane(16102001)

Aditya Joshi(16102017)

Dataset: *world_war_2_aircrafts*

Attributes : aircraft_type

aircraft_subtype

aircraft_name

year_in_service

country_of_origin

operators

```

$ cd /root/.aws
$ aws s3 cp s3://aws-logs-913611831131-us-east-1/s3/aws-logs-agent-installers/ /usr/share/doc/ --recursive
Linux aws2-aircrafts-m 4.9.0-11-amd64 #1 SMP Debian 4.9.189-3 (2019-09-02) x86_64

The programs included with the Debian GNU/Linux system are free software;
the distribution terms for each program are described in the
individual files in /usr/share/doc/*/copyright.

Debian GNU/Linux comes with ABSOLUTELY NO WARRANTY, to the extent
permitted by applicable law.

root@aws2-aircrafts-m:~# gsu2 ls gsf/wm2_csw_bucket/
gsf/wm2_csw_bucket/world_war_2-aircrafts.csv
root@aws2-aircrafts-m:~# gsu2 ls gsf/wm2_csw_bucket/
gsf/wm2_csw_bucket/world_war_2-aircrafts.csv
root@aws2-aircrafts-m:~# ssh root@aws2-aircrafts-m -i /root/.ssh/jdbc-hive2 //localhost:10000/default -n winlogexpgwm2-aircrafts-m -d org.ap
ache.hive.jdbc.HiveDriver
Connecting to jdbc:hive2://localhost:10000/default
Connected to Apache Hive (version 2.3.5)
Transaction isolation: TRANSACTION_READ_COMMITTED
Debian version 2.3.5 by Apache Hive
root@aws2-aircrafts-m:~# jdbc:hive2://localhost:10000/default; CREATE EXTERNAL TABLE dataset
(
  aircraft_type VARCHAR(25),
  aircraft_subtype VARCHAR(25),
  aircraft_name VARCHAR(35),
  year_in_service HIGINT,
  country_of_origin VARCHAR(25),
  operator VARCHAR(100)
)
ROW FORMATTED DELIMITED FIELDS TERMINATED BY ','
LOCATION 'gsf/wm2_csw_bucket';
root@aws2-aircrafts-m:~#
No rows affected (2.846 seconds)
root@aws2-aircrafts-m:~# jdbc:hive2://localhost:10000/default; DESCRIBE dataset
?
+-----+
| col_name | data_type | comment |
+-----+
| aircraft_type | varchar(25) | |
| aircraft_subtype | varchar(25) | |
| aircraft_name | varchar(35) | |
| year_in_service | higit | |
| country_of_origin | varchar(25) | |
| operator | varchar(100) | |
+-----+
6 rows selected (0.113 seconds)
root@aws2-aircrafts-m:~# jdbc:hive2://localhost:10000/default;

```

Query#2: Display the types of aircrafts along with their count

Command: select distinct aircraft_type, count(*) from dataset group by aircraft_type;

```
1 shcloud.google.com/projects/wendataanalytics/tables/central-1<instance>/aircrafts/author=3000-en_1000projectNumber=196347006470
2
3 Waco CG-3
4 Waco CG-4/Hedrian
5 Waco OUP-7/PT-14 CPTF trainer
6 Watanabe E2W(1)
7 Watanabe E2W-103
8 Weiss We-10 C1v
9 Westland Weekin high altitude inter
10 Westland Whirlwind
11 XP-3
12 Yakovlev UT-1
13 Yakovlev UT-2
14 Yakovlev Yak-5
15 Yakovlev Yak-9
16 Yakovlev Yak-18
17 Yakovlev Yak-2
18 Yokosuka B4Y torpedo bomber
19 Yokosuka D3Y
20
21 -----
22 | aircraft_name |
23 -----
24 | Yokosuka D4Y dive bomber/night figh |
25 | Yokosuka E4X(e) |
26 | Yokosuka B5Y |
27 | Yokosuka K4Y |
28 | Yokosuka K5Y |
29 | Yokosuka L3Y |
30 | Yokosuka M1T0 |
31 | Yokosuka M1T6 |
32 | Yokosuka M1Y7 Ghaa named missile |
33 | Yokosuka M1Y9/Pe-13 |
34 | Yokosuka P1Y |
35 | Yokosuka K2Y |
36 | Zilin Z-111 |
37 | Zmaj Fizir P3 |
38 | Zmaj Fizir P3-2 |
39 | Zmaj K-1 |
40 | Zveno SFB |
41
42 -----
43 | aircraft_name |
44 -----
45 | de Havilland Albatross |
46 | de Havilland Australia OMA-G |
47 | de Havilland DH.60 Moth |
48 | de Havilland DH.82 Queen Bee target |
49 | de Havilland Don |
50 | de Havilland Mosquito PR variants |
51 | de Havilland Mosquito night fighter |
52 | de Havilland Mosquito(4) |
53 | de Havilland Tiger Moth[notes 30] |
54
55 -----
56 527 rows selected (6.951 seconds)
57
58 01 jdbc:hive2://localhost:10000/default> select aircraft_type, count(*) from dataset group by aircraft_type;
59
60 -----
61 | aircraft_type | _c1 |
62 -----
63 | Blimps | 4 |
64 | Bomber and Attack aircraft | 149 |
65 | Fighter aircraft | 147 |
66 | Misc research | 3 |
67 | Prototypes | 282 |
68 | Research | 32 |
69 | Rockets and Drones | 15 |
70 | Rotorcraft | 4 |
71 | Seaplanes | 127 |
72 | Strategic and Photo-recon | 10 |
73 | Trainers | 151 |
74 | Transport and Utility air | 178 |
75 | aircraft_type |
76 -----
77
78 13 rows selected (6.424 seconds)
79
80 01 jdbc:hive2://localhost:10000/default>
```

