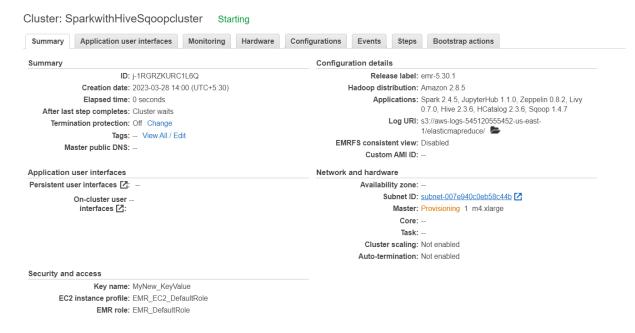# Capstone: Instant Health Alert System – Final Submission

A script to create an external hive table for patients' vital information and view data
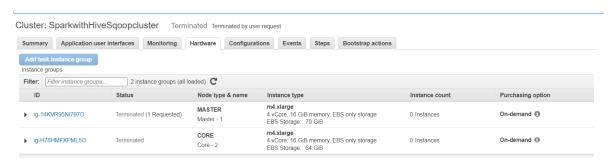
Patients_Vital_Info table in Hive

```
CREATE EXTERNAL TABLE IF NOT EXISTS Patients_Vital_Info (

    CustomerID int,

    BP int,

    HeartBeat int,

    Message_time timestamp

)

STORED AS PARQUET

LOCATION '/user/livy/output/date=2022-03-16';
```

Set up for the Hive Cluster

Cluster: SparkwithHiveSqoopcluster   Starting

| Summary | Application user interfaces | Monitoring | Hardware | Configurations | Events | Steps | Bootstrap actions |

Summary

ID: j-1RGRZKURC1L6Q
Creation date: 2023-03-28 14:00 (UTC+5:30)
Elapsed time: 0 seconds
After last step completes: Cluster waits
Termination protection: Off Change
Tags: -- View All / Edit
Master public DNS: --

Configuration details

Release label: emr-5.30.1
Hadoop distribution: Amazon 2.8.5
Applications: Spark 2.4.5, JupyterHub 1.1.0, Zeppelin 0.8.2, Livy 0.7.0, Hive 2.3.6, HCatalog 2.3.6, Sqoop 1.4.7
Log URI: s3://aws-logs-545120555452-us-east-1/elasticmapreduce/
EMRFS consistent view: Disabled
Custom AMI ID: --

Application user interfaces

Persistent user interfaces: --
On-cluster user interfaces: --

Network and hardware

Availability zone: --
Subnet ID: subnet-007e940c0eb58c44b
Master: Provisioning 1 m4.xlarge
Core: --
Task: --
Cluster scaling: Not enabled
Auto-termination: Not enabled

Security and access

Key name: MyNew_KeyValue
EC2 instance profile: EMR_EC2_DefaultRole
EMR role: EMR_DefaultRole

Screenshot of EMR Hardware Configuration *(with 1 master node and 2 core nodes)*

Cluster: SparkwithHiveSqoopcluster   Terminated Terminated by user request

| Summary | Application user interfaces | Monitoring | Hardware | Configurations | Events | Steps | Bootstrap actions |

Add task instance group

Instance groups

Filter: Filter instance groups ...   2 instance groups (all loaded)

ID	Status	Node type & name	Instance type	Instance count	Purchasing option
▶ ig-14KVR95NI797O	Terminated (1 Requested)	MASTER Master - 1	m4.xlarge 4 vCore, 16 GiB memory, EBS only storage EBS Storage: 70 GiB	0 Instances	On-demand ⓘ
▶ ig-H78HMFXPML5O	Terminated	CORE Core - 2	m4.xlarge 4 vCore, 16 GiB memory, EBS only storage EBS Storage: 64 GiB	0 Instances	On-demand ⓘ

In the Hive Shell

- Open Hive shell.

```
^C[hadoop@ip-172-31-82-68 ~]$ hive

Logging initialized using configuration in file:/etc/hive/conf.dist/hive-log4j2.
properties Async: true
```

- Create a database *patient_health_care*

```
create database if not exists patient_health_care;
```

```
hive> create database if not exists patient_health_care;
OK
Time taken: 0.855 seconds
```

- Use database patient_health_care

```
use patient_health_care;
```

```
hive> use patient_health_care;
OK
Time taken: 0.046 seconds
```

- Create external table named *Patients_Vital_Info*

```
CREATE EXTERNAL TABLE IF NOT EXISTS Patients_Vital_Info (
    CustomerID int,
    BP int,
    HeartBeat int,
    Message_time timestamp
)
STORED AS PARQUET
LOCATION '/user/livy/output/date=2022-03-16';
```

```
hive>
    >
    > CREATE EXTERNAL TABLE IF NOT EXISTS Patients_Vital_Info (
    >     CustomerID int,
    >     BP int,
    >     HeartBeat int,
    >     Message_time timestamp
    > )
    > STORED AS PARQUET
    > LOCATION '/user/livy/output/date=2022-03-16';
OK
Time taken: 0.333 seconds
```