## Расположение и описание тестового набора данных

1. UFC-Fight historical data from 1993 to 2019

<https://www.kaggle.com/rajeevw/ufcdata#data.csv>

/test\_datasets/ufcdata

3. Electric Motor Temperature

<https://www.kaggle.com/wkirgsn/electric-motor-temperature>

/test\_datasets/electric-motor-temperature

## Справочник по HiveQL

<https://cwiki.apache.org/confluence/display/Hive/LanguageManual>

DDL: <https://cwiki.apache.org/confluence/display/Hive/LanguageManual+DDL>

DML: <https://cwiki.apache.org/confluence/display/Hive/LanguageManual+DML>

SELECT: <https://cwiki.apache.org/confluence/display/Hive/LanguageManual+Select>

## Пример документации по SerDe

<https://cwiki.apache.org/confluence/display/Hive/CSV+Serde>

## Что сделать

1. Создание базы данных

create database hive\_db

1. Создание таблицы внутри базы данных и загрузка в нее данных

create table hive\_db.motor\_temperature

(

ambient float,

coolant float,

u\_dVoltage float,

u\_qVoltage float,

motor\_speed float,

torque float,

i\_d float,

i\_q float,

pm float,

stator\_yoke float,

stator\_tooth float,

stator\_winding float,

profile\_id int

)

ROW FORMAT SERDE

'org.apache.hadoop.hive.serde2.OpenCSVSerde'

STORED AS INPUTFORMAT

'org.apache.hadoop.mapred.TextInputFormat'

OUTPUTFORMAT

'org.apache.hadoop.hive.ql.io.HiveIgnoreKeyTextOutputFormat'

TBLPROPERTIES (

'serialization.null.format' = '',

'skip.header.line.count' = '1')

;

LOAD DATA INPATH '/user/dekan/pmsm\_temperature\_data.csv' INTO TABLE hive\_db.motor\_temperature

;

1. Посмотрим что внутри теперь

select \* from hive\_db.motor\_temperature

;

select count(\*), profile\_id

from hive\_db.motor\_temperature

group by profile\_id

;

1. Посмотрим как в действительности была создана таблица

show create table hive\_db.motor\_temperature

1. Посмотрим что там внутри

hdfs dfs –ls /user/hive/warehouse/hive\_db.db/motor\_temperature

1. Создание обычной таблицы внутри базы данных

drop table hive\_db.ufc\_fighter\_details

;

create table hive\_db.ufc\_fighter\_details

(

fighter\_name string,

Height string,

Weight string,

Reach string,

Stance string,

DOB string

)

ROW FORMAT SERDE

'org.apache.hadoop.hive.serde2.OpenCSVSerde'

LOCATION

'/ufctable/'

;

set mapred.job.queue.name=root.default

;

select \* from hive\_db.ufc\_fighter\_details

;

select count(\*) from hive\_db.ufc\_fighter\_details

;

1. Создание external таблицы в базе данных

create EXTERNAL table hive\_db.ufc\_fighter\_details

…….

1. Добавление данных

msck repair table hive\_db.ufc\_fighter\_details

;

1. Удаление обычной таблицы
2. Удаление external таблицы
3. Удаление базы