Query#3: Display the inaugural years for each aircraft types

Command: select aircraft_type, MIN(year_in_service) as inaugural_year from dataset group by aircraft_type;

```
01 jdbc:hive2://localhost:10000/default> select aircraft_type, MIN(year_in_service) as inaugural_year from dataset group by aircraft_type;
2
3 -----
4 | aircraft_type | inaugural_year |
5 -----
6 | Blimps | 1935 |
7 | Bomber and Attack aircraft | 1921 |
8 | Fighter aircraft | 1925 |
9 | Misc research | 1937 |
10 | Prototypes | 1933 |
11 | Research | 1937 |
12 | Rockets and Drones | 1935 |
13 | Rotorcraft | 1934 |
14 | Seaplanes | 1922 |
15 | Strategic and Photo-recon | 1952 |
16 | Trainers | 1913 |
17 | Transport and Utility air | 1925 |
18 | aircraft_type | NULL |
19 -----
20
21 13 rows selected (6.216 seconds)
```

Command: select aircraft_type, MAX(year_in_service) as final_year from dataset group by aircraft_type;

```

13 rows selected (6.216 seconds)
0: jdbc:hive2://localhost:10000/default: select aircraft_type, MAX(year_in_service) as final_year from dataset group by aircraft_type:
+-----+-----+
| aircraft_type | final_year |
+-----+-----+
| Bombs         | 1944       |
| Bomber and Attack aircraft | 1945       |
| Fighter aircraft | 1945       |
| Misc research  | 1944       |
| Prototypes    | 1945       |
| Research      | 1945       |
| Rockets and Drones | 1945       |
| Autocraft    | 1945       |
| Seaplanes     | 1945       |
| Strategic and Photo-recon | 1945       |
| Trainers      | 1944       |
| Transport and Utility air | 1945       |
| aircraft_type | NULL       |
+-----+-----+

13 rows selected (5.941 seconds)
0: jdbc:hive2://localhost:10000/default:

```

Command: SELECT aircraft_type from dataset where country_of_origin = 'UK';

[illegible]

```

[+] https://github.com/pentestit/elasticsearch-elasticsearch-central-2-elasticsearch-wed-aircraft-milauthor-08m-en-us@google-number-19694769947
2020
Prototypes |
Prototypes |
Prototypes |
Prototypes |
Prototypes |
Prototypes |
Prototypes |
Prototypes |
Prototypes |
Prototypes |
Prototypes |
Prototypes |
Prototypes |
Prototypes |
Prototypes |
Prototypes |
Prototypes |
Prototypes |
aircraft_type |
Prototypes |
Prototypes |
Prototypes |
Prototypes |
Prototypes |
Prototypes |
Prototypes |
Prototypes |
Prototypes |
Prototypes |
Prototypes |
Prototypes |
Research |
Research |
Research |
Research |
Research |
Research |
Research |
Research |
22 rows selected (0.15 seconds)
+-----+
| elasticsearch/localhost:1000/default |

```

Query#6: Display the duration of service of each aircraft types

Command: SELECT aircraft_type, (MAX(year_in_service) - MIN(year_in_service)) as Duration_of_Service FROM dataset group by aircraft_type;

```

-- Now retrieve (1/3rd seconds)
jdbchive1://localhost:10000/default?SELECT aircraft_type, (MAX(year_in_service) - MIN(year_in_service)) as Duration of Serv
FROM dataset group by aircraft_type)

```

aircraft_type	Duration of service
Slings	9
Bomber and Attack aircraft	24
Fighter aircraft	20
Misc research	7
Prototypes	12
Research	8
Rockets and Drones	10
Helicopters	11
Seaplanes	23
Strategic and Photo-recon	13
Trainers	31
Transport and Utility air	20
aircraft_type	NULL

```

-- rows selected (0.33 seconds)
jdbchive1://localhost:10000/default?

```

Query#6: Order aircraft by country

Command: SELECT aircraft_type, country_of_origin FROM dataset ORDER BY country_of_origin;

Strategic and Photo-recon	USSR	
Prototypes	USSR	
Prototypes	USSR	
Bomber and Attack aircraft	USSR	
Transport and Utility air	USSR	
Seaplanes	USSR	
Bomber and Attack aircraft	USSR	
Prototypes	USSR	
Prototypes	USSR	
Trainers	USSR	
Transport and Utility air	USSR	
Prototypes	USSR	
Prototypes	USSR	
Fighter aircraft	USSR	
Bomber and Attack aircraft	USSR	
Fighter aircraft	USSR	
Fighter aircraft	USSR	
Prototypes	USSR	
Prototypes	USSR	
Bomber and Attack aircraft	USSR	
Prototypes	USSR	
Prototypes	USSR	
Trainers	USSR	
Prototypes	VF I-16	
Prototypes	VL Rome	
Prototypes	VL Pyörremyrsky	
Prototypes	XP-1	
Trainers	Yugoslavia	
Prototypes	Yugoslavia	
Trainers	Yugoslavia	
Trainers	Yugoslavia	
Seaplanes	Yugoslavia	
Trainers	Yugoslavia	
Prototypes	Yugoslavia	
Prototypes	Yugoslavia	
Trainers	Yugoslavia	
Seaplanes	Yugoslavia	
Seaplanes	Yugoslavia	
Seaplanes	Yugoslavia	
aircraft type	country of origin	
Prototypes	de Havilland Vampire	

1,177 rows selected (4.309 seconds)

1: jdbc:hive2://localhost:10000/default>